Non-equilibrium Statistical Mechanics Studied Through the Lens of Stochastic Calculus

Dissertation
for the award of the degree
"Doctor rerum naturalium" (Dr. rer. nat)

of the Georg-August-Universität Göttingen
within the doctoral program
Physics of Biological and Complex Systems
of the Georg-August University School of Science

submitted by

Cai Vinzent Dieball

from Hildesheim

Göttingen, 2023

Thesis Advisory Committee

First referee and supervisor

Dr. Aljaž Godec
Max-Planck-Institut für Multidisziplinäre Naturwissenschaften, Göttingen

Second referee

Prof. Dr. Matthias Krüger
Institut für Theoretische Physik, Georg-August-Universität Göttingen

Prof. Dr. Peter Sollich
Institut für Theoretische Physik, Georg-August-Universität Göttingen

Examination Board

Dr. Aljaž Godec

Prof. Dr. Matthias Krüger

Prof. Dr. Peter Sollich

Prof. Dr. Ramin Golestanian
Max-Planck-Institut für Dynamik und Selbstorganisation, Göttingen

Prof. Dr. Timo Betz
Drittes Physikalisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Stefan Klumpp
Institut für Dynamik komplexer Systeme, Georg-August-Universität Göttingen

Date of oral examination: November 17, 2023

Acknowledgements

First, I deeply thank my supervisor Aljaž. I appreciate your personal and humorous style of supervision, and your striving for a fundamental understanding of physical processes. During the last four years, I have learned a lot from you, not only about physics, mathematics and scientific writing.

Second, I would like to thank all my colleagues, and former colleagues, in the *Mathematical Physics Group*—David, Alessio, Max, Kristian, Rick, Dima and Wolfgang, former Master's students Janik, Gerrit, Tassilo, and recently joined PhD students Felix, Francesco and Felipe. I always enjoyed how everyone is constantly open for scientific and non-scientific discussions, a coffee break and a good (or bad) joke, and I was always happy to be in this group. In particular, I thank Rick and Kristian for proofreading very large parts of this thesis.

Outside our group, I also want to thank many people in the *Theoretical and Computational Biophysics* department. In particular, Annke, Lars, Sara, Florian, and many others for great lunch breaks, and everyone who keeps the department running so well. I very much thank Frauke, Antje, Evi and Steffi for always being extremely helpful and pragmatic.

Beyond the department, I thank Matthias Krüger and Peter Sollich for giving input and asking critical questions from start to end of my PhD. In particular, I thank Matthias Krüger for writing the report on this thesis. I also thank Miguel, Raúl, Antonio, Gerrit, Matthias Weiss and Diego for great collaborations and co-authoring publications included in this thesis.

Last but not least, I deeply thank my family, especially my parents, friends, and Imke's family, for supporting me throughout the years. Clara, thank you for being you and being there. And most importantly, Imke, thank you so much for everything you do for me, and everything you mean to me.

Abstract

The microscopic world of atoms and molecules is continuously in motion, which is manifested on the macroscopic scale as a non-zero *temperature*. Due to collisions with surrounding molecules that perform thermal motion, many small systems undergo erratic, random motion, known as *Brownian motion*. In the presence of external forces, this motion is described by the *Langevin equation*, which forms a basis for studying thermodynamics of stochastic systems, and proved successful in biology, physics, chemistry and beyond. In particular, unlike classical thermodynamics, the Langevin description applies to systems far from thermal equilibrium, which is, e.g., vital to approach biological processes in living systems.

In this thesis, we investigate properties of stochastic dynamics by applying the mathematical framework of stochastic calculus, that directly addresses the notion of fluctuating trajectories. The major part of the thesis is concerned with fluctuations of time-averaged observables. We derive expressions for the variance and correlations of empirical densities and currents, which are observables central to statistical mechanics on the level of individual trajectories. Based on these results, we uncover the essential role of coarse graining in space, and explain the results by connecting them to a generalized time-reversal symmetry in non-equilibrium steady states. In particular, we use the stochastic-calculus approach to rederive an inequality called *thermodynamic uncertainty relation*, that allows to infer entropy production from measurements of time-averaged currents. Based on our simplified derivation, we are able to systematically determine when this inequality becomes saturated, i.e., when the thermodynamic inference is optimal.

Another large part of the thesis investigates how systems prepared at different temperatures relax to the temperature of the environment. While the classical laws of thermodynamics only apply close to thermal equilibrium and predict a relaxation that is indifferent with respect to the sign of the temperature change, it was recently found that for large temperature differences, i.e., far from equilibrium, thermal relaxation becomes asymmetric, and heating is faster than cooling. In a collaboration, we experimentally confirm this asymmetry, generalize this framework, and extend the study to relaxation towards non-equilibrium steady states.

Finally, we also address motion beyond Langevin dynamics, namely *two-state dynamics* that switches stochastically between different states of motion. We make predictions for experimentally relevant observables for diverse phys-

ical models, and apply the results to particle tracking data recorded in living cells.

Altogether, in this thesis, we study stochastic dynamics of systems far from thermal equilibrium. We push forward the framework of stochastic calculus to handle and understand randomly fluctuating trajectories, and to deduce thermodynamic insights and properties of experimentally relevant observables from this intrinsically stochastic dynamics.

Contents

1 Introduction **1**

1.1 Brownian motion 1
1.2 Physical approach: The Langevin equation 2
1.3 Mathematical treatment: Stochastic calculus 3
1.4 Equilibrium and non-equilibrium systems 5
1.5 Non-equilibrium steady-states 6
1.6 Time-integrated observables 7
1.7 A brief account of stochastic thermodynamics 7
1.8 Thermodynamic uncertainty relations 9
1.9 Selected further topics 10
1.10 Outline of the thesis 10

2 Fundamentals **13**

2.1 Probability measures and integration 13
2.2 Probability theory based on probability densities 18
2.3 Stochastic processes 23
2.4 Mathematical definition of Brownian motion 25
2.5 Stochastic calculus 26
2.6 The Fokker-Planck equation 35
2.7 Equilibrium and non-equilibrium steady states 38
2.8 Multiplicative noise 44
2.9 Stochastic entropy 46

3 Statistics of time-integrated densities and currents **51**

3.1 Definition of time-integrated densities and currents 51
3.2 Technical background and state of the art 53
3.3 Summary of results 60

3.4 Mathematical, thermodynamical, and experimental necessity for coarse graining empirical densities and currents in continuous space (Phys. Rev. Lett. 2022) . . . 61
3.5 Coarse graining empirical densities and currents in continuous-space steady states (Phys. Rev. Res. 2022) . . . 80
3.6 On correlations and fluctuations of time-averaged densities and currents with general time dependence (J. Phys. A 2022) . . . 126
3.7 Feynman-Kac theory of time-integrated functionals: Itô versus functional calculus (J. Phys. A 2023) . . . 139

4 Direct route to thermodynamic uncertainty relations and their saturation 153
4.1 Introduction to TURs . . . 153
4.2 Technical background . . . 159
4.3 State of the art . . . 160
4.4 Summary of results . . . 162
4.5 Direct route to thermodynamic uncertainty relations and their saturation (Phys. Rev. Lett. 2023) . . . 163

5 Thermal relaxation asymmetry 189
5.1 Introduction to thermal relaxation . . . 189
5.2 Quantifying thermal relaxation . . . 190
5.3 Summary of results . . . 194
5.4 Heating and cooling are fundamentally asymmetric and evolve along distinct pathways (Nat. Phys. 2024) . . . 195
5.5 Asymmetric thermal relaxation in driven systems: Rotations go opposite ways (Phys. Rev. Res. 2023) . . . 231

6 Scattering fingerprints of two-state dynamics 255
6.1 Introduction to two-state dynamics and anomalous diffusion . . . 255
6.2 Summary of results . . . 257
6.3 Scattering fingerprints of two-state dynamics (New J. Phys. 2022) . . . 257

7 Conclusion 287
7.1 Time-integrated observables and necessity of coarse graining . . . 287
7.2 Thermodynamic uncertainty relations . . . 290
7.3 Thermal relaxation far from equilibrium . . . 291
7.4 Two-state dynamics . . . 292
7.5 Final remarks . . . 293

Bibliography **295**

Chapter 1

Introduction

1.1 Brownian motion

A nanometer- to micrometer-sized particle immersed in water undergoes erratic motion in random directions, a phenomenon known as *Brownian motion*. This was originally observed under the microscope by Robert Brown [1], and theoretically explained in the seminal works of Albert Einstein [2] and Marian Smoluchowski [3]. Brownian motion was explained to be a consequence of collisions with water molecules that undergo random thermal motion. The explanation was thus based on the existence of atoms, which was still under heavy debate in the early 20th century. Combining observations with theory, one was able to see manifestations of the existence of atoms under a microscope. This striking insight, together with the quantitative experimental confirmation of Einstein's theory by Jean Perrin, finally settled the debate about the existence of atoms, and Jean Perrin was awarded with the Nobel Prize for Physics "for his work on the discontinuous structure of matter" [4, 5]. The establishment of the existence of atoms laid the foundation for a whole new area of physics, namely *statistical mechanics*, a general framework for understanding and predicting the behavior of systems consisting of many particles, seeking to bridge the gap between the microscopic description in terms of atoms and molecules, and the macroscopic world of observable phenomena. In particular, Einstein's approach paved the way for probabilistic descriptions of classical physical systems that proved successful in biology, physics, chemistry and beyond [6, 7].

1.2 Physical approach: The Langevin equation

For the scope of this thesis, we consider a *system* to be a set of coordinates coupled to a Markovian, i.e., memoryless, *heat bath*. The heat bath is the medium, e.g., water, and heat transfer occurs via a random force that originates from the probabilistically treated collisions with the surrounding molecules. Since the systems considered here are small, but at the same time large compared to the microscopic scale of the molecules in the bath, we say that such systems evolve on *mesoscopic* scales. A central equation for describing the stochastic dynamics of such systems, e.g., the position $x(t)$ of a Brownian particle at time t, on time and length scales between the scale of individual molecules and the macroscopic world, is the Langevin equation [6, 8]. In the case of a Brownian particle with mass m it reads

$$m\ddot{x}(t) = -\gamma\dot{x}(t) + F_{\text{external}}(x(t), t) + F_{\text{random}}(t)\,. \tag{1.2.1}$$

For simplicity we use a one-dimensional notation throughout the introduction. Equation (1.2.1) describes the acceleration $\ddot{x}$ due to a friction force with Stokes coefficient γ, external forces (e.g., electrical forces or chemical gradients), and the random force representing thermal fluctuations. The random influence encompasses the collisions with water molecules, that are treated probabilistically. A way to relate this stochastic description to the fundamental, deterministic Newtonian dynamics is given, e.g., by the Mori-Zwanzig projection formalism [9, 10]. Thereby, one rewrites the equations of motion and projects them onto the slow mesoscopic dynamics while the remaining motions are assumed to occur and relax much faster, such that they may be treated statistically.

A further simplification of the dynamics in Eq. (1.2.1), valid in particular for small biological systems, is that the friction leads to an effectively instantaneous relaxation of the momentum of the mesoscopic system. This approximation is typically[1] justified on timescales larger than the collision timescale m/γ, e.g., for particles of comparably low mass or highly viscous media. Taking this *overdamped* limit, the Langevin equation becomes [6, 7]

$$\gamma\dot{x}(t) = F_{\text{external}}(x(t), t) + F_{\text{random}}(t)\,. \tag{1.2.2}$$

In the absence of the random force, the motion $x(t)$ would simply follow from specifying initial conditions and solving the differential equation $\gamma\dot{x}(t) = F_{\text{external}}(x(t), t)$. However, the random force leads to a much richer phenomenology.

[1] In general, taking this limit is more complicated and may depend on further details of the dynamics, see, e.g., Ref. [11].

To understand the stochastic dynamics dictated by the overdamped Langevin equation (1.2.2), we turn our attention to the random force. Crucially, its statistical properties do not depend on the precise details of the unobserved particles, i.e., its probability distribution does not depend on the initial positions and momenta of the water molecules, assuming that these initial conditions are reasonably mixed. There is *a priori* no preferred direction of the impacts of the water molecules, i.e., they occur randomly from all directions, and therefore the average of the random force vanishes, denoted by $\langle F_{\mathrm{random}}(t)\rangle = 0$. In a medium like water, the molecular motions decorrelate much faster than the observed timescales of the mesoscopic system. For the random force, this implies that correlations at times differing by more than this short decorrelation time vanish, i.e., as a mathematical idealization $\langle F_{\mathrm{random}}(t)F_{\mathrm{random}}(t')\rangle = c\delta(t-t')$ with $c > 0$. Since the random force is composed of many, effectively uncorrelated, bounded contributions (the individual impacts of water molecules), the central limit theorem implies that it will converge to a Gaussian random distribution. The latter is completely specified by its mean and variance, such that the above properties determine the random force up to the constant prefactor c.

1.3 Mathematical treatment: Stochastic calculus

Based on these arguments, a mathematical formalization of Brownian motion was established, known as the Wiener process W_t [12]. It corresponds to Eq. (1.2.2) in the absence of external forces and for $c = \gamma = 1$, i.e., $\mathrm{d}W_t/\mathrm{d}t = F_{\mathrm{random}}(t)$. Formulating the properties of the random force in terms of the increments $W_{t+\mathrm{d}t} - W_t = \mathrm{d}W_t = F_{\mathrm{random}}(t)\mathrm{d}t$ for $\mathrm{d}t \to 0$, we note that $\mathrm{d}W_t$ follows a zero-mean Gaussian distribution with variance[2] $\langle \mathrm{d}W_t^2\rangle = \mathrm{d}t$, and increments are uncorrelated at unequal time $\langle \mathrm{d}W_t \mathrm{d}W_{t'}\rangle = 0$ for $t \neq t'$. Writing $W_t - W_0$ as the sum of these Gaussian uncorrelated increments, it follows that the Brownian motion, or Wiener process, $W_t - W_0$ obeys a zero-mean Gaussian distribution with variance t, i.e., even though the average position $\langle W_t\rangle$ remains unchanged, the Brownian motion leads to displacements along individual trajectories that increase over time. Considering the probability density $p(x,t)$ of $W_t = x$, or the density field of many non-interacting Brownian particles, the average over many random motions leads to deterministic *diffusion* in time, $\partial_t p(x,t) = \partial_x^2 p(x,t)/2$, a phenomenon, historically known much before Brownian motion [13].

A particular mathematical intricacy of Brownian motion that formally forbids to write Eqs. (1.2.1) and (1.2.2) becomes apparent directly from the variance of its increments, $\langle \mathrm{d}W_t^2\rangle = \mathrm{d}t$, which implies that $|\mathrm{d}W_t/\mathrm{d}t| \sim 1/\sqrt{\mathrm{d}t} \to \infty$ with probability one for $\mathrm{d}t \to 0$.

[2] For $\mathrm{d}t, \mathrm{d}t' \to 0$ this corresponds to $\langle \mathrm{d}W_t \mathrm{d}W_{t'}\rangle = \delta(t-t')\mathrm{d}t\mathrm{d}t'$ for $t = t'$, see Subsec. 2.5.2.

This means that Brownian motion is not differentiable in t, which appears problematic since the random force was characterized exactly as the derivative of Brownian motion, see Eq. (1.2.2) in the absence of external forces. This irregularity, together with the fluctuating nature of the process, demands a generalization of classical calculus.

Such a generalization was developed in *stochastic calculus* [6, 14–16] in the last century. Starting from the Brownian motion W_t, this field provides a framework for the analysis of *stochastic differential equations* such as the Langevin equation (1.2.2), in the mathematics literature written in a way that avoids the divergent term $\mathrm{d}W_t/\mathrm{d}t$ as

$$\mathrm{d}x_t = a(x_t, t)\mathrm{d}t + b(x_t, t)\mathrm{d}W_t\,. \tag{1.3.1}$$

To give a brief overview, the main practical difference between stochastic and classical calculus is that due to $\langle \mathrm{d}W_t^2 \rangle = \mathrm{d}t$, quadratic orders of $\mathrm{d}x_t$ are not negligible compared to $\mathrm{d}t$. Therefore, the classical chain rule $\mathrm{d}f(x) = f'(x)\mathrm{d}x$ is replaced by *Itô's Lemma* $\mathrm{d}f(x_t) = f'(x_t)\mathrm{d}x_t + f''(x_t)\mathrm{d}x_t^2/2$. Also, one can define *stochastic integrals*, e.g., $\int_{\tau=0}^{\tau=t} f(x_\tau)\mathrm{d}x_\tau$, but unlike in classical calculus, the result will depend on whether $f(x_\tau)\mathrm{d}x_\tau$ is interpreted as $f(x_\tau)(x_{\tau+\mathrm{d}\tau} - x_\tau)$ (*Itô*) or, e.g., $f((x_\tau + x_{\tau+\mathrm{d}\tau})/2)(x_{\tau+\mathrm{d}\tau} - x_\tau)$ (*Stratonovich*). Apart from stochastic calculus, the stochastic motion in Eqs. (1.2.2) and (1.3.1) may also be studied by *Fokker-Planck equations* (such as the diffusion equation for Brownian motion) or path integral techniques [6, 16, 17]. However, the approach that addresses the fluctuating trajectories most directly is the stochastic calculus. In this sense, stochastic calculus is the most natural approach, and a consistent way to avoid unclear interpretations of the complex dynamics.

This stochastic dynamics with an *unusual* calculus may at first appear quite specific. However, it is surprisingly widespread, largely due to the universality of the central limit theorem for dynamics influenced by many small contributions. We introduced it for the example of a Brownian particle, but actually many soft matter and biological systems, like macromolecules, viruses and individual cells, are on the relevant *mesoscopic* scale[3] governed by the Langevin equation. Apart from biological systems, the Langevin equation and stochastic calculus have been applied in chemistry [18], finance [19], fluctuations in electrical circuits [20] and a variety of other topics.

Coming back to thermodynamic systems, such as the Brownian particle, there is another general property of the Langevin equation (1.2.2). Intuitively, the friction and the random

[3] These are on the nanometer to micrometer scale, such that they are small enough to feel the random force but large enough to be subject to many, effectively uncorrelated collisions on the relevant timescale.

force are related, since both originate from collisions with the molecules of the surrounding medium. This fixes the diffusion coefficient, i.e., the prefactor of the variance of the random force, which is known as the *Einstein relation*, and was first derived by means of osmotic pressure and diffusive currents [2]. It also follows from requiring consistency with thermodynamical concepts such as the equipartition theorem for the kinetic energy [8], linear response theory [21–23], or Boltzmann statistics (see Sec. 2.7). Using this Einstein relation and the Wiener process to write the random force explicitly, the Langevin equation (1.2.2) in the notation of Eq. (1.3.1) reads

$$\mathrm{d}x_t = \gamma^{-1} F_{\mathrm{external}}(x_t, t)\mathrm{d}t + \sqrt{2k_{\mathrm{B}}T\gamma^{-1}}\mathrm{d}W_t\,, \tag{1.3.2}$$

where k_{B} is the Boltzmann constant and T the temperature of the surrounding medium.

1.4 Equilibrium and non-equilibrium systems

If the external force stems from a confining potential energy function $\phi(x)$ such that $F_{\mathrm{external}}(x) = -\partial_x \phi(x)$, the Langevin dynamics will approach thermal equilibrium for long times, where the probability density of x_t will approach the invariant Boltzmann steady-state density, $p(x,t) \overset{t\to\infty}{\longrightarrow} p_{\mathrm{s}}(x) \propto \exp(-\phi(x)/k_{\mathrm{B}}T)$ [6]. Thermal equilibrium is characterized by a vanishing entropy production and the absence of probability currents. For the considered systems, thermal equilibrium is equivalent to *time-reversal symmetry*, i.e., to the statement that the overdamped Langevin trajectories $(x_\tau)_{0\leq\tau\leq t}$ have the same probabilistic properties as the time-reversed trajectories $(x_{t-\tau})_{0\leq\tau\leq t}$ [6].

Systems in thermal equilibrium are generally much better understood than non-equilibrium systems. They are known to approach Boltzmann equilibrium determined by the potential energy $\phi(x)$, on average they neither produce entropy nor dissipate energy into the heat bath, and they can be fully described in terms of thermodynamic potentials such as the Helmholtz free energy. Also, small deviations from equilibrium are very well understood, in particular within *linear response theory* [21, 24] and *linear irreversible thermodynamics* [25–27].

All these theories cease to hold for systems and processes far from thermal equilibrium. Compared to the (close-to-)equilibrium case, very little is known about non-equilibrium processes. Except for some recent developments (see, e.g., Secs. 1.7 and 1.8), there are no general laws that describe non-equilibrium dynamics, and systems usually have to be approached on a case-by-case basis, solving the detailed equations of motion. This is not

only difficult in theory, but very often also impossible in practice, since there may be many dissipative degrees of freedom, and usually not all of them can be observed.

Strikingly, the description via the Langevin equation (1.3.2) allows to study systems arbitrarily far from equilibrium, i.e., the description holds true for large values of the entropy production, as long as the underlying assumption that dynamics in the heat bath relax sufficiently faster than the dynamics of the system still holds. There are several ways in which the Langevin equation can be used to describe non-equilibrium systems. First, there are *non-equilibrium steady states* that, like equilibrium systems, feature an invariant probability density $p_{\mathrm{s}}(x)$, but unlike equilibrium systems possess non-zero probability currents and a non-vanishing entropy production breaking time-reversal symmetry. Moreover, there are *transient* processes that start from a probability density $p(x, t = 0)$ different from $p_{\mathrm{s}}(x)$, and approach equilibrium or non-equilibrium steady states for long times. For finite times, such systems are always out of equilibrium, and for initial conditions markedly differing from p_{s}, possibly even very far from equilibrium. Allowing for explicit time dependence in the parameters of the Langevin equation, we can also study periodic steady states, or even Langevin dynamics with completely arbitrary time dependence with or without confining potentials.

1.5 Non-equilibrium steady-states

We now focus on non-equilibrium steady states (NESS) [6, 28] as a particular class of processes that may operate far from thermal equilibrium. To this end, consider confining, time-homogeneous (i.e., time-independent) external forces that do not need to simply stem from a confining potential as in the equilibrium case. As for the equilibrium Langevin dynamics mentioned above, $p(x, t)$ approaches a steady-state density $p_{\mathrm{s}}(x)$, but there will also be a non-vanishing steady-state current $j_{\mathrm{s}}(x) \neq 0$ characterizing ongoing flows of probability that leave $p_{\mathrm{s}}(x)$ invariant[4]. Note that for a given temperature, there are infinitely many different NESS (different possible choices of $j_{\mathrm{s}}(x)$) sharing the same invariant density $p_{\mathrm{s}}(x)$ of a single equilibrium state ($j_{\mathrm{s}}(x) = 0$ for all x). In addition to non-zero currents, NESS also have a strictly positive entropy production. NESS are particularly relevant for biological systems, since living systems are inherently out of equilibrium. A paradigmatic example are molecular motors [28–32].

An important concept for steady-state dynamics is *ergodicity*. For the considered steady-state Langevin dynamics, this implies that the average over many independent copies of the

[4] We usually consider multidimensional NESS since in one-dimensional space the conditions for NESS can only be met for periodic boundary conditions.

dynamics sampled from the steady-state density $p_{\mathrm{s}}(x)$, i.e., the *ensemble average* $\langle f(x_t)\rangle_{\mathrm{s}}$ for any $t \geq 0$ and well-behaved function f, agrees with time-averages over long times, i.e., $\langle f(x_t)\rangle_{\mathrm{s}} = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(x_\tau)\mathrm{d}\tau$ [16]. Note that the right-hand side of the equation is a random variable but equality holds with probability one.

1.6 Time-integrated observables

Given the measurement of a very long trajectory, or several short trajectories, sampled from the steady state, we may attempt to infer the steady-state density $p_{\mathrm{s}}(x)$ and current $j_{\mathrm{s}}(x)$. A central object for such inference are *time-integrated*, or more precisely *time-averaged*, densities and currents [33, 34]

$$\begin{aligned} \rho_t &\equiv \frac{1}{t}\int_0^t f(x_\tau)\mathrm{d}\tau\,, \\ J_t &\equiv \frac{1}{t}\int_{\tau=0}^{\tau=t} f(x_\tau)\circ \mathrm{d}x_\tau\,, \end{aligned} \tag{1.6.1}$$

for some function f. The $\circ$-symbol denotes the Stratonovich convention $f(x_\tau)\circ \mathrm{d}x_\tau = f((x_\tau + x_{\tau+\mathrm{d}\tau})/2)(x_{\tau+\mathrm{d}\tau} - x_\tau)$ required to completely define this integral within stochastic calculus. For $f(y) \to \delta(x-y)$, in case of long times or by averaging over many measurements, ρ_t and J_t tend to the steady-state density $p_{\mathrm{s}}(x)$ and current $j_{\mathrm{s}}(x)$, respectively. Since small stochastic systems are subject to random fluctuations, and the number and duration of measurements is restricted, an important part of the analysis of the observables in Eq. (1.6.1) is understanding their fluctuations. Since ρ_t and J_t are functionals of the stochastic dynamics $x_\tau, 0 \leq \tau \leq t$ for finite times t, these functionals have complicated statistical properties. The time-integrated density $t\rho_t$, also known as local time or occupation time, was studied in different scenarios [33, 35–41], often using Feynman-Kac theory [33, 42]. Time-integrated currents are mathematically even more challenging than densities. Together, ρ_t and J_t, also known as *empirical density and current* or *dynamical functionals*, are addressed in the literature mainly in the large deviation limit (large t limit) [34, 43–46] or in terms of inequalities [47–49].

1.7 A brief account of stochastic thermodynamics

We here give a brief introduction to stochastic thermodynamics, focusing on the topics that are relevant for this thesis. As mentioned above, the Langevin equation (1.3.2) can describe systems that are far from equilibrium, e.g., a Brownian particle driven by large external forces. Therefore, it provides a very promising framework to generalize thermodynami-

cal concepts to fluctuating trajectories and systems that are arbitrarily far from thermal equilibrium[5]. This instigated the recent development of the field of *stochastic thermodynamics* [28,50], where the thermodynamic notions of heat, work, energy, and entropy, and the classical laws of thermodynamics, were generalized to individual, fluctuating trajectories for processes arbitrarily far from equilibrium connected to a heat bath at temperature T (the medium/water for the Brownian particle) [51–54]. A central assumption of stochastic thermodynamics is that of *local equilibrium* or *local detailed balance* [55], stating that all unobserved degrees of freedom evolve sufficiently fast compared to the observed ones, such that the former can be considered to be in equilibrium at the temperature of the heat bath. In terms of the Langevin description in Eq. (1.3.2) this means that we need to consider all slow (mesoscopic) degrees of freedom in x, which is challenging for large numbers of degrees of freedom.

A key quantity to describe and quantify non-equilibrium processes is the *total entropy production* $\Delta S_{\rm tot} = S_{\rm tot}(t) - S_{\rm tot}(0)$. This encompasses all contributions to the entropy production, e.g., in the case of a mesoscopic system in some medium, we account for the contributions to the entropy production in the system and the medium. Following the principles of stochastic thermodynamics, we can assign a fluctuating entropy production $\Delta s_{\rm tot}$ to individual trajectories, where the deterministic quantity $\Delta S_{\rm tot} = \langle \Delta s_{\rm tot} \rangle$ corresponds to the average over many realizations of the process [54].

As an illustration of how stochastic thermodynamics and Langevin dynamics (1.3.2) extend beyond the classical concepts of thermodynamics, consider the *integrated fluctuation theorem* for the total entropy production [28, 54],

$$\langle \exp(-\Delta s_{\rm tot}) \rangle = 1\,. \tag{1.7.1}$$

Equation (1.7.1) belongs to the class of fluctuation theorems in stochastic thermodynamics [28, 56] that was initiated by seminal works in the 1990s [57–60] and proved particularly useful for inferring free energy differences from non-equilibrium measurements via the Jarzynski [61–63] and Crooks relations [64–66]. As mentioned above, the tools of equilibrium thermodynamics no longer apply far from equilibrium, and little is known about general concepts. In this sense, fluctuation theorems represent a conceptual advance to formulate non-system-specific features of non-equilibrium dynamics. Considering the particular fluctuation theorem in Eq. (1.7.1), we see that, on the one hand, it reflects a physical law for the probability distribution of the stochastic entropy production along in-

[5] The same holds true for another class of stochastic processes, namely Markov jump dynamics. Since they operate on a discrete state space, they are however less fundamental in terms of the connection to Newtonian dynamics.

dividual stochastic paths, while classical thermodynamics only considers its average ΔS_{tot}. On the other hand, it also contains the classical statement, since by applying Jensen's inequality $\langle \exp(-\Delta s_{\mathrm{tot}}) \rangle \geq \exp(-\langle \Delta s_{\mathrm{tot}} \rangle)$ the second law of thermodynamics follows from Eq. (1.7.1), i.e., $\Delta S_{\mathrm{tot}} \geq 0$. This illustrates that the second law only holds on the level of average values, while Eq. (1.7.1) implies that (unless $\Delta s_{\mathrm{tot}} = 0$ with probability one) there will always be instances of $\Delta s_{\mathrm{tot}} < 0$.

1.8 Thermodynamic uncertainty relations

An important recent development in the field of stochastic thermodynamics are diverse trade-off relations between precision and entropy production. Most notably, there are *thermodynamic uncertainty relations* (TURs) [47,48,67] for the precision of stochastic currents. In an NESS, the TUR states that the product of the relative variation of a current and the average entropy production ΔS_{tot}[6] is bounded from below by $2k_{\mathrm{B}}$, i.e., for J_t as in Eq. (1.6.1),

$$\frac{\mathrm{variance}(J_t)}{\langle J_t \rangle^2} \Delta S_{\mathrm{tot}} \geq 2k_{\mathrm{B}} \,. \tag{1.8.1}$$

This inequality highlights the fundamental role of fluctuations in stochastic systems by showing that the variance of currents is bounded by the entropy production. Thus, we see that fluctuations do not merely complicate the inference of averages, but instead they contain valuable information about thermodynamical properties of non-equilibrium systems in practice.

In resemblance to the fluctuation theorem (1.7.1), the TUR (1.8.1) can also be seen as an extension of the second law of thermodynamics, $\Delta S_{\mathrm{tot}} \geq 0$, since for any J_t with $\langle J_t \rangle \neq 0$, Eq. (1.8.1) gives a strictly positive lower bound for ΔS_{tot}. In contrast to the fluctuation theorem, however, the TUR is more directly applicable in practice. Considering a trajectory x_τ with $0 \leq \tau \leq t$, one can evaluate J_t from Eq. (1.6.1), and estimate its average and variance from repeated measurements. Then the TUR (1.8.1) directly gives a lower bound for ΔS_{tot} which yields valuable, and otherwise hardly accessible, information on how far a given process operates from equilibrium [68, 69]. Importantly for practical relevance, the TUR also applies to partially observed dynamics, i.e., we do not need to measure the dynamics of all slow degrees of freedom to apply the TUR. Thus, we obtain information on the entropy production of the full system by observing only a subset of degrees of freedom.

[6] In the particular case of an NESS, entropy is actually only produced in the medium, and $T\Delta S_{\mathrm{tot}}$ corresponds to the heat dissipated into the medium, see Sec. 2.9 or Ref. [28].

Despite a variety of techniques applied to prove and generalize the TUR [48, 49, 70–72], a direct derivation solely based on stochastic-calculus techniques was so far elusive.

1.9 Selected further topics

Another emerging topic in non-equilibrium physics is the far-from-equilibrium dynamics of heating and cooling [73–76], which can be described by the Langevin dynamics for arbitrarily large temperature differences. Close to equilibrium, i.e., for small temperature differences, linear irreversible thermodynamics predicts that heating and cooling are symmetric. Quantifying heating and cooling of a Langevin system in terms of an excess free energy, it was recently found that in harmonic potentials beyond the linear regime (i.e., for large temperature differences) thermal relaxation becomes asymmetric and heating towards an equilibrium state occurs faster than the corresponding cooling process [73].

On another note, there is a vast variety of non-equilibrium stochastic processes relevant to biophysical systems beyond the Langevin dynamics in Eq. (1.3.2), e.g., *active* particles or field theories [77–80]. In particular, if the unobserved degrees of freedom in the medium do not relax on a timescale much faster than the observed motions, the system will develop *memory*, i.e., the evolution at time t no longer only depends on the current state x_t as in Eq. (1.3.2), but also on preceding dynamics. A prominent example is the diffusive motion of particles in crowded environments, such as the interior of biological cells, giving rise to *anomalous diffusion* with $\mathrm{variance}(x_t) \propto t^\alpha$ with $\alpha > 0$ where classical Brownian motion (normal diffusion) only covers the case $\alpha = 1$ [81, 82].

1.10 Outline of the thesis

An overarching theme of this thesis is to address a selection of open questions through the lens of stochastic calculus. Even though this requires some initial effort to develop the necessary mathematical methods, we will see how it can lead to conveniently *direct* rederivations of known results, and, in particular, to a variety of new results.

The thesis is organized as follows. In Ch. 2, we provide the basic background on probability theory, stochastic calculus and stochastic thermodynamics. The following Chs. 3-6 treat specific topics and therefore contain separate dedicated introductions, and refine the technical background developed in Ch. 2, to the respective topics. Each chapter encompasses one or several of the publications from Refs. [83–90] reproduced from different journals. Chapter 3 employs stochastic calculus to address the (co)variance and coarse graining of

the time-integrated densities and currents introduced in Eq. (1.6.1). Based on the methods developed in Ch. 3, we rederive and generalize the TUR (1.8.1) in Ch. 4, and discuss when the inequality becomes saturated. The investigation of fluctuations of time-integrated observables in Chs. 3 and 4 makes up the major part of the thesis. Chapter 5 addresses the asymmetry of heating and cooling based on the Langevin equation by providing experimental evidence for the results of Ref. [73] and generalizing the theory to relaxation towards an NESS. Going beyond the Langevin dynamics in Eq. (1.3.2), in Ch. 6 we address dynamics that stochastically switches between different Langevin motions, and even between motions featuring anomalous diffusions, and apply the results to particle tracking data recorded in living cells in Ref. [91]. The thesis is completed by a conclusion, including an outlook, in Ch. 7.

Chapter 2

Fundamentals

Here, we lay the mathematical and physical foundation for the coming chapters. For a large part of this chapter, we recite textbook knowledge, but selected and arranged in a way that is most suitable for the scope of this thesis. To keep the mathematical introductions practical and pedagogical, we state many mathematical results and concepts in weaker or less general versions, sometimes even in a less precise version. The foundation laid in this chapter is to be completed by more specialized technical introductions given for the individual topics in the following chapters.

This chapter covers a selection of basic topics in probability theory and stochastic processes, in particular the mathematical formulation of Brownian motion, i.e., the Wiener process. We then give a practical introduction to stochastic calculus and the Fokker-Planck equation with a focus on non-equilibrium steady states. We conclude this predominantly mathematical chapter by presenting the physical concept of entropy production for Langevin dynamics.

2.1 Probability measures and integration

To properly define probability measures and integrals over random variables, we require a brief review of measure theory and integration. Details on all mentioned facts can, e.g., be found in Ref. [92]. Although we will not explicitly need these technical definitions in the later parts of the thesis, it will sometimes be helpful to be able to reduce problems back to fundamental mathematical definitions in order to avoid confusion or mistakes. In particular, we present a solid basis for dealing with technical details of limits and integrals towards the end of this section, which will be used throughout the work.

2.1.1 Probability spaces

Here, we formally define probability measures and probability spaces. Consider a set Ω called *sample space*. For rolling a dice, this would be $\Omega = \{1, 2, 3, 4, 5, 6\}$. To assign probabilities to different *events*, e.g., the event $E = \{1, 3, 5\}$ to roll an odd number, we define a set $\mathcal{A}$ containing subsets of Ω. In the example of the dice, we would include all subsets of Ω in $\mathcal{A}$, i.e., $E \in \mathcal{A}$ for all subsets $E \subset \Omega$. If the set Ω is uncountably infinite, e.g., an interval in $\mathbb{R}$, we would run into problems when assigning probabilities to all subsets of Ω [92]. To formally define a useful set of events, we therefore define a certain class of $\mathcal{A}$. We call a set $\mathcal{A}$ of subsets of Ω a *σ-algebra* in Ω if it contains Ω, and is closed under the formation of complements ($E^c \equiv \Omega \setminus E$, i.e., the set that remains if we remove all elements of E from Ω) and countable unions, i.e.,

1. $\Omega \in \mathcal{A}$,
2. $E \in \mathcal{A}$ implies $E^c \in \mathcal{A}$,
3. $E_1, E_2, \dots \in \mathcal{A}$ implies $\bigcup_i E_i \in \mathcal{A}$.

In probability theory, the σ-algebra is called *event space*. To assign probabilities to events $E \in \mathcal{A}$, we define a *probability measure* as a function $\mathbb{P}$ that satisfies

1. $0 \leq \mathbb{P}(E) \leq 1$ for all $E \in \mathcal{A}$,
2. $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$,
3. if $E_1, E_2, \dots \in \mathcal{A}$ is a disjoint (i.e., $E_i \cap E_j = \emptyset$ for $i \neq j$) sequence of sets in $\mathcal{A}$, then

$$\mathbb{P}\left(\bigcup_i E_i\right) = \sum_i \mathbb{P}(E_i)\,. \tag{2.1.1}$$

For the example of a dice, defining $\mathbb{P}(\{i\}) = 1/6$ for $i = 1, 2, \dots, 6$ implies via Eq. (2.1.1) that $\mathbb{P}(\{1, 3, 5\}) = 1/2$.

A triple $(\Omega, \mathcal{A}, \mathbb{P})$ of a set Ω, a sigma-algebra $\mathcal{A}$ in Ω, and a probability measure $\mathbb{P}$ on $\mathcal{A}$ is called *probability space*.

2.1.2 Measure spaces and the Lebesgue integral

To address a slightly more general setting, we drop the condition $\mathbb{P}(E) \leq 1$ that we had for probability measures to generally define a *measure* μ satisfying

1. $0 \leq \mu(E)$ for all $E \in \mathcal{A}$,
2. $\mu(\emptyset) = 0$,
3. if $E_1, E_2, \cdots \in \mathcal{A}$ is a disjoint (i.e., $E_i \cap E_j = \emptyset$ for $i \neq j$) sequence of sets in $\mathcal{A}$, then

$$\mu\left(\bigcup_i E_i\right) = \sum_i \mu(E_i)\,. \tag{2.1.2}$$

A pair $(\Omega, \mathcal{A})$ is called *measurable space* and a triple $(\Omega, \mathcal{A}, \mu)$ is called *measure space*.

Apart from the probability spaces defined above (which are a special case of measure spaces), for the scope of the thesis we only need a particular class of measure spaces, namely those where Ω is an interval in $\mathbb{R}$, or a product space of intervals such that $\Omega \subset \mathbb{R}^d$, e.g., $\Omega = [0,1]$, $\Omega = [0,\infty)$, $\Omega = \mathbb{R}$ or $\Omega = [0,\infty) \times [0,\infty) \subset \mathbb{R}^2$. In such cases, we choose $\mathcal{A}$ to be the *Borel σ-algebra* $\mathcal{B}(\Omega)$, which can be defined as the smallest σ-algebra containing all open intervals in Ω. Note that this also contains all closed sets since σ-algebras are closed under the formation of complements (in particular, $\mathcal{B}(\Omega)$ contains all sets $\{x\}$ for $x \in \Omega$). Given two measurable spaces $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{B}(\Omega'))$ with $\Omega' \subset \mathbb{R}^d$, we define a function $f\colon \Omega \to \Omega'$ to be a *measurable function* if the preimages of $\mathcal{B}(\Omega')$ are contained in $\mathcal{A}$, i.e., if $f^{-1}(E') \in \mathcal{A}$ for all $E' \in \mathcal{B}(\Omega')$. If $\Omega \subset \mathbb{R}^d$ with $\mathcal{A} = \mathcal{B}(\Omega)$, the class of measurable functions is very large, e.g., it contains all pointwise limits of sequences of continuous functions, which is why one in practice may assume all occurring functions to be measurable.

The standard measure used on $\mathcal{B}(\Omega)$ is the *Lebesgue measure* λ that measures the volume of sets in $\mathcal{B}(\Omega)$. For $\Omega \subset \mathbb{R}$, it can be uniquely defined by requiring for intervals $(a, b]$ that it measures the length $\lambda((a,b]) = b - a$. In higher-dimensional space, it is defined via the volumes of boxes, e.g., $\lambda((a,b] \times (c,d]) = (b-a)(d-c)$ for $\Omega \subset \mathbb{R}^2$. Note that the above definitions imply that $\lambda((a,b)) = \lambda([a,b]) = b - a$ and $\lambda(\{x\}) = 0$ for all $x \in \Omega$, i.e., points have Lebesgue measure 0, agreeing with the geometrical intuition of the length of an interval. Moreover, lines have Lebesgue measure 0 in $\mathbb{R}^d$ for $d \geq 2$ and in general, n-dimensional surfaces have Lebesgue measure 0 in $\mathbb{R}^d$ for $d \geq n+1$.

The Lebesgue measure is very useful for defining integration via measure theory, as an alternative to the Riemann integral. In short, the Lebesgue integral over a function f is formally defined by requiring linearity and writing f as a limit of linear combinations of indicator functions (we use the symbol $\equiv$ for definitions throughout this thesis),

$$\mathbb{1}_E(x) \equiv \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{otherwise} \end{cases}\,, \tag{2.1.3}$$

where the integral over indicator functions is set to give $\int 1_E \mathrm{d}\lambda = \lambda(E)$. In this way, the Lebesgue integral can be defined for all positive measurable functions, and for all measurable functions (with arbitrary signs) as long as $\int |f| \, \mathrm{d}\lambda < \infty$. As required for consistency, the Riemann and the Lebesgue integral give the same result for any setting where both are properly defined, which is why one may also write $\int f(x)\mathrm{d}x$ for $\int f \mathrm{d}\lambda$.

2.1.3 Random variables and expectation values

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a Borel measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, a measurable function $X \colon \Omega \to \mathbb{R}^d$ is called *random variable*. For the purpose of this thesis, all considered probability measures possess a *probability density* $p \colon \mathbb{R}^d \to [0, \infty)$ which relates the probability measure $\mathbb{P}$ to the Lebesgue measure λ, and thus to integrals $\mathbb{P}(E) = \int_E p(x)\mathrm{d}x$. This allows to define the *expectation value*, also called *mean* or *average*, of $f(X)$ for any measurable function $f \colon \mathbb{R}^d \to \mathbb{R}^{d'}$ as (in the multidimensional case, we use $\mathrm{d}x = \mathrm{d}x_1 \cdots \mathrm{d}x_{d'}$ to denote the d'-dimensional volume integral)

$$\langle f(X) \rangle = \int_{\mathbb{R}^{d'}} f(x) p(x) \mathrm{d}x \,. \tag{2.1.4}$$

Since expectation values are special cases of integrals, they are linear, i.e., for random variables X, Y and $\alpha \in \mathbb{R}$ we have $\langle \alpha X + Y \rangle = \alpha \langle X \rangle + \langle Y \rangle$.

Apart from probability densities and measures, one may also use *probability distributions* defined, e.g., for $d = 1$ as $P(x) \equiv \mathbb{P}((-\infty, x])$ or, if the density exists, as $P(x) \equiv \int_{-\infty}^{x} p(x)\mathrm{d}x$.

2.1.4 Interchanging integrations, expectation values, limits and derivatives

We are usually allowed to interchange the order of integrations, expectation values, limits and derivatives. If we, e.g., consider functions $f(t)$, $f(s,t)$ or $f_n(t)$ depending on parameters $s, t \in \mathbb{R}^d$ and $n \in \mathbb{N}$, we usually have

$$\begin{aligned} \lim_{n\to\infty} \int \mathrm{d}s f_n(s) &= \int \mathrm{d}s \lim_{n\to\infty} f_n(s) \\ \partial_t \int \mathrm{d}s f(s,t) &= \int \mathrm{d}s \, \partial_t f(s,t) \\ \int \mathrm{d}s \left[\int \mathrm{d}t f(s,t) \right] &= \int \mathrm{d}t \left[\int \mathrm{d}s f(s,t) \right] . \end{aligned} \tag{2.1.5}$$

As we saw above, expectation values can be understood as integrals, see Eq. (2.1.4), such that for random variables X the above Eq. (2.1.5) becomes

$$\begin{aligned} \lim_{n\to\infty} \langle X_n \rangle &= \left\langle \lim_{n\to\infty} X_n \right\rangle \\ \partial_t \langle X(t) \rangle &= \langle \partial_t X(t) \rangle \\ \left\langle \int \mathrm{d}t X(t) \right\rangle &= \int \mathrm{d}t \langle X(t) \rangle \,. \end{aligned} \tag{2.1.6}$$

These calculation rules, and strict conditions on when they hold, can be very well formulated in terms of measure-theoretic integration.

To this end, we consider the *monotone convergence theorem* (MCT) [92, Theorem 16.2] and the *dominated convergence theorem* [92, Theorem 16.4]. These theorems state that limits and integrals can be interchanged (i.e., that the first lines of Eqs. (2.1.5) and (2.1.6) hold) under the condition that f_n converges to f (except possibly on sets of measure zero) and that either $0 \leq f_1 \leq f_2 \leq \dots$ (monotone convergence) or that there exists a g with $\int |g(x)| \, \mathrm{d}x < \infty$ and $|f_n| \leq g$ for all n (dominated convergence), which are fairly mild assumptions. By defining $f_n(s) = n[f(s, t+1/n) - f(s,t)]$ and using linearity of the integral, we obtain the second lines of Eqs. (2.1.5) and (2.1.6), respectively, as a consequence of the first lines. The third lines of Eqs. (2.1.5) and (2.1.6) hold under the related assumption that either $f(s,t) \geq 0$ (everywhere except possibly on sets of measure zero; the value ∞ is allowed for the result of the integral) or that we have $\int \mathrm{d}t \int \mathrm{d}s \, |f(s,t)| < \infty$, see Fubini's theorem [92, Theorem 18.3].

Due to the weak assumptions under which the mentioned theorems hold, one often uses Eqs. (2.1.5) and (2.1.6) without explicitly mentioning them. Moreover, examples where the assumptions, and accordingly Eqs. (2.1.5) and (2.1.6), do not hold are rather intuitive. E.g., for $f(s,t) = \sin(t)$, we have that $\int_0^\infty \mathrm{d}t \int_0^\infty \mathrm{d}s \, |\sin(t)| = \infty$ and we cannot say that $\int_0^\infty \mathrm{d}t \int_0^\infty \mathrm{d}s \sin(t) = \int_0^\infty \mathrm{d}s \int_0^\infty \mathrm{d}t \sin(t)$ (actually, we cannot compute either side of the equation).

2.1.5 A note on the delta distribution

Note that the delta distribution $\delta(x)$, defined in the sense of Schwartz distributions (i.e., through the action on test functions f) via $\int f(x)\delta(x)\mathrm{d}x = f(0)$, is not a Lebesgue integrable function. This is because $\delta(x) \neq 0$ holds only in the point $x = 0$ which has Lebesgue measure 0, which would imply for the Lebesgue integral that $\int \delta \mathrm{d}\lambda = 0$ if it was defined. However, one may always view the delta distribution as a limit of integrable functions, e.g., $\delta(x) = \lim_{\varepsilon\to 0} g_\varepsilon(x)$ for $g_\varepsilon(x) \equiv \exp(-x^2/2\varepsilon^2)/\sqrt{2\pi\varepsilon^2}$, and then understand the integral

as $\int f(x)\delta(x)\mathrm{d}x = \int f(x)\lim_{\varepsilon\to 0} g_\varepsilon(x)\mathrm{d}x \overset{!}{=} \lim_{\varepsilon\to 0}\int f(x)g_\varepsilon(x)\mathrm{d}x$. The condition that the integral and the limit commute, as indicated in the last step, is necessary for the distributional expression to be well-defined, i.e., the left side of the equation is only defined as the limit of the integrated expression $\int f(x)g_\varepsilon(x)\mathrm{d}x$.

2.2 Probability theory based on probability densities

From now we consider the basic concepts of σ-algebra, measurability of functions and random variables, and the existence of densities for probability measures as given. Under these simplifying assumptions, we introduce further basic concepts of probability theory.

2.2.1 Moments and cumulants

For a random variable X with probability density function $p(x)$[1], the information about the distribution of X is completely contained in $p(x)$. For example, a one-dimensional *Gaussian* (also called *normal* or *normally distributed*) random variable X is specified by the probability density

$$p_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{2.2.1}$$

with $\mu\in\mathbb{R}$ and $\sigma>0$. Although all information is contained in Eq. (2.2.1), it is often useful to state this information, or parts of it, in different forms. In most practical cases, only parts of this information are available. The arguably most important piece of information about a random variable X is its *expectation value*, also called *average* or *mean*,

$$\langle X\rangle = \int xp(x)\mathrm{d}x. \tag{2.2.2}$$

In the example Eq. (2.2.1) we have $\langle X\rangle = \mu$. The most important expression to quantify fluctuations around this mean value is the *variance*

$$\mathrm{var}(X) \equiv \langle (X-\langle X\rangle)^2\rangle = \langle X^2\rangle - \langle X\rangle^2 \geq 0\,. \tag{2.2.3}$$

In the example of Eq. (2.2.1) we have $\mathrm{var}(X)=\sigma^2$. We define the k-th *moment* of X or of $p(x)$ as $M_k \equiv \langle X^k\rangle = \int x^k p(x)\mathrm{d}x$ for $k\in\mathbb{N}$. If one chooses to include $k=0$, the zeroth moment $\langle 1\rangle = 1$ trivially reflects the normalization of $p(x)$. The variance is an expression

[1] Note that x is a meaningless dummy variable here, and the notation $p(x)$ actually refers to the function $p\colon \mathbb{R}^d\to\mathbb{R}$.

of the mean $M_1 = \langle X \rangle$, and of $M_2 = \langle X^2 \rangle$. Although the first two moments contain much less information than $p(x)$, they already suffice to make statements about the average of X and fluctuations around it. For example, $\mathrm{var}(X) = 0$ would imply that $X = \langle X \rangle$ with probability one. Note, however, that the average value is not always the typical or most likely value (only in special cases like Eq. (2.2.1) it is the case that $p(x)$ peaks at $x = \langle X \rangle$), such that the interpretation of fluctuations *around* the average is not always the correct picture.

To obtain a more complete, but still reduced, picture of the properties of X one can include further moments. For example, the Gaussian example in Eq. (2.2.1) has $\langle (X-\mu)^4 \rangle / \sigma^4 = 3$. With this equation, knowledge of the first four moments of X allows to evaluate the expression $\langle (X - M_1)^4 \rangle / (M_2 - M_1^2)^2$, also known as *kurtosis*, and use its deviation from 3 as a measure of non-Gaussianity.

If all moments for $k \in \mathbb{N}$ of X exist (i.e., if all $M_k < \infty$) and if the series $\sum_k M_k s^k / k!$ has a positive radius of convergence in s, then $p(x)$ can in theory be reconstructed from them [92, Theorem 30.1]. A way to do understand this is via the *moment generating function* (also known as *characteristic function*), which is defined as [92, Eq. (21.21)][2]

$$M(s) \equiv \langle \mathrm{e}^{\mathrm{i}sX} \rangle = \int \mathrm{e}^{\mathrm{i}sx} p(x) \mathrm{d}x \,, \tag{2.2.4}$$

where the integration runs over the range of X. We stick to one-dimensional X for now, but the definition in Eq. (2.2.4) generalizes to multidimensional X by interpreting sX as a scalar product of two vectors.

Taking derivatives at $s = 0$ shows why $M(s)$ is called moment generating function,

$$\left(\frac{\mathrm{d}}{\mathrm{d}s} \right)^k_{s=0} M(s) = \left\langle \mathrm{i}^k X^k \mathrm{e}^{\mathrm{i}sX} |_{s=0} \right\rangle = \mathrm{i}^k M_k \,. \tag{2.2.5}$$

If the M_k are known and the Taylor series converges, one can thus write $M(s) = \sum_{k=0}^{\infty} \mathrm{i}^k s^k M_k / k!$ which determines $M(s)$ uniquely in terms of the moments. From the inverse Fourier transform, one can then (at least in theory) reconstruct $p(x)$.

For the Gaussian example (2.2.1) the characteristic function reads

$$M_G(s) = \int_{-\infty}^{\infty} \mathrm{e}^{isx} p_G(x) \mathrm{d}x = \mathrm{e}^{\mathrm{i}\mu s - \frac{1}{2}\sigma^2 s^2} \,, \tag{2.2.6}$$

[2] For the scope of the thesis we may assume that the integral exists for all $s \in \mathbb{R}$.

from which one can get the moments via Eq. (2.2.5) $M_1 = \mu$, $M_2 = \sigma^2 + \mu^2$ and so on. For $\mu = 0$, all odd moments vanish, but even moments do not.

The exponential form of Eq. (2.2.6) suggests another concept, closely related to moments, namely the *cumulants*, where the k-th cumulant is defined in analogy to the moments in Eq. (2.2.5) as

$$\kappa_k \equiv \mathrm{i}^{-k} \left(\frac{\mathrm{d}}{\mathrm{d}s} \right)^k_{s=0} \log[M(s)] \,. \tag{2.2.7}$$

In general, the k-th cumulant is a function of the first k moments, e.g., $\kappa_1 = M_1$ and $\kappa_2 = M_2 - M_1^2 = \mathrm{var}(X)$. The Gaussian density (2.2.1) can be defined via the property that the first two cumulants are $\kappa_1 = \mu$ and $\kappa_2 = \sigma^2$, while all higher cumulants vanish, i.e., $\kappa_k = 0$ for $k \geq 3$.

2.2.2 Joint and conditional densities

We mentioned above that the information about the distribution of a random variable is contained in the density $p(x)$, or alternatively in the moments, cumulants or moment generating function. However, if we consider more than one random variable, e.g., X and Y, there may also be correlations between them which are not contained in the individual $p_X(x)$ and $p_Y(y)$. Instead, we have to consider the *joint density* $p_{X,Y}(x, y)$ that for intervals (or in higher dimensions products of intervals) $I_{1,2}$ quantifies the joint probability that X is in I_1 and Y in I_2, i.e.,

$$\mathbb{P}(X \in I_1;\, Y \in I_2) = \int_{I_1} \mathrm{d}x \int_{I_2} \mathrm{d}y \, p_{X,Y}(x, y). \tag{2.2.8}$$

Accordingly, one can also define joint expectation values. In general, $p_{X,Y}(x, y)$ contains more information than $p_X(x)$ and $p_Y(y)$ together. The individual densities are recovered by *marginalization*, i.e., by integrating out the other variable as, e.g., in $p_X(x) = \int \mathrm{d}y p_{X,Y}(x, y)$.

Closely related is the concept of *conditional* probabilities. For the probability of $X \in I_1$ conditioned on the event $Y \in I_2$, Bayes' theorem states that[3] $\mathbb{P}(X \in I_1 | Y \in I_2) = \mathbb{P}(X \in I_1;\, Y \in I_2)/\mathbb{P}(Y \in I_2)$. The conditional density of X given that $Y = y$ accordingly reads

$$\mathbb{P}_{X|Y}(x|y) \equiv \frac{p_{X,Y}(x, y)}{p_Y(y)} \,. \tag{2.2.9}$$

[3] Technically, the first $\mathbb{P}$ on the right-hand side is the product measure on the product of the two probability spaces and the second $\mathbb{P}$ is the probability measure on Y.

2.2.3 Independent random variables and the convolution theorem

Random variables X and Y are called *independent* if for all events E_1 and E_2 we have $\mathbb{P}(X \in E_1; Y \in E_2) = \mathbb{P}(X \in E_1)\mathbb{P}(Y \in E_2)$. Intuitively, this means that the events happen independently, i.e., fixing the value of X does not alter the distribution of Y and vice versa. For independent variables, the joint density does not contain more information than the individual densities and simply factorizes, $p_{X,Y}(x, y) = p_X(x)p_Y(y)$.

More than two random variables are called independent if their joint density factorizes, i.e., $p_{X_1,X_2,X_3,\dots}(x_1, x_2, x_3, \dots) = p_{X_1}(x_1)p_{X_2}(x_2)p_{X_3}(x_3)\cdots$.

When adding independent random variables X, Y, their densities are convoluted, i.e., the density of $X + Y$ reads

$$p_{X+Y}(z) = \int \mathrm{d}x p_{X,Y}(x, z - x) = \int \mathrm{d}x p_X(x)p_Y(z - x)\,, \tag{2.2.10}$$

which via the convolution theorem for Fourier transforms and the definition of the moment generating function in Eq. (2.2.4) implies that their joint moment generating function factorizes,

$$M_{X+Y}(s) = M_X(s)M_Y(s)\,. \tag{2.2.11}$$

Applying this iteratively to several independent variables yields

$$M_{\sum_{k=1}^n X_k}(s) = \prod_{k=1}^n M_{X_k}(s)\,. \tag{2.2.12}$$

For example, when adding two independent Gaussian random variables with density as in Eq. (2.2.1) with means $\mu_{x,y}$ and variances $\sigma^2_{x,y}$, the moment generating functions Eq. (2.2.6) are multiplied such that

$$M_{X+Y}(s) = \exp[\mathrm{i}(\mu_x + \mu_y)s - (\sigma_x^2 + \sigma_y^2)s^2/2]\,, \tag{2.2.13}$$

which implies that $X + Y$ is a Gaussian random variable with mean $\mu_x + \mu_y$ and variance $\sigma_x^2 + \sigma_y^2$.

2.2.4 The central limit theorem

As mentioned in the Introduction in Ch. 1, a very powerful statement about the sum of independent random variables is the *central limit theorem*, which in a simple form states the following [92, Theorem 27.1]. Consider one-dimensional random variables X_i, $i \in \mathbb{N}$, that are independent and identically distributed (i.e., they have the same density) with mean μ_1 and variance σ_1^2, then[4]

$$\frac{\sum_{i=1}^n X_i - n\mu_1}{\sigma_1\sqrt{n}} \overset{n\to\infty}{\longrightarrow} \mathcal{N}(0,1), \tag{2.2.14}$$

where $\mathcal{N}(\mu,\sigma^2)$ denotes the normal distribution with mean μ and variance σ^2 (i.e., with density as in Eq. (2.2.1)). Thus, the central limit theorem states that the distribution of the sum of random variables approaches a Gaussian which is the reason why this distribution is so universal, in particular, for the Brownian motion (see Secs. 1.1 and 1.2).

The central limit theorem can be proven in a simple way via the moment generating function $M(s)$. For a quick sketch of this proof, consider identically distributed, independent variables Y_i with mean zero, i.e., set $Y_i = X_i - \mu_1$ above, and variance σ^2. Then the moment generating function for $s \to 0$ reads

$$M_{Y_i}(s) = 1 - \frac{\sigma^2}{2}s^2 + O(s^3)\,. \tag{2.2.15}$$

By definition of the moment generating function, see Eq. (2.2.4), we get for $c \in R$ that $M_{cY_i}(s) = M_{Y_i}(cs)$, such that using Eq. (2.2.12) we have

$$M_{\sum_{i=1}^n Y_i/\sqrt{n}\sigma}(s) = \left[1 - \frac{s^2}{2n} + O\left(\frac{s^3}{n^{3/2}}\right)\right]^n \overset{n\to\infty}{\longrightarrow} \mathrm{e}^{-s^2/2}\,, \tag{2.2.16}$$

where we used $(1+x/n)^n \to \mathrm{e}^x$. The obtained exponential is the moment generating function of a normal distribution of mean 0 and variance 1, see Eq. (2.2.6). According to Lévy's continuity theorem [92, Theorem 26.3], convergence in the moment generating function implies convergence of the probability density of the random variable, which completes the proof of the central limit theorem.

The central limit theorem can be generalized in several ways, allowing for non-identically distributed random variables (see, e.g., [92, Theorem 27.2]) or for weakly correlated random variables (see, e.g., [93, Theorem 18.1.1] and [94, Ch. 7]). Crucial, however, is the

[4] Technically, the convergence in Eq. (2.2.14) is *in distribution*, i.e., the probability distribution of the left-hand side converges to that of a Gaussian. Note that one cannot say that the left-hand side of Eq. (2.2.14) becomes a Gaussian random variable since, e.g., its *large deviations* remain non-Gaussian, see Subsec. 3.2.3.

assumption of a finite variance. Limit theorems for variables with infinite variances also exist [95], however, they will no longer give Gaussian distributions and are therefore not direct generalizations of Eq. (2.2.14).

2.3 Stochastic processes

Instead of simple random variables, we are usually interested in random processes in time, e.g., a Brownian particle with position coordinate X_t that stochastically evolves in time $t \geq 0$. Formally, a *stochastic process* in continuous time t is defined as a collection $\{X_t \colon t \geq 0\}$ of random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is a specific time-ordered set of σ-algebras called *filtration*. As in the previous section, we consider measurability and the existence of densities to be given, such that we do not need to go into details regarding the filtration. In analogy to the example of a Brownian particle, we often refer to X_t as the *position* at time t.

The distribution of the process is often quantified in terms of the *finite-dimensional distributions* $\mathbb{P}[(X_{t_1}, \ldots, X_{t_k}) \in E]$ for some (Borel) measurable sets $E \subset \mathbb{R}^k$ and $k \in \mathbb{N}$. Although the finite-dimensional distributions do not specify a process uniquely (see [92, Eq. (36.1)]), they usually contain the relevant information about the process and correlations within it. In terms of probability densities, the finite-dimensional distributions correspond to the k-point joint densities $p(x_1, t_1; \ldots; x_k, t_k)$.

2.3.1 Markov processes

Most of the stochastic processes considered in this thesis belong to the class of *Markov processes*, which simplifies the description of their relevant properties. A stochastic process is called *Markov process* if for $k \geq 1$ and $0 \leq t_1 \leq \cdots \leq t_k \leq t$, the conditional density fulfills

$$p(x, t|x_1, t_1; \ldots; x_k, t_k) = p(x, t|x_k, t_k)\,. \tag{2.3.1}$$

This means that for Markov processes, the time-evolution after time t_k only depends on the position at t_k, and not on earlier positions.

We define a Markov process to be *time homogeneous* if the conditional density $p(x, t|x', t')$ only depends on the time difference $t - t'$, i.e., if we always have $p(x, t|x', t') = p(x, t - t'|x') \equiv p(x, t - t'|x', 0)$. In this case, the joint density at times $0 \leq t_1 \leq \cdots \leq t_k$ can be

written in terms of the *transition density* $p(x,t|x')$ (also called *Green's function* and written as $G(x,t|x')$) as [16, Eq. (2.23)]

$$p(x_1,t_1;\ldots;x_k,t_k) = p(x_1,t_1)p(x_2,t_2-t_1|x_1)\cdots p(x_k,t_k-t_{k-1}|x_{k-1})\,. \tag{2.3.2}$$

A central equation for Markov processes is the *Chapman-Kolmogorov equation*, which in terms of the transition density for time-homogeneous Markov processes for any $0<\tau<t$ reads [16, Eq. (2.26)]

$$p(y,t|x) = \int_{\mathbb{R}^d} p(y,t-\tau|z)p(z,\tau|x)\mathrm{d}z\,. \tag{2.3.3}$$

This is very intuitive since it integrates over all possible intermediate points z at time τ when going from x to y, but note that this crucially requires $p(y,t|z,\tau;x,0) = p(y,t|z,\tau)$ which is why it only generally holds for Markov processes. Note that while the Markov property implies the Chapman-Kolmogorov equation, the latter does not imply for a process to be Markovian [96, 97]. Generalizations of Eqs. (2.3.2) and (2.3.3) to Markov processes that are not time homogeneous are immediate [16] but are not relevant to the thesis.

2.3.2 Gaussian processes

A stochastic process is called a *Gaussian process* if all joint densities are mulitvariate (i.e., multidimensional) Gaussian densities. Note that this is a stricter requirement than just demanding that $p(x,t)$ is a Gaussian density for all $t\geq 0$. For time-homogeneous Markov processes, according to Eq. (2.3.2) it suffices if the initial condition and the transition density are Gaussian.

2.3.3 Steady-state density and ergodic processes

For long times, many processes approach an invariant probability measure with density called *steady-state density* p_s. Such processes are typically[5] ergodic. For our purpose, we define that a stochastic process x_t to be *ergodic* if a steady-state density exists and [16, Eqs. (2.38)-(2.39)]

$$\langle f(x_t)\rangle_\mathrm{s} = \lim_{T\to\infty}\frac{1}{T}\int_0^T f(x_\tau)\mathrm{d}\tau\,, \tag{2.3.4}$$

where $\langle\cdot\rangle_\mathrm{s}$ denotes the average over the steady-state density $\langle f(x_t)\rangle_\mathrm{s} \equiv \int \mathrm{d}x\, p_\mathrm{s}(x)f(x)$.

[5] Examples for non-ergodic processes with an invariant density can be constructed by introducing time-inhomogeneous probability currents that do not alter p_s.

Note that, e.g., a Brownian motion in an infinite domain (e.g., on the real line; detailed definition in Sec. 2.4) is not an ergodic process[6] since the probability density does not converge to a normalized density for long times (the density spreads out indefinitely such that we would get $p_\mathrm{s}(x) = 0$ for all x but this is not normalizable). However, a Brownian motion confined in a finite interval or in a sufficiently confining potential is an ergodic process.

2.4 Mathematical definition of Brownian motion

A *Brownian motion* or *Wiener process* is a stochastic process $[W_t : t \geq 0]$ with the following properties [16, Def. 1.8]

- The paths of W_t start at $W_0 = 0$ and are continuous in t.
- The increments are independent, i.e., for $t_0 < t_1 < \cdots < t_n$ the random variables $W_{t_1} - W_{t_0}$, $W_{t_2} - W_{t_1}$,..., $W_{t_n} - W_{t_{n-1}}$ are independent.
- For $0 \leq s \leq t$ the increment $W_t - W_s$ is normally distributed with mean 0 and variance $t - s$.

We already mentioned that $|\mathrm{d}W_t/\mathrm{d}t|$ diverges as $\mathrm{d}t \to 0$ (see Sec. 1.3) such that W_t is with probability one not differentiable in t. One can even show that W_t is with probability one nowhere differentiable, i.e., for no τ with $0 < \tau < t$ [92, Theorem 37.3].

For later use, we here mention some consequences of the definition of the independence properties of the Brownian motion. First, note that W_t is a Markov process. Moreover, for all f and $t_1 \leq t_2 \leq t_3$ we have $\langle f(W_{t_1})(W_{t_3} - W_{t_2})\rangle = \langle f(W_{t_1})\rangle\langle(W_{t_3} - W_{t_2})\rangle = 0$. If we consider the process at discrete points $t_k \equiv k\Delta t$ with $k \in \mathbb{N}$ we have $\langle(W_{t_{k+1}} - W_{t_k})(W_{t_{k'+1}} - W_{t_{k'}})\rangle = \delta_{kk'}\Delta t$ where $\delta_{kk'} = 1$ for $k = k'$ and $\delta_{kk'} = 0$ for $k \neq k'$ is the *Kronecker delta*.

A *d-dimensional Brownian motion* $\mathbf{W}_t = (W_t^1, W_t^2, \ldots, W_t^d)^T$ is defined as a vector of d independent one-dimensional Brownian motions.

[6] Some authors would call this process *weakly* (but not strongly) ergodic since the statistics of $x_{t+\tau} - x_t$ does not depend on x_0 as $t \to \infty$ [98].

2.5 Stochastic calculus

In this section, we present the mathematical framework to correctly address the Langevin equation. We introduce stochastic differential equations, calculation rules for stochastic calculus, and different conventions for stochastic integration.

2.5.1 Definition of stochastic differential equations and stochastic integration

Now that we defined Brownian motion, we turn to its generalization in terms of the Langevin equation, see Sec. 1.2, i.e., we consider Brownian dynamics under the influence of external forces that deterministically depend on the position $\mathbf{x}_t$ (from now on we write multidimensional quantities in boldface) and time t. In general, also the amplitude of the random fluctuations may depend on $\mathbf{x}_t$ and t. As outlined above, Brownian motion is not differentiable, which is why we write the *Langevin equation* in d-dimensional space $\mathbb{R}^d$ in the increment notation as [16, Eq. (3.4)]

$$\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(t)\mathrm{d}\mathbf{W}_t\,, \tag{2.5.1}$$

where $\mathbf{a}$ and $\boldsymbol{\sigma}$ are (real) vector- and matrix-valued functions of $\mathbf{x}_t, t$, respectively. We always assume that $\boldsymbol{\sigma}$ has no eigenvalues equal to 0. For now, we assume that $\boldsymbol{\sigma}$ does not depend on $\mathbf{x}_t$. We later relax this assumption and explain all details that have to be considered in this more general case in Subsec. 2.8. Equation (2.5.1) is called a *stochastic differential equation* [14–16, 99]. From Eq. (2.5.1) we already see that $\mathbf{x}_t$ is a Markov process since the evolution of $\mathbf{x}_t$ for $t_+ \geq t$ does not depend on values $\mathbf{x}_{t_-}$ with $t_- < t$.

We stressed that $\mathrm{d}W_t/\mathrm{d}t$ diverges. However, as $\mathrm{d}t \to 0$ we also have $\mathrm{d}\mathbf{W}_t \to 0$ (in the sense that $\mathbb{P}(|\mathrm{d}\mathbf{W}_t| > \varepsilon) \to 0$ for any $\varepsilon > 0$). The completely correct way to state the dynamics $\mathbf{x}_t$ (which is meant by Eq. (2.5.1)) is therefore

$$\mathbf{x}_t = \mathbf{x}_0 + \int_0^t \mathbf{a}(\mathbf{x}_\tau, \tau)\mathrm{d}\tau + \int_{\tau=0}^{\tau=t} \boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}_\tau\,, \tag{2.5.2}$$

where we now need to specify the definition and interpretation of the stochastic integrals. Note that throughout the thesis we write $\int_{\tau=0}^{\tau=t}$ explicitly to show that integration runs from $\tau = 0$ to $\tau = t$ even though it is not a $\mathrm{d}\tau$-integral. A process $\mathbf{x}_t$ that fulfills Eq. (2.5.2) is called the *strong* solution of the stochastic differential equation [16, Def. 3.2], which is known to be unique for reasonable initial conditions if $\mathbf{a}$ and $\boldsymbol{\sigma}$ do not vary too quickly (for details see, e.g., [16, Theorem 3.1]).

The first integral, i.e., the integral in $\mathrm{d}\tau$, is generally defined as a Lebesgue-Stieltjes integral which is a slight generalization of Riemann and Lebesgue integration. For the purpose of the thesis it suffices to think of the ordinary Riemann integral, where we discretize the time interval into $K \in \mathbb{N}$ steps of equal size Δt such that $\Delta t = \Delta t(K) \equiv t/K$. We then evaluate the integrand $\mathbf{f}_\tau$, which may generally be a random scalar-, vector-, or matrix-valued function of τ, $\mathbf{W}_\tau$ and $\mathbf{x}_\tau$, at discrete times $t_k = t_k(K) \equiv k\Delta t(K)$ with $k = 0, 1, \ldots, K-1$, and define the integral as the regular Riemann integral

$$\begin{aligned}\int_0^t \mathbf{f}_\tau \mathrm{d}\tau &\equiv \lim_{K\to\infty} \sum_{k=0}^{K-1} \mathbf{f}_{t_k}(t_{k+1} - t_k) \\ &= \lim_{K\to\infty} \Delta t(K) \sum_{k=0}^{K-1} \mathbf{f}_{t_k}\,. \end{aligned} \tag{2.5.3}$$

For any such function $\mathbf{f}$ that fulfills the very mild assumption $\langle \int_0^t ||\mathbf{f}_\tau||_2^2 \,\mathrm{d}\tau \rangle < \infty$, the $\mathrm{d}\mathbf{W}_\tau$ integral in Eq. (2.5.2) can be defined analogously as a Riemann sum [16, Eq. (3.13)] (here, vector- or matrix-valued $\mathbf{f}$ are assumed to have appropriate dimensionalities)

$$\int_{\tau=0}^{\tau=t} \mathbf{f}_\tau \cdot \mathrm{d}\mathbf{W}_\tau \equiv \lim_{K\to\infty} \sum_{k=0}^{K-1} \mathbf{f}_{t_k} \cdot (\mathbf{W}_{t_{k+1}} - \mathbf{W}_{t_k})\,. \tag{2.5.4}$$

This is known as the Itô stochastic integral and we will meet other, related but slightly different, stochastic integrals in Subsec. 2.5.3. These definitions fully specify the meaning of a solution $\mathbf{x}_t$ in Eq. (2.5.2) of the stochastic differential equation (2.5.1).

Integration over $\mathrm{d}\mathbf{x}_t$ is defined as sums over $\mathrm{d}t$ and $\mathrm{d}\mathbf{W}_\tau$ integrals according to Eq. (2.5.1), i.e., $\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(t)\mathrm{d}\mathbf{W}_t$, as

$$\int_{\tau=0}^{\tau=t} \mathbf{f}_\tau \cdot \mathrm{d}\mathbf{x}_\tau \equiv \int_0^t \mathbf{f}_\tau \cdot \mathbf{a}(\mathbf{x}_\tau, \tau)\mathrm{d}\tau + \int_{\tau=0}^{\tau=t} \mathbf{f}_\tau \cdot \boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}_\tau\,. \tag{2.5.5}$$

The stochastic integrals in Eqs. (2.5.3), (2.5.4), and (2.5.5) share most properties of standard integrals, like linearity, and that we can usually swap integrals with other integrals, expectation values, limits, and derivatives as described in Subsec. 2.1.4[7].

However, the $\mathrm{d}\mathbf{W}_\tau$ integrals in Eq. (2.5.4) and Eq. (2.5.5) have particular properties not appearing in standard Riemann integrals that we explain in the following two subsections.

[7] To rigorously generalize Subsec. 2.1.4 to stochastic integration, use the *stochastic dominated converge theorem* [99, Proposition 5.8]

2.5.2 Calculation rules for Itô stochastic calculus

We here list some important calculation rules for Itô integrals that will prove useful later.

Average values of stochastic integrals

From the definition of Brownian motion, see Sec. 2.4, we see that the increments $\mathbf{W}_{t_{k+1}} - \mathbf{W}_{t_k}$ in Eq. (2.5.4) are independent of $\mathbf{f}_{t_k}$ such that for $\Delta\tau \geq 0$ always

$$\langle \mathbf{f}_\tau \cdot (\mathbf{W}_{\tau+\Delta\tau} - \mathbf{W}_\tau) \rangle = \langle \mathbf{f}_\tau \rangle \cdot \langle \mathbf{W}_{\tau+\Delta\tau} - \mathbf{W}_\tau \rangle = 0\,, \tag{2.5.6}$$

which gives

$$\left\langle \int_0^t \mathbf{f}_\tau \cdot \mathrm{d}\mathbf{W}_\tau \right\rangle = 0\,. \tag{2.5.7}$$

For $\mathrm{d}\tau$ integrals over functions $\mathbf{f}(\mathbf{x}_\tau)$, we can write in terms of the time-dependent probability density of $\mathbf{x}_\tau$ by interchanging the expecation value and integral that ($\int \mathrm{d}\mathbf{x}$ denotes the volume integral $\int \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_d$)

$$\left\langle \int_0^t \mathbf{f}(\mathbf{x}_t) \mathrm{d}\tau \right\rangle = \int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x}\, \mathbf{f}(\mathbf{x}) p(\mathbf{x}, t)\,. \tag{2.5.8}$$

Calculation rules for stochastic increments

We now explain some calculation rules for the increments $\mathrm{d}t$, $\mathrm{d}\mathbf{W}_t$ and $\mathrm{d}\mathbf{x}_t$ that are justifiable by rigorous arguments on the corresponding stochastic integrals with the mathematical concept of quadratic variation [14,15,99], or equivalently by L_2-norm limits of the integrals, see Eq. (2.5.9) below (see also [6, Sec. 4.2.5]).

The central calculation rules for stochastic integrals in one-dimensional space read "$\mathrm{d}W_t^2 = \mathrm{d}t$" and "$\mathrm{d}W_t^{2+N} = 0$" for $N > 0$ [6]. They are also written as $\mathrm{d}W_t = O(\sqrt{\mathrm{d}t})$ and $\mathrm{d}W_t^N = O(\mathrm{d}t^{N/2})$ where orders of $O(\mathrm{d}t^{3/2})$ or higher can be neglected. To make sense of these expressions one first has to note that they are not meant literally but instead in the sense of the corresponding stochastic integrals. The equation "$\mathrm{d}W_t^2 = \mathrm{d}t$" holds true in terms of $\langle \mathrm{d}W_t^2 \rangle = \mathrm{d}t$ (recall that $\mathrm{d}W_t \equiv W_{t+\mathrm{d}t} - W_t$ has variance $\mathrm{d}t$, see Sec. 2.4) but this does not encompass the whole meaning of the rule. The expression "$\mathrm{d}W_t^2 = \mathrm{d}t$" is by no means an equation as written since $\mathrm{d}W_t^2$ is a random variable and, e.g., we have for any value of $\Delta t > 0$ that $\langle [(W_{t+\Delta t} - W_t)^2/\Delta t - 1]^2 \rangle = 2$ (this holds as a property of the Gaussian distribution, see also [6, Sec. 4.2.5]), so $\mathrm{d}W_t^2 = \mathrm{d}t$ does not hold as an equality for $\Delta t \to \mathrm{d}t$. The actual meaning of the expressions "$\mathrm{d}W_t^2 = \mathrm{d}t$" and "$\mathrm{d}W_t^2 = 0$" can be

phrased in a well-defined sense as follows (where the occurring integrals are defined as Riemann sums as in Eqs. (2.5.3) and (2.5.4), and f_τ can again be a function of τ, W_τ and x_τ) [6, Sec. 4.2.5],

$$\int_{\tau=0}^{\tau=t} (\mathrm{d}W_\tau)^{2+N} f_\tau = \begin{cases} \int_0^t f_\tau \mathrm{d}\tau & \text{for } N = 0 \\ 0 & \text{for } N > 0 \end{cases} . \tag{2.5.9}$$

Here, equality means that the limits agree in the L^2-norm, i.e., if we denote the discretized version (see Eqs. (2.5.3) and (2.5.4)) of the integrals as $I_{1,2}(\Delta t)$ then $\lim_{\Delta t\to 0}\langle[I_1(\Delta t) - I_2(\Delta t)]^2\rangle = 0$. The fact that orders higher than $\mathrm{d}W_t^2$, i.e., higher than $\mathrm{d}t$, can be neglected corresponds to the fact that the summation in the discretized integrals contain a number of $K = t/\Delta t \to t/\mathrm{d}t$ terms, such that terms of order $\mathrm{d}t$ give finite results after summation while terms of higher order vanish in the limit $\Delta t \to 0$.

For a multidimensional Brownian motion the i-th and j-th components are independent for $i \neq j$, and the calculation rules are "$\mathrm{d}W_t^i \mathrm{d}W_t^j = \delta_{ij}\mathrm{d}t$", and that products of more than two (same or different) components of $\mathrm{d}\mathbf{W}_t$ always give vanishing integrals.

Itô's isometry

What the calculation rules above do not cover are products of increments at different times, i.e., $\mathrm{d}W_{t_1}\mathrm{d}W_{t_2}$ at $t_1 \neq t_2$. In such a case we use *Itô's isometry* [16, Eq. (3.18)], that states in one-dimensional space for functions of t_i, W_{t_i}, x_{t_i} denoted by f_{t_i} as above

$$\left\langle \int_{t_1=0}^{t_1=t} f_{t_1}\mathrm{d}W_{t_1} \int_{t_2=0}^{t_2=t} g_{t_2}\mathrm{d}W_{t_2} \right\rangle = \left\langle \int_0^t f_\tau g_\tau \mathrm{d}\tau \right\rangle . \tag{2.5.10}$$

If the details of the x_t process are known, the right-hand side can be evaluated using Eq. (2.5.8). Equation (2.5.10) can be directly proven from the discrete-time definition in Eq. (2.5.4) using $\langle(W_{t_{k+1}} - W_{t_k})(W_{t_{k'+1}} - W_{t_{k'}})\rangle = \delta_{kk'}\Delta t$, together with the fact that $W_{t_{\max(k,k')+1}} - W_{t_{\max(k,k')}}$ is independent of $f_{t_k}g_{t_{k'}}$. Itô's isometry can be phrased as the calculation rule $\langle \mathrm{d}W_{t_1}\mathrm{d}W_{t_2}\rangle = \delta(t_1 - t_2)\mathrm{d}t_1\mathrm{d}t_2$ which has to be understood as a rule for double integrals as in Eq. (2.5.10). This delta-correlation property of the noise increments is also known as the *white-noise* property [6]. Note that unlike the calculation rules above, this only holds on average, i.e., for mean values of such double integrals.

For double integrals involving $\mathrm{d}t_1\mathrm{d}t_2$ there is no such calculation rule needed as they are already in the simplest form and expectation values of those are directly accessible if the joint density of x_{t_1} and x_{t_2} at different times is known (for details see Ch. 3).

Averages of double integrals involving $\mathrm{d}W_{t_1}\mathrm{d}t_2$ are significantly more challenging, and exceed the textbook knowledge presented here. Namely, for $t_2 > t_1$ we cannot argue with independence properties of $\mathrm{d}W_{t_2}$ such that non-trivial terms remain in the integration. This will be the main mathematical challenge addressed in Ch. 3.

Itô's lemma

Another consequence of $\mathrm{d}W_t^2 = O(\mathrm{d}t)$ is a correction term to the conventional chain rule of differentiation $\mathrm{d}f(x(t))/\mathrm{d}t = f'(x)\mathrm{d}x/\mathrm{d}t$. The adapted chain rule for stochastic calculus is known as *Itô's lemma*, that states for a twice-differentiable function $f(W_t)$ with first and second derivatives f' and f'' that [6, Sec. 4.3.3]

$$\mathrm{d}f(W_t) = f'(W_t)\mathrm{d}W_t + \frac{1}{2}f''(W_t)\mathrm{d}t\,. \tag{2.5.11}$$

This can be thought of as the Taylor expansion for $f(W_t + \mathrm{d}W_t)$ up to second order with $\mathrm{d}W_t^2 = \mathrm{d}t$. For the example $f(x) = x^2$ with $f'(x) = 2x$ and $f''(x) = 2$ written out in terms of integrals, Eq. (2.5.11) states

$$W_t^2 - W_0^2 \equiv \int_{\tau=0}^{\tau=t} \mathrm{d}(W_\tau^2) = 2\int_{\tau=0}^{\tau=t} W_\tau \mathrm{d}W_\tau + \int_0^t \mathrm{d}\tau\,. \tag{2.5.12}$$

Using that by definition $W_0^2 = 0$ this gives $2\int_{\tau=0}^{\tau=t} W_\tau \mathrm{d}W_\tau = W_t^2 - t$ which agrees with the result obtained by directly computing $\int_{\tau=0}^{\tau=t} W_\tau \mathrm{d}W_\tau$ from the properties of W_t and the definition of the $\mathrm{d}W_\tau$ integral in Eq. (2.5.4), see [6, Sec. 4.2.2]. Itô's lemma generalizes to multiple dimensions or functions of multiple components exactly according to the notion of the Taylor series up to second order, alongside the calculation rules for increments. For a scalar, twice differentiable function of stochastic dynamics $\mathbf{x}_t$ specified by Eq. (2.5.1) in d-dimensional space it therefore reads (throughout the thesis, curly brackets $\{\cdots\}$ denote that derivatives only act inside the brackets, and not further to the right)

$$\mathrm{d}f(\mathbf{x}_t) = \{\nabla f(\mathbf{x}_t)\}\cdot \mathrm{d}\mathbf{x}_t + \frac{1}{2}\sum_{i,j=1}^{d}\{\partial_i\partial_j f(\mathbf{x}_t)\}\mathrm{d}x_t^i \mathrm{d}x_t^j\,. \tag{2.5.13}$$

Using for the components of $\mathrm{d}x_t^i = \sum_{i'}\sigma_{ii'}(t)\mathrm{d}W_t^{i'} + O(\mathrm{d}t)$ the calculation rule $\mathrm{d}W_t^i\mathrm{d}W_t^j = \delta_{ij}\mathrm{d}t$ as well as that higher orders vanish, we get $\mathrm{d}x_t^i\mathrm{d}x_t^j = \sum_{i',j'}\sigma_{ii'}\sigma_{jj'}\delta_{i'j'}\mathrm{d}t = \sum_k \sigma_{ik}\sigma_{jk}\mathrm{d}t = 2D_{ij}\mathrm{d}t$ where $\mathbf{D}(t) \equiv \boldsymbol{\sigma}(t)\boldsymbol{\sigma}^T(t)/2$ is the *diffusion coefficient*. Using that

we assumed $\mathbf{D}(t)$ to not depend on $\mathbf{x}_t$ we can commute $\mathbf{D}$ through the derivatives to obtain Itô's lemma in the useful form

$$\mathrm{d}f(\mathbf{x}_t) = \mathbf{a}(\mathbf{x}_t, t) \cdot \nabla f(\mathbf{x}_t)\mathrm{d}t + \nabla \cdot \mathbf{D}(t)\nabla f(\mathbf{x}_t)\mathrm{d}t + \nabla f(\mathbf{x}_t) \cdot \boldsymbol{\sigma}(t)\mathrm{d}\mathbf{W}_t\,. \tag{2.5.14}$$

2.5.3 Different conventions for stochastic integration

Although the definitions of the $\mathrm{d}\mathbf{W}_\tau$ and $\mathrm{d}\mathbf{x}_\tau$ integrals in Eqs. (2.5.4) and (2.5.5) are completely sufficient to introduce and define the stochastic dynamics described by the stochastic differential equation in Eq. (2.5.1), there are also other useful definitions of stochastic integration. The alternative definitions can be traced back to $\mathrm{d}\mathbf{W}_\tau$ being of order $\sqrt{\mathrm{d}t}$ such that products of increments in integrals do not vanish.

Different conventions for $\mathrm{d}\mathbf{W}_\tau$ integrals

We extend the definition of the $\mathrm{d}\mathbf{W}_\tau$ integral from Eq. (2.5.4) for scalar-, vector-, or matrix-valued functions $\mathbf{f}(\mathbf{W}_t)$ to a definition depending on a parameter $\lambda \in [0, 1]$ [16, Eq. (3.13)]

$$\int_0^t \mathbf{f}(\mathbf{W}_\tau) \cdot \mathrm{d}_\lambda \mathbf{W}_\tau \equiv \lim_{K\to\infty} \sum_{k=0}^{K-1} \mathbf{f}[(1-\lambda)\mathbf{W}_{t_k} + \lambda \mathbf{W}_{t_{k+1}}] \cdot (\mathbf{W}_{t_{k+1}} - \mathbf{W}_{t_k})\,. \tag{2.5.15}$$

To understand this definition, think of the regular Riemann integral where this would correspond to

$$\begin{aligned}\int_0^t \mathbf{f}(\tau) \cdot \mathrm{d}_\lambda \tau &\equiv \lim_{K\to\infty} \sum_{k=0}^{K-1} \mathbf{f}[(1-\lambda)t_k + \lambda t_{k+1}](t_{k+1} - t_k) \\ &= \lim_{K\to\infty} \sum_{k=0}^{K-1} \mathbf{f}[t_k + \lambda\Delta t(K)]\Delta t(K)\,,\end{aligned} \tag{2.5.16}$$

i.e., $\lambda = 0$ and $\lambda = 1$ correspond to evaluating $\mathbf{f}$ at left and right edges of the interval, respectively, and $\lambda \in (0, 1)$ linearly interpolates between these[8]. For simplicity, we first consider one-dimensional notation. For functions with derivative f', we use for $\Delta t \to 0$ that $f[t_k + \lambda\Delta t(K)]\Delta t(K) = f(t_k)\Delta t(K) + \lambda f'(t_k)\Delta t(K)^2 + O(\Delta t(K)^3)$. As $\Delta t \to 0$, the summation over k in Eq. (2.5.16) only cancels an order of Δt^1 such that higher orders

[8] Considering the notion of evaluating $\mathbf{f}$ at intermediate points in the time interval, one might be tempted to define Eq. (2.5.15) with $\mathbf{W}_{t_k+\lambda\Delta t}$ instead of $(1-\lambda)\mathbf{W}_{t_k} + \lambda\mathbf{W}_{t_{k+1}}$ which should in theory give the same result. However, this alternative definition is less preferable since in both, theory and practice (i.e., in measurements), we approach the integral by values at the discrete times t_k and $\mathbf{W}_{t_k+\lambda\Delta t}$ is not included in this discrete set.

vanish. This shows that the limit $K \to \infty$ in Eq. (2.5.16) is independent of λ such that it always agrees with the standard definition ($\lambda = 0$) in Eq. (2.5.3).

This is different for the $\mathrm{d}\mathbf{W}_\tau$ integral defined in Eq. (2.5.15). The same approach here gives, using the calculation rule $\mathrm{d}W_t^2 = \mathrm{d}t$ as stated in Eq. (2.5.9) (we generalize to multidimensional space below in Eqs. (2.5.20)-(2.5.22)),

$$\begin{aligned}\int_0^t f(W_\tau)\mathrm{d}_\lambda W_\tau - \int_0^t f(W_\tau)\mathrm{d}W_\tau &= \int_0^t \left[f'(W_\tau)\lambda \mathrm{d}W_\tau^2 + O(\mathrm{d}W_\tau^3)\right] \\ &= \lambda \int_0^t f'(W_\tau)\mathrm{d}\tau\,. \end{aligned} \tag{2.5.17}$$

Therefore, the choice of λ in Eq. (2.5.15) unlike in Eq. (2.5.16) does make a difference. The most relevant choices are the *Itô* convention ($\lambda = 0$, i.e., as in Eq. (2.5.4)), the *Stratonovich* convention ($\lambda = 1/2$, from now on denoted by $\circ\mathrm{d}W_t$), and the *Klimontovich* or *anti-Itô* convention ($\lambda = 1$). Equation (2.5.17) shows how the different conventions for λ can be rephrased in terms of the Itô integral, for which we can use the calculation rules of the previous subsection.

Often, the Itô convention is most practical since $\mathrm{d}W_t$ is independent of $f(W_t)$, but not of $f[(1-\lambda)W_{t_k} + \lambda W_{t_{k+1}}]$, which implies that calculation rules as $\langle \int f(W_t)\mathrm{d}W_t\rangle = 0$ and Itô's isometry in Eq. (2.5.10) only hold for the Itô convention. At the same time, the Stratonovich convention $\lambda = 1/2$ has the advantage that it behaves more natural under time-reversal, since applying the definition in Eq. (2.5.15) backwards maps λ to $1-\lambda$. Most importantly, the Stratonovich convention has the advantage that the stochastic chain rule, i.e., Itô's lemma $\mathrm{d}f(W_t) = f'(W_t)\mathrm{d}W_t + \frac{1}{2}f''(W_t)\mathrm{d}t$ in Eq. (2.5.11), according to Eq. (2.5.17) applied to f' simplifies to

$$\mathrm{d}f(W_t) = f'(W_t) \circ \mathrm{d}W_t\,. \tag{2.5.18}$$

In other words, the Stratonovich integral fulfills the fundamental theorem of calculus $\int_{\tau=0}^{\tau=t} f'(W_\tau) \circ \mathrm{d}W_\tau = f(W_t) - f(W_0)$.

Revisiting the example in Eq. (2.5.12) with the Stratonovich chain rule Eq. (2.5.18) for $f(x) = x^2$, we obtain that

$$2\int_{\tau=0}^{\tau=t} W_\tau \circ \mathrm{d}W_\tau = W_t^2 - W_0^2\,. \tag{2.5.19}$$

The equivalence to the Itô result in Eq. (2.5.12) is given by Eq. (2.5.17) with $\lambda = 1/2$ since $\int_{\tau=0}^{\tau=t} W_\tau \circ \mathrm{d}W_\tau = \int_{\tau=0}^{\tau=t} W_\tau \mathrm{d}W_\tau + \int_0^t \mathrm{d}\tau/2$. Note that, on the one hand, Eq. (2.5.19) is

simplified compared to the Itô result in Eq. (2.5.12), but, on the other hand, $\langle \int_0^t W_\tau \mathrm{d}W_\tau \rangle = 0$ is simpler than the non-trivial $\langle \int_0^t W_\tau \circ \mathrm{d}W_\tau \rangle = t/2 \neq 0$.

The above calculation rules directly generalize to multidimensional space, i.e.,

$$\int_0^t f(\mathbf{W}_\tau)\mathrm{d}_\lambda \mathbf{W}_\tau - \int_0^t f(\mathbf{W}_\tau)\mathrm{d}\mathbf{W}_\tau = \lambda \int_0^t \nabla f(\mathbf{W}_\tau)\mathrm{d}\tau\,, \tag{2.5.20}$$

and for d-dimensional vector-valued $\mathbf{f}$,

$$\int_0^t \mathbf{f}(\mathbf{W}_\tau)\cdot \mathrm{d}_\lambda \mathbf{W}_\tau - \int_0^t \mathbf{f}(\mathbf{W}_\tau)\cdot \mathrm{d}\mathbf{W}_\tau = \lambda \int_0^t \nabla\cdot\mathbf{f}(\mathbf{W}_\tau)\mathrm{d}\tau\,. \tag{2.5.21}$$

Moreover, we have for the Stratonovich integral with scalar product denoted by "$\cdot \circ \mathrm{d}\mathbf{W}_t$" that

$$\mathrm{d}f(\mathbf{W}_t) = \nabla f(\mathbf{W}_t)\cdot \circ \mathrm{d}\mathbf{W}_t\,. \tag{2.5.22}$$

Different conventions for $\mathrm{d}\mathbf{x}_t$ integrals

The conventions for $\mathrm{d}\mathbf{W}_\tau$ integrals directly translate to functions of $\mathbf{x}_\tau$ and $\mathrm{d}\mathbf{x}_\tau$ integrals for $\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t,t)\mathrm{d}t + \boldsymbol{\sigma}(t)\mathrm{d}\mathbf{W}_t$ as in Eq. (2.5.1), since the convention can be computed as above for the $\mathrm{d}\mathbf{W}_\tau$ integral and does not alter the $\mathrm{d}\tau$ integral as in Eq. (2.5.16). Using $\mathrm{d}\mathbf{x}_t = \boldsymbol{\sigma}(t)\mathrm{d}\mathbf{W}_t + O(\mathrm{d}t)$ and for the i,j-th components the calculation rule "$\mathrm{d}W_t^i \mathrm{d}W_t^j = \delta_{ij}\mathrm{d}t$", we compute in analogy to above

$$\begin{aligned}\int_{\tau=0}^{\tau=t} f(\mathbf{x}_\tau)\mathrm{d}_\lambda \mathbf{W}_\tau - \int_{\tau=0}^{\tau=t} f(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau &= \lambda \int_0^t [\nabla f(\mathbf{x}_\tau)\cdot \mathrm{d}\mathbf{x}_\tau]\,\mathrm{d}\mathbf{W}_\tau \\ &= \lambda \int_0^t [\nabla f(\mathbf{x}_\tau)\cdot \boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}_\tau]\,\mathrm{d}\mathbf{W}_\tau \\ &= \lambda \int_0^t \boldsymbol{\sigma}^T(\tau)\nabla f(\mathbf{x}_\tau)\mathrm{d}\tau\,. \end{aligned} \tag{2.5.23}$$

For $\mathrm{d}\mathbf{x}_t$ integrals, we use that "$\mathrm{d}W_t^i \mathrm{d}W_t^j = \delta_{ij}\mathrm{d}t$" gives $[\nabla f\cdot\boldsymbol{\sigma}\mathrm{d}\mathbf{W}_\tau]\boldsymbol{\sigma}\mathrm{d}\mathbf{W}_\tau = \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_\tau[\mathrm{d}\mathbf{W}_\tau\cdot \boldsymbol{\sigma}^T\nabla f] = \boldsymbol{\sigma}\boldsymbol{\sigma}^T\nabla f = 2\mathbf{D}\nabla f$ to obtain

$$\begin{aligned}\int_{\tau=0}^{\tau=t} f(\mathbf{x}_\tau)\mathrm{d}_\lambda \mathbf{x}_\tau - \int_{\tau=0}^{\tau=t} f(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau &= \lambda \int_0^t [\nabla f(\mathbf{x}_\tau)\cdot \mathrm{d}\mathbf{x}_\tau]\,\mathrm{d}\mathbf{x}_\tau \\ &= \lambda \int_0^t [\nabla f(\mathbf{x}_\tau)\cdot \boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}_\tau]\,\boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}_\tau \\ &= 2\lambda \int_0^t \mathrm{d}\tau \mathbf{D}(\tau)\nabla f(\mathbf{x}_\tau)\,. \end{aligned} \tag{2.5.24}$$

For vector-valued functions we obtain (where $\mathbf{e}_i$ denotes the i-th unit vector)

$$\begin{aligned}\int_{\tau=0}^{\tau=t} \mathbf{f}(\mathbf{x}_\tau)\cdot \mathrm{d}_\lambda \mathbf{x}_\tau - \int_{\tau=0}^{\tau=t} \mathbf{f}(\mathbf{x}_\tau)\cdot \mathrm{d}\mathbf{x}_\tau &= \lambda \sum_{i=1}^{d} \int_0^t [\nabla f_i(\mathbf{x}_\tau)\cdot \mathrm{d}\mathbf{x}_\tau]\, \mathrm{d}x_\tau^i \mathbf{e}_i \\ &= \lambda \sum_{i,j,k,l=1}^{d} \int_0^t \partial_j f_i(\mathbf{x}_\tau)\sigma_{jk}(\tau)\mathrm{d}W_\tau^k \sigma_{il}(\tau)\mathrm{d}W_\tau^l \mathbf{e}_i \\ &= \lambda \sum_{i,j,k=1}^{d} \int_0^t \mathrm{d}\tau \partial_j f_i(\mathbf{x}_\tau)\sigma_{jk}(\tau)\sigma_{ik}(\tau)\mathbf{e}_i \\ &= 2\lambda \int_0^t \nabla\cdot \mathbf{D}(\tau)\mathbf{f}(\mathbf{x}_\tau)\,. \end{aligned} \tag{2.5.25}$$

The Stratonovich correction term cancels the second order term in Itô's lemma also for $\mathrm{d}\mathbf{x}_t$ integrals, such that the Stratonovich chain rule generalizes as [16]

$$\mathrm{d}f(\mathbf{x}_t) = \nabla f(\mathbf{x}_t)\cdot \circ \mathrm{d}\mathbf{x}_t\,. \tag{2.5.26}$$

2.5.4 On different notations and approaches

To conclude this section on stochastic calculus, we briefly mention a different, commonly used notation. Often, the random part in the Langevin equation, denoted by $\mathrm{d}\mathbf{W}_t$ here, is written as $\boldsymbol{\xi}(t)\mathrm{d}t$, where $\boldsymbol{\xi}(t)$ is the *white noise* $\mathrm{d}\mathbf{W}_t/\mathrm{d}t$. With this, the Langevin equation (2.5.1) can be written as $\dot{\mathbf{x}}(t) = \mathbf{a}[(\mathbf{x}(t),t)] + \boldsymbol{\sigma}(t)\boldsymbol{\xi}(t)$. As mentioned in Sec. 1.2 and Subsec. 2.5.1, this notation is mathematically misleading since $\boldsymbol{\xi}(t) = \mathrm{d}\mathbf{W}_t/\mathrm{d}t$ formally diverges as[9] $\mathrm{d}t \to 0$. However, the notation with $\boldsymbol{\xi}(t)$ is not *per se* wrong since it generally should be understood in the sense of $\mathrm{d}\mathbf{W}_t$ in Eq. (2.5.1), or even more precisely in the integrated sense as in Eq. (2.5.2).

To give a concrete example where ambiguities due to sloppy notation might occur, consider the simple case of a one-dimensional Brownian motion in the notation $\dot{x}(t) = \xi(t)$ with the white noise property $\langle \xi(t)\xi(\tau)\rangle = \delta(t-\tau)$, see Sec. 1.2, i.e., in this example $\langle \dot{x}(t)\dot{x}(\tau)\rangle = \delta(t-\tau)$. To be able to apply the standard chain rule in Eq. (2.5.26) one often interprets integrals written as $\int_0^t f[x(\tau)]\dot{x}(\tau)\mathrm{d}\tau$ in the Stratonovich sense (although the clarified notation $\int_0^t f[x(\tau)]\circ\dot{x}(\tau)\mathrm{d}\tau$ is also used). However, when writing $\langle \int_0^t f_1[x(\tau)]\dot{x}(\tau)\mathrm{d}\tau \int_0^t f_2[x(\tau')]\dot{x}(\tau')\mathrm{d}\tau'\rangle = \int_0^t f_1[x(\tau)]f_2[x(\tau)]\mathrm{d}\tau$ using independence properties and $\langle \xi(t)\xi(\tau)\rangle = \delta(t-\tau)$ for the Brownian motion, one necessarily has to interpret the stochastic integrals in the Itô sense.

[9] In particular, one should rather refrain from plotting $\xi(t)$ (or only with explicitly stating its dependence on $\mathrm{d}t > 0$) since doing so misleadingly suggests that it exists.

On another note, in writing a formal expression for the path probability density of Langevin trajectories one often encounters the notation $\dot{x}(t)$ that denotes the derivative of a non-differentiable path. E.g., for one-dimensional Brownian motion with constant drift, $\dot{x}(t) = b + \xi(t)$, the *Freidlin-Wentzell* or *Onsager-Machlup path measure* is often written as $\mathbb{P}([x(\tau)]_{0\leq\tau\leq t}) \propto \exp(-\int_0^t [\dot{x}(\tau) - b]^2 \mathrm{d}\tau/4)$. Since $|\mathrm{d}x_t/\mathrm{d}t|$ diverges as $\mathrm{d}t \to 0$ this expression either has to be understood in discretized time, or as a means to evaluate the probability that the trajectory tightly follows a smooth path $[y(\tau)]_{0\leq\tau\leq t}$ where $\dot{x}(\tau)$ is substituted by the (actually existing) derivative $\dot{y}(\tau)$, see [14, Theorem 9.1]. From logarithms of ratios of the path measure one may also arrive at $\dot{x}(\tau)\mathrm{d}\tau$ integrals where it is not *a priori* clear whether these should be interpreted in the Stratonovich sense (as consistent with the approximation by $\dot{y}(\tau)\mathrm{d}\tau$ and required to compute the total entropy production from the path measure [54, Eq. (14)], see also [59]) or Itô sense (as required, e.g., for the derivation of the thermodynamic uncertainty relation via the Cramér-Rao inequality in Ref. [70]).

Although these are by no means unresolvable issues, they should highlight that the stochastic calculus approach emphasized in this thesis, in particular, bypassing $\dot{x}$ and the path measure, represents a direct route to avoid confusions in this mathematically challenging field.

2.6 The Fokker-Planck equation

So far, we described the stochastic dynamics of individual paths in terms of the Langevin equation (2.5.1). Alternatively, one can describe the stochastic properties and the time-evolution of the dynamics in terms of the evolution of the probability density function $p(\mathbf{x}, t)$ over an ensemble of paths. In this section, the equation governing the time evolution $\partial_t p(\mathbf{x}, t)$ is derived, and selected properties of this equation are discussed.

2.6.1 Derivation of the Fokker-Planck equation

Starting from the Langevin dynamics in Eq. (2.5.1), we derive the *Fokker-Planck equation* for the time evolution of the probability density $p(\mathbf{x}, t)$ using the calculation rules presented above (see also [6, Sec. 4.3.4]). Even though there are different ways to derive this equation (e.g., the *Kramers-Moyal expansion* [6]), we prefer the derivation from the Langevin equation to highlight that all information on the stochastic dynamics is contained in Eq. (2.5.1). Moreover, this derivation is very simple and direct if the methods developed in the previous section are known.

Consider a twice-differentiable test function f that obeys the appropriate boundary conditions (i.e., vanishes at the boundaries for infinite space and is periodic for periodic boundary conditions). We can write the time evolution of an expectation value as (see Subsec. 2.1.4 for interchanging derivative and integral)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle f(\mathbf{x}_t)\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}\mathbf{x} f(\mathbf{x})p(\mathbf{x},t) = \int \mathrm{d}\mathbf{x} f(\mathbf{x})\partial_t p(\mathbf{x},t)\,. \tag{2.6.1}$$

Alternatively, we may also swap derivative and expectation value (see Subsec. 2.1.4) and use Itô's lemma in the form Eq. (2.5.14), and $\langle g(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau\rangle = 0$, see Eq. (2.5.6), to obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle f(\mathbf{x}_t)\rangle &= \left\langle \frac{\mathrm{d}f(\mathbf{x}_t)}{\mathrm{d}t}\right\rangle \\ &= \langle \mathbf{a}(\mathbf{x}_t,t)\cdot\nabla f(\mathbf{x}_t) + \nabla\cdot\mathbf{D}(t)\nabla f(\mathbf{x}_t)\rangle \\ &= \int \mathrm{d}\mathbf{x} p(\mathbf{x},t)\left[\mathbf{a}(\mathbf{x},t)\cdot\nabla f(\mathbf{x}) + \nabla\cdot\mathbf{D}(t)\nabla f(\mathbf{x})\right]\,.\end{aligned} \tag{2.6.2}$$

Integrating by parts, using $\mathbf{D}^T \equiv (\boldsymbol{\sigma}\boldsymbol{\sigma}^T/2)^T = \mathbf{D}$ and that f obeys appropriate boundary conditions, gives (note that derivatives are always, except in curly brackets (e.g., $\{\partial_x f(x)\}$), understood to act all the way to the right, i.e., $-\nabla$ acts on $\mathbf{a}(\mathbf{x},t)$ and $p(\mathbf{x},t)$)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle f(\mathbf{x}_t)\rangle = \int \mathrm{d}\mathbf{x} f(\mathbf{x})\left[-\nabla\cdot\mathbf{a}(\mathbf{x},t) + \nabla\cdot\mathbf{D}(t)\nabla\right]p(\mathbf{x},t)\,. \tag{2.6.3}$$

Since Eqs. (2.6.1) and (2.6.3) agree for all admissible choices of f, we conclude the *Fokker-Planck equation* [6, 17]

$$\partial_t p(\mathbf{x},t) = -\nabla\cdot\mathbf{a}(\mathbf{x},t)p(\mathbf{x},t) + \nabla\cdot\mathbf{D}(t)\nabla p(\mathbf{x},t)\,. \tag{2.6.4}$$

Note that this equation only depends on $p(\mathbf{x},t)$ and not, e.g., on $p(\mathbf{x},\tau)$ for $\tau < t$. This shows again that the considered dynamics $\mathbf{x}_t$ is a Markov process, such that the theory developed in Subsec. 2.3.1, e.g., the Chapman-Kolmogorov equation (2.3.3), applies to $p(\mathbf{x},t)$ obtained from Eq. (2.6.4).

For later use, we define based on Eq. (2.6.4) the *Fokker-Planck operator* $\hat{L}(\mathbf{x},t)$, its adjoint operator $\hat{L}^\dagger(\mathbf{x},t)$ (also called *generator* of the Markov process [16, Sec. 2.3]), the *probability density current* or *probability current density* $\mathbf{j}(\mathbf{x},t)$, and the Fokker-Planck equation in the divergence form $\partial_t p(\mathbf{x},t) = -\nabla\cdot\mathbf{j}(\mathbf{x},t)$ as

$$\begin{aligned}\partial_t p(\mathbf{x},t) &= \hat{L}(\mathbf{x},t)p(\mathbf{x},t) \\ \hat{L}(\mathbf{x},t) &\equiv -\nabla\cdot\mathbf{a}(\mathbf{x},t) + \nabla\cdot\mathbf{D}(t)\nabla\end{aligned}$$

$$\begin{aligned}\hat{L}^\dagger(\mathbf{x},t) &\equiv \mathbf{a}(\mathbf{x},t)\cdot\nabla + \nabla\cdot\mathbf{D}(t)\nabla \\ \mathbf{j}(\mathbf{x},t) &\equiv [\mathbf{a}(\mathbf{x},t) - \mathbf{D}(t)]p(\mathbf{x},t) \\ \partial_t p(\mathbf{x},t) &= -\nabla\cdot\mathbf{j}(\mathbf{x},t)\,.\end{aligned} \tag{2.6.5}$$

Since $\hat{L}(\mathbf{x},t)$ is independent of $\mathbf{x}_0$, the Fokker-Planck equation equally applies to the transition density, i.e.,

$$\partial_t p(\mathbf{x},t|\mathbf{x}_0) = \hat{L}(\mathbf{x},t)p(\mathbf{x},t|\mathbf{x}_0)\,. \tag{2.6.6}$$

In the absence of a drift term, i.e., for $\mathbf{a}(\mathbf{x},t)=\mathbf{0}$, the Fokker-Planck equation simplifies to the *diffusion equation* $\partial_t p(\mathbf{x},t) = \nabla\cdot\mathbf{D}(t)\nabla p(\mathbf{x},t)$.

Often, we consider the time-homogeneous case, where $\mathbf{a}$ and $\mathbf{D}$ do not explicitly depend on time t. There, $p(\mathbf{x},t)$ will generally depend on time but $\hat{L}(\mathbf{x},t) = \hat{L}(\mathbf{x})$ is time homogeneous, and, by solving $\partial_t p(\mathbf{x},t) = \hat{L}(\mathbf{x})p(\mathbf{x},t)$ for $p(\mathbf{x},0|\mathbf{x}_0) = \delta(\mathbf{x}-\mathbf{x}_0)$, the transition function $p(\mathbf{x},t|\mathbf{x}_0)$ can be formally expressed as

$$\mathrm{e}^{\hat{L}(\mathbf{x})t}\delta(\mathbf{x}-\mathbf{x}_0) = p(\mathbf{x},t|\mathbf{x}_0)\,. \tag{2.6.7}$$

The *time-evolution operator* or *propagator* $\mathrm{e}^{\hat{L}(\mathbf{x})t}$ is said to *propagate* the final state $\mathbf{x}$. By repeatedly using $\partial_{x^i}\delta(\mathbf{x}-\mathbf{x}_0) = -\partial_{x_0^i}\delta(\mathbf{x}-\mathbf{x}_0)$, Eq. (2.6.7) implies that the adjoint operator $\hat{L}^\dagger(\mathbf{x}_0)$, see Eq. (2.6.5), propagates initial condition $\mathbf{x}_0$, i.e.,

$$\mathrm{e}^{\hat{L}^\dagger(\mathbf{x}_0)t}\delta(\mathbf{x}-\mathbf{x}_0) = p(\mathbf{x},t|\mathbf{x}_0)\,. \tag{2.6.8}$$

2.6.2 Average values of Stratonovich $\mathrm{d}\mathbf{x}_\tau$ integrals

We stated calculation rules for average values of stochastic integrals in Eqs. (2.5.7) and (2.5.8). Based on these equations, we can state another useful expression in terms of the probability density current $\mathbf{j}(\mathbf{x},t)$ defined based on the Fokker-Planck equation in Eq. (2.6.5) as follows. From the Stratonovich correction term in Eq. (2.5.24) and an integration by parts using $\mathbf{D}^T = \mathbf{D}$ we that

$$\begin{aligned}\left\langle \int_{\tau=0}^{\tau=t} f(\mathbf{x}_\tau,\tau)\circ\mathrm{d}\mathbf{x}_\tau \right\rangle &= \left\langle \int_{\tau=0}^{\tau=t} f(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{x}_\tau + \int_0^t \mathrm{d}\tau\mathbf{D}(\tau)\nabla f(\mathbf{x}_\tau,\tau)\right\rangle \\ &= \int \mathrm{d}\mathbf{x}\int_0^t \mathrm{d}\tau f(\mathbf{x},\tau)\left[\mathbf{a}(\mathbf{x},\tau) - \mathbf{D}(\tau)\nabla\right]p(\mathbf{x},\tau) \\ &= \int \mathrm{d}\mathbf{x}\int_0^t \mathrm{d}\tau f(\mathbf{x}_\tau,\tau)\mathbf{j}(\mathbf{x},\tau)\,,\end{aligned} \tag{2.6.9}$$

and equivalently using Eq. (2.5.25) for vector-valued $\mathbf{f}$ that

$$\left\langle \int_{\tau=0}^{\tau=t} \mathbf{f}(\mathbf{x}_\tau, \tau) \cdot \circ \mathrm{d}\mathbf{x}_\tau \right\rangle = \int \mathrm{d}\mathbf{x} \int_0^t \mathrm{d}\tau \mathbf{f}(\mathbf{x}, \tau) \cdot \mathbf{j}(\mathbf{x}, \tau)\,. \tag{2.6.10}$$

These equations show that the average values of Stratonovich displacements $\circ\mathrm{d}\mathbf{x}_\tau$ for Langevin dynamics are described by the current $\mathbf{j}(\mathbf{x}, t)$ in the Fokker-Planck equation.

2.7 Equilibrium and non-equilibrium steady states

As sketched in Sec. 1.4, a particularly interesting class of Langevin dynamics are steady-state dynamics, emerging for confined Langevin motion $\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t$ with time-homogeneous drift and diffusion coefficients. Such processes approach an invariant density $p_\mathrm{s}(\mathbf{x})$ and are ergodic, see Subsec. 2.3.3.

2.7.1 Equilibrium steady states

First, we introduce equilibrium steady states and some important properties. Afterwards, we turn to the non-equilibrium case.

Equilibrium Boltzmann-Gibbs density

In the simplest case, consider the drift $\mathbf{a}(\mathbf{x})$ to arise from a potential $\phi(\mathbf{x})$[10], such that the force $-\nabla\phi(\mathbf{x})$ translates via the inverse mobility with Einstein relation $\boldsymbol{\gamma}^{-1} = \mathbf{D}/k_\mathrm{B}T$, see Eq. (1.3.2), to a drift $\mathbf{a}(\mathbf{x}) = -\mathbf{D}\nabla\phi(\mathbf{x})/k_\mathrm{B}T$. The Fokker-Planck equation (2.6.5) for this motion reads

$$\begin{aligned} \partial_t p(\mathbf{x}, t) &= \hat{L}(\mathbf{x}) p(\mathbf{x}, t) = -\nabla \cdot \mathbf{j}(\mathbf{x}, t) \\ \mathbf{j}(\mathbf{x}, t) &= -\mathbf{D}\left[\{\nabla\phi(\mathbf{x})\}/k_\mathrm{B}T + \nabla\right] p(\mathbf{x}, t)\,, \end{aligned} \tag{2.7.1}$$

where curly brackets denote that the derivative only acts on U. From Eq. (2.7.1), we see that $\mathbf{j}(\mathbf{x}) = \mathbf{0}$ if, and only if, $p(\mathbf{x}) \propto \exp[-\phi(\mathbf{x})/k_\mathrm{B}T]$. Therefore, the normalized density $p_\mathrm{s}(\mathbf{x})$ obeys

$$\begin{aligned} Z &\equiv \int \mathrm{d}\mathbf{x}' \exp[-\phi(\mathbf{x}')/k_\mathrm{B}T] \\ p_\mathrm{s}(\mathbf{x}) &\equiv Z^{-1} \exp[-\phi(\mathbf{x})/k_\mathrm{B}T]\,, \end{aligned} \tag{2.7.2}$$

[10] One may think here of a potential energy. If entropic effects from hidden dimensions (i.e., degeneracy of states) play into $\phi(\mathbf{x})$, one should more precisely think of a potential of *free energy*.

and we have $\{\hat{L}(\mathbf{x})p_\mathrm{s}(\mathbf{x})\} = 0$, such that according to the Fokker-Planck equation $\partial_t p(\mathbf{x},t) = \hat{L}(\mathbf{x})p(\mathbf{x},t)$ it is invariant in time, i.e., $\mathrm{e}^{\hat{L}(\mathbf{x})t}p_\mathrm{s}(\mathbf{x}) = p_\mathrm{s}(\mathbf{x})$. Then, for any initial condition $p(\mathbf{x},0)$, one obtains $\lim_{t\to\infty} p(\mathbf{x},t) = p_\mathrm{s}(\mathbf{x})$ [6]. For $p(\mathbf{x},0) = p_\mathrm{s}(\mathbf{x})$ the system is in the steady state at all times $t \geq 0$.

The density $p_\mathrm{s}(\mathbf{x})$ is called *steady-state density* or *Boltzmann-Gibbs density*, since it has the familiar form $\propto \exp[-\phi(\mathbf{x})/k_\mathrm{B}T]$ [6]. In the scenario considered here, it is even the *equilibrium density*, since setting $p(\mathbf{x},t) = p_\mathrm{s}(\mathbf{x})$ in Eq. (2.7.1) gives $\mathbf{j} = 0$, and steady states with $\mathbf{j} = 0$ are in equilibrium (we will see in Sec. 2.9 that the entropy production vanishes for $\mathbf{j} = 0$) [6].

Einstein relation from Boltzmann density

If one would not assume the Einstein relation, Eq. (2.7.1) in one-dimensional notation would read $j(x,t) = -D[(D\gamma)^{-1}\{\partial_x\phi(x)\}p(x,t) + \partial_x]p(x,t)$, which would require $p_\mathrm{s}(x) \propto \exp[-\phi(x)/D\gamma]$ to have $j = 0$. Thus, if it is given that $p_\mathrm{s}(x) \propto \exp[-\phi(x)/k_\mathrm{B}T]$, e.g., from the concept of free energy minimization, see Subsec. 5.2.1, one can deduce the Einstein relation $D\gamma = k_\mathrm{B}T$ from the given p_s.

Rewriting the Fokker-Planck operator

For Langevin systems that approach equilibrium, the Fokker-Planck equation as in Eq. (2.7.1) can, using $p_\mathrm{s}(\mathbf{x})$ as in Eq. (2.7.2), be rewritten as (see Subsec. 3.5.3 or [16, Eq. (4.99)])

$$\hat{L}(\mathbf{x})p_\mathrm{s}(\mathbf{x}) = \nabla_\mathbf{x} \cdot p_\mathrm{s}(\mathbf{x})\mathbf{D}(\mathbf{x})\nabla_\mathbf{x}\,. \tag{2.7.3}$$

Rewritten in this form, it becomes directly apparent that $\hat{L}(\mathbf{x})p_\mathrm{s}(\mathbf{x}) = [\hat{L}(\mathbf{x})p_\mathrm{s}(\mathbf{x})]^\dagger$.

Equivalently to Eq. (2.7.3) we may also write

$$\hat{\mathcal{L}} \equiv p_\mathrm{s}^{-1/2}(\mathbf{x})\hat{L}(\mathbf{x})p_\mathrm{s}^{1/2}(\mathbf{x}) = p_\mathrm{s}^{-1/2}(\mathbf{x})\nabla_\mathbf{x} \cdot p_\mathrm{s}(\mathbf{x})\mathbf{D}(\mathbf{x})\nabla_\mathbf{x}p_\mathrm{s}^{-1/2}(\mathbf{x})\,, \tag{2.7.4}$$

which shows that the Fokker-Planck operator $L(\mathbf{x})$ can be brought into a self-adjoint form $\hat{\mathcal{L}} = \hat{\mathcal{L}}^\dagger$, which is important for the spectral properties of the Fokker-Planck equation. In particular, this implies that $\hat{L}(\mathbf{x})$ is diagonalizable with real eigenvalues and eigenfunctions.

Time-reversal symmetry and detailed balance condition for the conditional density

Using that Eq. (2.7.3) implies $\hat{L}(\mathbf{x})p_\mathrm{s}(\mathbf{x}) = [\hat{L}(\mathbf{x})p_\mathrm{s}(\mathbf{x})]^\dagger$, which in turn implies $p_\mathrm{s}(\mathbf{x})\mathrm{e}^{\hat{L}^\dagger t} = \mathrm{e}^{\hat{L}t}p_\mathrm{s}(\mathbf{x})$, we can use Eqs. (2.6.7) and (2.6.8) to obtain (see Subsec. 3.5.3 for details)

$$p(\mathbf{x},t|\mathbf{y})p_\mathrm{s}(\mathbf{y}) = p(\mathbf{y},t|\mathbf{x})p_\mathrm{s}(\mathbf{x})\,. \tag{2.7.5}$$

This strong symmetry of the transition and steady-state densities is known as the *detailed balance* condition for conditional densities [6, Sec. 5.3.4]. Since this is a defining characteristic of equilibrium dynamics, the terms *detailed balance* and *equilibrium* are often used interchangeably. We moreover call dynamics obeying Eq. (2.7.5) *reversible dynamics* or *reversible systems*, also for times before equilibrium is approached.

For steady-state initial conditions, $p(\mathbf{x},0) = p_\mathrm{s}(\mathbf{x})$, Eq. (2.7.5) implies for two-point (joint) densities that $p(\mathbf{x},t+\Delta t;\mathbf{y},t) = p(\mathbf{x},t;\mathbf{y},t+\Delta t)$. Since steady-state dynamics are time-translation invariant [16], this also implies $p(\mathbf{x},t+\Delta t;\mathbf{y},t) = p(\mathbf{x},t-\Delta t;\mathbf{y},t)$, which in turn implies that correlation functions are symmetric forward and backwards in time. Repeating the argument for n-point densities for $n \to \infty$[11] shows that the process is *symmetric under time reversal*, i.e., that forward and backwards probabilities of paths of length t agree for equilibrium dynamics, i.e., $\mathbb{P}[(\mathbf{x}_\tau)_{0\leq\tau\leq t}] = \mathbb{P}[(\mathbf{x}_{t-\tau})_{0\leq\tau\leq t}]$. This symmetry reflects that in equilibrium there is *no arrow of time*, as dynamics forward and backwards in time cannot be distinguished. If this symmetry is broken, the forward dynamics will be characterized by a strictly positive entropy production, distinguishing it from the backwards dynamics.

2.7.2 Non-equilibrium steady states

We now introduce the case of non-equilibrium steady states (NESS). Deeper theory on NESS will be developed in later chapters, in particular in Secs. 3.5 and 5.5.

Steady-state currents

If the dynamics is confined in a potential or finite space, but the drift is not solely due to the potential (as it was for equilibrium states, see Eq. (2.7.1)) the dynamics will approach an NESS [6, 16] with invariant density $p_\mathrm{s}(\mathbf{x})$ that will not be in equilibrium. That is, there

[11] The limit $n \to \infty$ is mathematically not trivial to take. For a more rigorous approach to time reversal, see, e.g., Ref. [100].

will be a non-zero *steady-state current* (non-zero for some $\mathbf{x}$, not necessarily for all $\mathbf{x}$; for equilibrium steady states $\mathbf{j}_\mathrm{s}(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x}$)

$$\mathbf{j}_\mathrm{s}(\mathbf{x}) \equiv [\mathbf{a}(\mathbf{x}) - \mathbf{D}\nabla]\, p_\mathrm{s}(\mathbf{x}) \neq \mathbf{0}\,. \tag{2.7.6}$$

This reflects that even in the steady state, there is a flow of probability and the detailed balance condition (2.7.5) and time-reversal symmetry are broken, and entropy is produced as we will see in Sec. 2.9.

Since the steady-state density still fulfills $\{\hat{L}(\mathbf{x})p_\mathrm{s}(\mathbf{x}) = 0\}$ for all $\mathbf{x}$ (otherwise it would not be invariant, see Eq. (2.6.5)), the divergence form of the Fokker-Planck equation in Eq. (2.6.5) implies for all $\mathbf{x}$

$$\nabla \cdot \mathbf{j}_\mathrm{s}(\mathbf{x}) = 0\,, \tag{2.7.7}$$

i.e., the steady-state current is *divergence free* (or *incompressible*).

Physical interpretation of the invariant density

In equilibrium, we wrote the Boltzmann-Gibbs equilibrium density $p_\mathrm{s}(\mathbf{x}) \propto \exp[-\phi(\mathbf{x})/k_\mathrm{B}T]$ in Eq. (2.7.2). For non-equilibrium steady states, $p_\mathrm{s}(\mathbf{x})$ might still be of this form, e.g., for dynamics that are driven out of equilibrium by imposing a current $\mathbf{j}_\mathrm{s}(\mathbf{x})$ that leaves $p_\mathrm{s}(\mathbf{x})$ invariant (as for the examples that will be considered in Secs. 3.4 and 3.5). However, in general, a drift that forces the system to depart from equilibrium will also alter $p_\mathrm{s}(\mathbf{x})$, i.e., it will cause the system to settle into a steady state with another density $p_\mathrm{s}^\mathrm{driven}(\mathbf{x}) \neq p_\mathrm{s}^\mathrm{equilibrium}(\mathbf{x})$. Thus, in the general case, the interpretation of $p_\mathrm{s}(\mathbf{x}) \propto \exp[-\phi(\mathbf{x})/k_\mathrm{B}T]$ with a potential $\phi(\mathbf{x})$ is lost.

Even though the physical interpretation is not as direct as in the equilibrium case, we will nevertheless write $p_\mathrm{s}(\mathbf{x}) \propto \exp[-\phi(\mathbf{x})/k_\mathrm{B}T]$ also for NESS, but we should be aware that $\phi(\mathbf{x})$ is no longer the potential energy function. Instead, we simply consider $\phi(\mathbf{x})$ to be an *effective potential*, i.e., we define $\phi(\mathbf{x}) \equiv -k_\mathrm{B}T \ln p_\mathrm{s}(\mathbf{x})$.

Decomposed Langevin equation for NESS

In terms of $p_\mathrm{s}(\mathbf{x})$ and $\mathbf{j}_\mathrm{s}(\mathbf{x})$, the drift term $\mathbf{a}(\mathbf{x})$ in the Langevin equation $\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t$ and in the corresponding Fokker-Planck equation (2.6.4) can be decomposed into *reversible* and *irreversible* components[12] as (for details see Sec. 3.5 and [6, 16])

$$\begin{aligned}\mathbf{a}(\mathbf{x}) &= \mathbf{a}_\mathrm{rev}(\mathbf{x}) + \mathbf{a}_\mathrm{irr}(\mathbf{x})\\ \mathbf{a}_\mathrm{rev}(\mathbf{x}) &= \mathbf{D}\nabla \ln p_\mathrm{s}(\mathbf{x})\\ \mathbf{a}_\mathrm{irr}(\mathbf{x}) &= \mathbf{j}_\mathrm{s}(\mathbf{x})/p_\mathrm{s}(\mathbf{x})\,.\end{aligned} \tag{2.7.8}$$

Note that even though it is easy to verify this decomposition, there is no general way to obtain this decomposition for a given drift field $\mathbf{a}(\mathbf{x})$ where $p_\mathrm{s}(\mathbf{x})$ is not known [101].

One can, however, use this decomposition to construct different NESS corresponding with the same $p_\mathrm{s}(\mathbf{x})$. For a given diffusion matrix, one obtains all such possible NESS by setting $\mathbf{a}_\mathrm{irr}(\mathbf{x}) = \mathbf{A}(\mathbf{x})\nabla \ln p_\mathrm{s}(\mathbf{x}) + \nabla\cdot\mathbf{A}(\mathbf{x})$ for any asymmetric matrix-valued function $\mathbf{A}(\mathbf{x}) = -\mathbf{A}^T(\mathbf{x})$ [101]. We will use this construction and choose a constant $\mathbf{A}(\mathbf{x}) = \mathbf{A}$ in Secs. 3.5 and 5.5.

Further properties of NESS

We will discuss more properties of non-equilibrium steady-states in the following chapters, especially in Secs. 3.5 and 5.5. In particular, in Subsec. 3.5.3 we will generalize the symmetric form of the Fokker-Planck operator in Eq. (2.7.3) to NESS and present a generalization of the time-reversal symmetry and detailed balance condition (2.7.5).

2.7.3 Langevin motion in harmonic potentials

A standard example for steady-state dynamics is Langevin motion in a harmonic potential $\phi(\mathbf{x}) = k_\mathrm{B}T\mathbf{x}^T\mathbf{A}_\mathrm{sym}\mathbf{x}/2D$ with a symmetric, positive definite matrix $\mathbf{A}_\mathrm{sym}$. The equilibrium dynamics in this potential with constant, isotropic diffusion coefficient $\mathbf{D}(\mathbf{x},t) = D\mathbb{1}$ reads

$$\begin{aligned}\mathrm{d}\mathbf{x}_t &= -D\nabla\phi(\mathbf{x}_t)\mathrm{d}t/k_\mathrm{B}T + \sqrt{2D}\mathrm{d}\mathbf{W}_t\\ &= -\mathbf{A}_\mathrm{sym}\mathbf{x}_t\mathrm{d}t + \sqrt{2D}\mathrm{d}\mathbf{W}_t\,.\end{aligned} \tag{2.7.9}$$

[12] The names *reversible* and *irreversible* reflect that the system without the irreversible component would settle into equilibrium, see Eq. (2.7.1)

To address more general motion in harmonic potentials, consider

$$\mathrm{d}\mathbf{x}_t = -\mathbf{A}\mathbf{x}_t \mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t\,, \tag{2.7.10}$$

for any constant matrices $\mathbf{A}$ and $\boldsymbol{\sigma}$ where eigenvalues of $\mathbf{A}$ have positive real parts (reflecting the confining potential) and eigenvalues of $\boldsymbol{\sigma}$ are non-zero (note that the matrices do not have to be diagonalizable, see Sec. 5.5). Compared to Eq. (2.7.9) this is a significant generalization since it allows to study NESS, i.e., systems with non-zero entropy production and broken time-reversal symmetry in the steady state.

The dynamics specified in Eqs. (2.7.9) and (2.7.10) is also known as the *Ornstein-Uhlenbeck process* [102] or *linear diffusion* as the drift term is linear in $\mathbf{x}_t$. Writing Eq. (2.7.10) as $\mathbf{x}_{(n+1)\mathrm{d}t} = (\mathbb{1} - \mathbf{A})\mathbf{x}_{n\mathrm{d}t}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_{n\mathrm{d}t}$, we see from the independence property and Gaussian distribution of $\mathrm{d}\mathbf{W}_{n\mathrm{d}t}$ that, if $\mathbf{x}_{n\mathrm{d}t}$ is Gaussian, then also $\mathbf{x}_{(n+1)\mathrm{d}t}$ is Gaussian as it is the sum of two independent Gaussians, see Eq. (2.2.13). Therefore, whenever $p(\mathbf{x}, 0)$ is Gaussian, $p(\mathbf{x}, t)$ will be a Gaussian density for all t. Due to the time-homogeneity of the process this suffices to conclude that for any Gaussian initial condition $\mathbf{x}_0$ (or any initial condition that is localized in a point), the process $\mathbf{x}_t$ is a Gaussian process, see Subsec. 2.3.2.

For a Gaussian initial condition, solving for the time-dependent density $p(\mathbf{x}, t)$ therefore reduces to computing the first two cumulants which specify the Gaussian density, see Eq. (2.2.6). Following the stochastic-calculus approach developed above, we compute the cumulant as follows (for details see Subsecs. 5.5.8 and 6.3.5). Note for the first cumulant that taking the mean value gives $\mathrm{d}\langle\mathbf{x}_t\rangle = -\mathbf{A}\langle\mathbf{x}_t\rangle\mathrm{d}t$, which implies $\langle\mathbf{x}_t\rangle = \mathrm{e}^{-\mathbf{A}t}\langle\mathbf{x}_0\rangle$. For the second cumulant, the covariance matrix $\boldsymbol{\Sigma}(t) \equiv \langle\mathbf{x}_t\mathbf{x}_t^T\rangle - \langle\mathbf{x}_t\rangle\langle\mathbf{x}_t^T\rangle$, we use Itô's lemma (2.5.14) to obtain $\frac{\mathrm{d}}{\mathrm{d}t}\langle\mathbf{x}_t\mathbf{x}_t^T\rangle = -\mathbf{A}\langle\mathbf{x}_t\mathbf{x}_t^T\rangle - \langle\mathbf{x}_t\mathbf{x}_t^T\rangle\mathbf{A}^T + 2\mathbf{D}$ which in turn gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Sigma}(t) = -\mathbf{A}\boldsymbol{\Sigma}(t) - \boldsymbol{\Sigma}(t)\mathbf{A}^T + 2\mathbf{D}\,. \tag{2.7.11}$$

This is known as the a *differential Lyapunov equation*. In a confining potential, $\boldsymbol{\Sigma}(t)$ will approach a steady-state covariance $\boldsymbol{\Sigma}_\mathrm{s}$ for long times, for which Eq. (2.7.11) implies the *Lyapunov equation*

$$\mathbf{A}\boldsymbol{\Sigma}_\mathrm{s} + \boldsymbol{\Sigma}_\mathrm{s}\mathbf{A}^T = 2\mathbf{D}\,. \tag{2.7.12}$$

If this $\boldsymbol{\Sigma}_\mathrm{s}$ is given, then for an initial covariance $\boldsymbol{\Sigma}(0)$, Eq. (2.7.11) is solved as (easily shown by plugging Eq. (2.7.13) into Eq. (2.7.11), see Subsec. 5.5.8)

$$\boldsymbol{\Sigma}(t) = \boldsymbol{\Sigma}_\mathrm{s} + \mathrm{e}^{-\mathbf{A}t}\left[\boldsymbol{\Sigma}(0) - \boldsymbol{\Sigma}_\mathrm{s}\right]\mathrm{e}^{-\mathbf{A}^T t}\,. \tag{2.7.13}$$

We can write the density $p(\mathbf{x},t)$ and the transition density $p(\mathbf{x},t|\mathbf{x}_0)$ as the respective multivariate Gaussian densities with mean and covariance given by $\langle\mathbf{x}_t\rangle = \mathrm{e}^{-\mathbf{A}t}\langle\mathbf{x}_0\rangle$ and $\boldsymbol{\Sigma}(t)$ as in Eq. (2.7.13). These expressions will allow for several explicit calculations in the following chapters.

2.8 Multiplicative noise

Recall that in the simple case of a spherical Brownian particle in a medium at a constant temperature, one expects the diffusion matrix to be isotropic and constant in space and time, i.e., $\mathbf{D}(\mathbf{x},t) = D\mathbb{1}$, where $\mathbb{1}$ is the d-dimensional unit matrix, and in terms of the friction constant γ, we have the Einstein relation $D = k_\mathrm{B}T/\gamma$, see Eq. (1.3.2). However, since the components of the vector $\mathbf{x}_t$ may generally describe arbitrary degrees of freedom, the isotropy of $\mathbf{D}$ (i.e., $\mathbf{D}(\mathbf{x},t) = D(\mathbf{x},t)\mathbb{1}$) is not a general feature and not part of the assumptions that we will make.

Space- or time-dependent diffusion could, e.g., reflect that the temperature or viscosity of the medium is not constant in space or time. Moreover, space-dependent noise can occur due to non-linear coordinate transformations [17], or due to the adiabatic elimination of hidden degrees of freedom [103]. We will often assume that $\mathbf{D}(\mathbf{x},t) = \mathbf{D}$, or equivalently $\boldsymbol{\sigma}(\mathbf{x},t) = \boldsymbol{\sigma}$, but for the sake of generality we will state the results in Chs. 3 and 4 also for general $\mathbf{D}(\mathbf{x},t)$ (retaining, however, the condition that $\boldsymbol{\sigma}$ is a real matrix with eigenvalues unequal to zero which implies that $\mathbf{D} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T/2$ is always positive definite and symmetric).

For simplicity, the results in the previous sections were stated for space-independent diffusion matrices $\mathbf{D}(t)$, also called *additive noise*. In this section we generalize the previous statements to *multiplicative noise*, i.e., to noise amplitudes $\boldsymbol{\sigma}(\mathbf{x},t)$ and diffusion matrices $\mathbf{D}(\mathbf{x},t)$ that have explicit $\mathbf{x}$-dependence.

There are two complications that appear for multiplicative noise. First, since $\nabla\cdot\mathbf{D}(\mathbf{x},t)$ no longer vanishes, we cannot commute derivatives and $\mathbf{D}$, such that the Fokker-Planck equation (2.6.4) slightly changes. Second, since $\nabla\cdot\boldsymbol{\sigma}(\mathbf{x},t)$ no longer vanishes, the Langevin equation gives rise to different dynamics depending on the interpretation of the noise term $\boldsymbol{\sigma}\mathrm{d}_\lambda\mathbf{W}_t$.

2.8.1 Conditions for Boltzmann equilibrium and detailed balance

We first consider the adapted Fokker-Planck equation, for now starting from the Langevin equation $\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(\mathbf{x}_t, t)\mathrm{d}\mathbf{W}_t$ interpreted with the Itô convention. Repeating the derivation of the Fokker-Planck equation (2.6.4) would give a Fokker-Planck operator

$$\begin{aligned}\hat{L}_{\text{Itô}}(\mathbf{x}, t) &= -\nabla \cdot \mathbf{a}(\mathbf{x}, t) + \sum_{i,j=1}^{d} \partial_i \partial_j \mathbf{D}_{ij}(\mathbf{x}, t) \\ &= -\nabla \cdot [\mathbf{a}(\mathbf{x}, t) - \{\nabla \cdot \mathbf{D}(\mathbf{x}, t)\}] + \nabla \cdot \mathbf{D}(\mathbf{x}, t)\nabla \,. \end{aligned} \tag{2.8.1}$$

For the physical concepts concerning the Boltzmann equilibrium (see Eqs. (2.7.1) and (2.7.2)) and the mean values of displacements in terms of the probability current (see Eq. (2.6.9) and (2.6.10)) to be valid, we require that the Fokker-Planck operator in the multidimensional case is still of the form $-\nabla \cdot \mathbf{a}(\mathbf{x}, t) + \nabla \cdot \mathbf{D}(\mathbf{x}, t)\nabla$ as for additive noise, see Eq. (2.6.5), where $\mathbf{a}(\mathbf{x}, t)$ is the drift resulting from external forces. Even beyond the properties so far, it is known that the Fokker-Planck equation is required to be of this desired form, see, e.g., Ref. [17, Eq. (6.96)].

From Eq. (2.8.1) we see that a general way to achieve this is to postulate that for multitplicative noise, the Langevin equation with the Itô interpretation has to be written as (see, e.g., [60, Eq. (5.2)] and [104])

$$\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t, t)\mathrm{d}t + \nabla \cdot \mathbf{D}(\mathbf{x}, t)\mathrm{d}t + \boldsymbol{\sigma}(\mathbf{x}_t, t)\mathrm{d}\mathbf{W}_t \,, \tag{2.8.2}$$

i.e., there is an extra drift term emerging for multiplicative noise that ensures

$$\hat{L}(\mathbf{x}, t) = -\nabla \cdot \mathbf{a}(\mathbf{x}, t) + \nabla \cdot \mathbf{D}(\mathbf{x}, t)\nabla \,. \tag{2.8.3}$$

The disadvantage of this simple remedy is that the physical interpretation of this drift is not clear since it is only justifiable *a posteriori*, i.e., it is necessary for the correct physical results, but its microscopic origin is not obvious.

2.8.2 Anti-Itô interpretation of the Langevin equation

An approach to explain the extra drift term in Eq. (2.8.2) for certain cases is interpreting the Langevin equation with a convention $\lambda \neq 0$,

$$\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(\mathbf{x}_t, t)\mathrm{d}_\lambda \mathbf{W}_t \,. \tag{2.8.4}$$

Following the same calculation as for Eq. (2.5.25) but for a matrix-valued function we can rewrite the $\mathrm{d}_\lambda \mathbf{W}_t$ expression in Eq. (2.8.4) in terms of an Itô increment ($\lambda = 0$) as [16, Eq. (3.31)]

$$\begin{aligned} \mathrm{d}\mathbf{x}_t &= [\mathbf{a}(\mathbf{x}_t, t) + \lambda \mathbf{h}(\mathbf{x}_t, t)]\mathrm{d}t + \boldsymbol{\sigma}(\mathbf{x}_t, t)\mathrm{d}\mathbf{W}_t \\ \mathbf{h}(\mathbf{x}, t) &= 2\nabla \cdot \mathbf{D}(\mathbf{x}, t) - \boldsymbol{\sigma}(\mathbf{x}, t)\left[\nabla \cdot \boldsymbol{\sigma}^T(\mathbf{x}, t)\right] , \end{aligned} \tag{2.8.5}$$

where the divergence of a matrix is $(\nabla \cdot \mathbf{A})_i \equiv \sum_{k=1}^{d} \frac{\partial A_{ik}}{\partial x_k}$, and as always $\mathbf{D} \equiv \boldsymbol{\sigma}\boldsymbol{\sigma}^T/2$. Hence, the dynamics for different conventions differ by a correction in the drift term. Comparing Eq. (2.8.5) to the desired form in Eq. (2.8.2) we see that this form is obtained whenever $\lambda \mathbf{h}(\mathbf{x}, t) = \nabla \cdot \mathbf{D}(\mathbf{x}, t)$.

For one-dimensional dynamics we have $\partial_x D(x,t) = \partial_x[\sigma^2(x,t)/2] = \sigma(x,t)\partial_x\sigma(x,t)$ such that $h(x,t) = \partial_x D(x,t)$. This means that for multiplicative noise in one-dimensional space the Klimontovich or anti-Itô interpretation $\lambda = 1$ always gives rise to the desired dynamics in Eqs. (2.8.2) and (2.8.3).

For $\lambda = 1$ in multidimensional space we need the further condition that

$$\nabla \cdot \mathbf{D}(\mathbf{x}, t) \stackrel{!}{=} \boldsymbol{\sigma}(\mathbf{x}, t)\left[\nabla \cdot \boldsymbol{\sigma}^T(\mathbf{x}, t)\right] , \tag{2.8.6}$$

which does not hold in general. Whenever this condition holds, the correct multidimensional Langevin dynamics for multiplicative noise are obtained from the Klimontovich or anti-Itô interpretation. For examples where this condition does not hold, one should start from the postulated dynamics in Eq. (2.8.2) instead. From a microscopic point of view, this remains an open question.

2.9 Stochastic entropy

2.9.1 Definitions

We introduced the framework of stochastic thermodynamics in Sec. 1.7. For the scope of the thesis the most important concept from this framework is the entropy production, in particular the total entropy $\Delta S_{\mathrm{tot}}([0,t]) \geq 0$ that is produced in the time interval $[0,t]$, see Sec. 1.7.

For the Langevin dynamics $\mathrm{d}\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t, t)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t$ (for multiplicative noise, see, e.g., Ref. [105]) with Einstein relation $k_{\mathrm{B}}T\boldsymbol{\gamma}^{-1} = \mathbf{D} \equiv \boldsymbol{\sigma}\boldsymbol{\sigma}^T/2$ and external force $\mathbf{F}(\mathbf{x}, t) = \boldsymbol{\gamma}\mathbf{a}(\mathbf{x}, t) = k_{\mathrm{B}}T\mathbf{D}^{-1}\mathbf{a}(\mathbf{x}, t)$ (see Eq. (1.3.2)), a stochastic entropy production in $[0,t]$ can be

assigned to individual trajectories $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$ as follows [28,54]. First, consider the entropy production connected to the probability density $p(\mathbf{x},\tau)$ of the system. The expected (i.e., average) *entropy production in the system* described by a difference of the familiar *Gibbs-Shannon entropies* of initial and final states,

$$\Delta S_{\mathrm{sys}}([0,t]) \equiv -k_{\mathrm{B}}\int \mathrm{d}\mathbf{x} p(\mathbf{x},t)\ln[p(\mathbf{x},t)] + k_{\mathrm{B}}\int \mathrm{d}\mathbf{x} p(\mathbf{x},0)\ln[p(\mathbf{x},0)]\,. \tag{2.9.1}$$

Note that this difference vanishes for the steady-state dynamics, i.e., for dynamics as in Sec. 2.7 with steady-state initial condition $p(\mathbf{x},0) = p_{\mathrm{s}}(\mathbf{x})$. A trajectory-dependent[13] stochastic version Δs_{sys} such that $\langle \Delta s_{\mathrm{sys}}\rangle = \Delta S_{\mathrm{sys}}$ can be defined as

$$\Delta s_{\mathrm{sys}}[(\mathbf{x}_\tau)_{0\leq\tau\leq t}] \equiv -k_{\mathrm{B}}\ln[p(\mathbf{x}_t,t)] + k_{\mathrm{B}}\ln[p(\mathbf{x}_0,0)]\,, \tag{2.9.2}$$

where we insert the stochastic position $\mathbf{x}_t$. Note that to evaluate this expression for a single trajectory we nevertheless need information on the ensemble of trajectories since $p(\mathbf{x},t)$ and $p(\mathbf{x},0)$ cannot be inferred from single trajectories.

In addition to the entropy production in the system, entropy can also be produced in the medium due to a heat transfer from the system into the medium. We assume that the medium remains at a constant temperature and that all heat transferred into the medium therefore constitutes an *entropy production in the medium*. We can write this heat as an integral of the external force $\mathbf{F}$ over the displacement such that [28]

$$\begin{aligned}\Delta s_{\mathrm{med}}[(x_\tau)_{0\leq\tau\leq t}] &\equiv T^{-1}\int_{\tau=0}^{\tau=t}\mathbf{F}(\mathbf{x}_\tau,\tau)\cdot\circ\mathrm{d}\mathbf{x}_\tau \\ &= k_{\mathrm{B}}\int_{\tau=0}^{\tau=t}[\mathbf{D}^{-1}\mathbf{a}(\mathbf{x}_\tau,\tau)]\cdot\circ\mathrm{d}\mathbf{x}_\tau\,.\end{aligned} \tag{2.9.3}$$

Since this integral reflects that force arises as a change in energy along a stochastic path, the correct choice is the Stratonovich integral. In particular, only this convention gives $\int_{\tau=0}^{\tau=t}\{-\nabla\phi(\mathbf{x}_\tau)\}\circ\mathrm{d}\mathbf{x}_\tau = \phi(\mathbf{x}_0)-\phi(\mathbf{x}_t)$ for an external force arising from a potential $\phi(\mathbf{x})$. The average entropy production in the medium can be computed from Eq. (2.6.10) as

$$\Delta S_{\mathrm{med}}([0,t]) \equiv \langle \Delta s_{\mathrm{med}}\rangle = k_{\mathrm{B}}\int \mathrm{d}\mathbf{x}\int_0^t \mathrm{d}\tau[\mathbf{D}^{-1}\mathbf{a}(\mathbf{x},\tau)]\cdot\mathbf{j}(\mathbf{x},\tau)\,. \tag{2.9.4}$$

As expected, this quantity vanishes in equilibrium since there $\mathbf{j}=\mathbf{0}$, see Sec. 2.7.

[13] This particular trajectory-dependent object actually only depends on $\mathbf{x}_0$ and $\mathbf{x}_t$, and not on the whole trajectory $(x_\tau)_{0\leq\tau\leq t}$.

The stochastic and average total entropy production are then defined as the sum of the entropy production in the system and in the medium [28, 54],

$$\begin{aligned}\Delta s_{\mathrm{tot}}[(\mathbf{x}_\tau)_{0\le\tau\le t}] &\equiv \Delta s_{\mathrm{sys}}[(\mathbf{x}_\tau)_{0\le\tau\le t}] + \Delta s_{\mathrm{med}}[(\mathbf{x}_\tau)_{0\le\tau\le t}]\,,\\ \Delta S_{\mathrm{tot}}([0,t]) &\equiv \Delta S_{\mathrm{sys}}([0,t]) + \Delta S_{\mathrm{med}}([0,t])\,.\end{aligned} \tag{2.9.5}$$

Alternatively, the total entropy production Δs_{tot} for overdamped Langevin dynamics may equivalently be defined as the logarithm of the ratio of the path probability of $(\mathbf{x}_\tau)_{0\le\tau\le t}$ and the path probability of the time reversed path $(\mathbf{x}_{t-\tau})_{0\le\tau\le t}$ [54, 69]. In this alternative definition, the time-reversal symmetry (see paragraph below Eq. (2.7.5)) reflects a vanishing entropy production in equilibrium.

2.9.2 A useful formula

We can rewrite the definitions above to obtain a compact formula for ΔS_{tot}. Using $\partial_t \int p(x,t) = \partial_t 1 = 0$ and the Fokker-Planck equation in divergence form $\partial_t p(\mathbf{x},t) = -\nabla\cdot\mathbf{j}(\mathbf{x},t)$, see Eq. (2.6.5), we have

$$\begin{aligned}\partial_\tau \int \mathrm{d}\mathbf{x}\ln[p(\mathbf{x},\tau)]p(\mathbf{x},\tau) &= \int \mathrm{d}\mathbf{x}\ln[p(\mathbf{x},\tau)]\partial_t p(\mathbf{x},\tau)\\ &= -\int \mathrm{d}\mathbf{x}\ln[p(\mathbf{x},\tau)]\nabla\cdot\mathbf{j}(\mathbf{x},\tau)\\ &= \int \mathrm{d}\mathbf{x}\,\mathbf{j}(\mathbf{x},\tau)\cdot\frac{\nabla p(\mathbf{x},\tau)}{p(\mathbf{x},\tau)}\,.\end{aligned} \tag{2.9.6}$$

Using $\mathbf{j}(\mathbf{x},t) = \mathbf{a}(\mathbf{x},t)p(\mathbf{x},t) - \mathbf{D}\nabla p(\mathbf{x},t)$, see Eq. (2.6.5), we can substitute $\nabla p(\mathbf{x},t) = \mathbf{D}^{-1}[\mathbf{a}(\mathbf{x},t)p(\mathbf{x},t) - \mathbf{j}(\mathbf{x},t)]$ above. Integrating the expression over time gives

$$\Delta S_{\mathrm{sys}}([0,t]) = k_{\mathrm{B}}\int_0^t \mathrm{d}\tau\int \mathrm{d}\mathbf{x}\,\mathbf{j}(\mathbf{x},\tau)\cdot\mathbf{D}^{-1}\left[\mathbf{a}(\mathbf{x},\tau) - \frac{\mathbf{j}(\mathbf{x},\tau)}{p(\mathbf{x},\tau)}\right]. \tag{2.9.7}$$

Adding the entropy production in the medium as computed in Eq. (2.9.4) yields the desired formula for the average total entropy production

$$\Delta S_{\mathrm{tot}}([0,t]) = \Delta S_{\mathrm{sys}}([0,t]) + \Delta S_{\mathrm{med}}([0,t]) = k_{\mathrm{B}}\int_0^t \mathrm{d}\tau\int \mathrm{d}\mathbf{x}\,\frac{\mathbf{j}(\mathbf{x},\tau)\cdot\mathbf{D}^{-1}\mathbf{j}(\mathbf{x},\tau)}{p(\mathbf{x},\tau)}\,. \tag{2.9.8}$$

For the case of steady-state dynamics this formula simplifies to

$$\Delta S_{\mathrm{tot}}([0,t]) = tk_{\mathrm{B}}\int \mathrm{d}\mathbf{x}\,\frac{\mathbf{j}_{\mathrm{s}}(\mathbf{x})\cdot\mathbf{D}^{-1}\mathbf{j}_{\mathrm{s}}(\mathbf{x})}{p_{\mathrm{s}}(\mathbf{x})}\,. \tag{2.9.9}$$

Moreover, note that for steady-state dynamics $\Delta S_{\mathrm{sys}} = 0$ such that $T\Delta S_{\mathrm{tot}} = T\Delta S_{\mathrm{med}}$ is the heat dissipated into the medium.

Chapter 3

Statistics of time-integrated densities and currents

After establishing the necessary background in Ch. 2, we now come to the first research topic of this thesis, addressing *time-integrated* and *time-averaged densities and currents* for Langevin dynamics. This chapter contains four publications reproduced in Secs. 3.4-3.7 and, together with the application of the theory following in Ch. 4, makes up the major part of the thesis.

In the following sections, we give a short introduction in Sec. 3.1, and then extend the technical background from Ch. 2 deeper into the specific field of time-integrated observables, arriving at the state of the art in Sec. 3.2. In Sec. 3.3, we summarize how our results contribute to this field. Afterwards, in Secs. 3.4-3.7, four related publications are reproduced. Since each publication contains an introduction and mentions the contribution to the field, the sections preceding the publications are kept rather short and are meant to provide an overview. Only the technical background and state of the art presented here will be somewhat elaborate, since these parts are often assumed to be known to the specialized readership of the papers, and therefore not discussed in the publications.

3.1 Definition of time-integrated densities and currents

As presented in the Introduction in Sec. 1.6, a key type of observables in measurements of fluctuating dynamics, in particular for the inference of steady-state density and current

from individual trajectories are the *time-averaged density and current* (where overlines denote time averages)

$$\begin{aligned}\overline{\rho_t} &\equiv \frac{1}{t}\int_0^t U(\mathbf{x}_\tau)\mathrm{d}\tau \\ \overline{\mathbf{J}_t} &\equiv \frac{1}{t}\int_{\tau=0}^{\tau=t} U(\mathbf{x}_\tau)\circ\mathrm{d}\mathbf{x}_\tau\,.\end{aligned} \tag{3.1.1}$$

First, we note that one could equivalently study the time-*accumulated* or time-*integrated* observables

$$\begin{aligned}\rho_t &\equiv t\overline{\rho_t} = \int_0^t U(\mathbf{x}_\tau)\mathrm{d}\tau \\ \mathbf{J}_t &\equiv t\overline{\mathbf{J}_t} = \int_{\tau=0}^{\tau=t} U(\mathbf{x}_\tau)\circ\mathrm{d}\mathbf{x}_\tau\,.\end{aligned} \tag{3.1.2}$$

To more directly address the notion of averaging over time, we use the definition in Eq. (3.1.1) for now, but both types of observables are completely equivalent whenever t is just some specified constant. Since the main complications lie in the time-integration, and the division by t is merely an optional detail, we use *time-integrated observables* as an umbrella term for Eqs. (3.1.1) and (3.1.2) as in the title of this chapter. The time-integrated observables are also known as *empirical density and current*, *dynamical functionals*, *additive functionals*, and in the case of ρ_t as *local time* or *occupation time*.

The functionals in Eqs. (3.1.1) and (3.1.2) are directly accessible from measured trajectories $(\mathbf{x}_{i\Delta t})_{i=0,1,\ldots,N}$ with a sufficiently small time resolution Δt via the Riemann sums in Eqs. (2.5.3) and (2.5.15) and are therefore very practical to use. The function $U(\mathbf{x}_\tau)$ can thereby be chosen arbitrarily (as long as it is an integrable function), and different choices of $U(\mathbf{x}_\tau)$ can be applied to the same measurement to investigate different properties of the process. To infer spatially local densities and currents of the process, we will often choose $U(\mathbf{x}_\tau) = U^h_{\mathbf{x}}(\mathbf{x}_\tau)$ where $U^h_{\mathbf{x}}$ is a localized function of length scale h (interpreted as coarse-graining scale) around a point $\mathbf{x}$. In Ch. 4 we will also discuss how to choose U for optimal inference of entropy production.

We will explain the observables in Eq. (3.1.1) in more detail in Secs. 3.4 and 3.5, connecting them to time or displacement accumulated in a region, to coarse graining in space and to fluctuating histograms, see in particular Figs. 3.4.1, 3.5.1 and 3.5.2. What is, however, barely shown in the publications reproduced in the later sections are the technical background, and in particular the state of the art of techniques to approach such observables, which we outline in the following.

3.2 Technical background and state of the art

We here present the technical background and state of the art, focusing on the most common methods used to approach the time-averaged or time-integrated functionals. As mentioned in Sec. 1.6, the statistics of these observables is typically addressed using Feynman-Kac theory [33,42] or large deviation theory [43,44,46,106,107], or in terms of thermodynamic uncertainty relations [47–49]. In this section, we first recall and apply the technical background from the previous chapter. Then, to give an overview of the state of the art (with a focus on methodology), we address these three approaches individually. Note that the approach that we pursue in Secs. 3.4, 3.5 and 3.6 will be primarily the use of stochastic calculus instead.

3.2.1 Average values and a first look at fluctuations

Throughout this section and Secs. 3.4 and 3.5 we consider steady-state dynamics (see Sec. 2.7; the generalization beyond steady-state dynamics is given in Sec. 3.6). In this case we use Eqs. (2.5.8) and (2.6.9) to obtain for the steady-state average values of the observables in Eq. (3.1.1)

$$\begin{aligned}\langle\overline{\rho_t}\rangle_{\mathrm{s}} &= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{z} U(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}) \\ &= \int \mathrm{d}\mathbf{z} U(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z})\,, \end{aligned} \tag{3.2.1}$$

and with the steady-state current $\mathbf{j}_{\mathrm{s}}(\mathbf{z})$ from Eq. (2.7.6),

$$\langle\overline{\mathbf{J}_t}\rangle_{\mathrm{s}} = \int \mathrm{d}\mathbf{z} U(\mathbf{z}) \mathbf{j}_{\mathrm{s}}(\mathbf{z})\,. \tag{3.2.2}$$

Therefore, for localized functions $U(\mathbf{z}) \to \delta(\mathbf{z}-\mathbf{x})$, the time-averaged density and current are (unbiased) estimators of $p_{\mathrm{s}}(\mathbf{x})$ and $\mathbf{j}_{\mathrm{s}}(\mathbf{x})$, respectively. Together with the fact that $\overline{\rho_t}$ and $\overline{\mathbf{J}_t}$ take all available information in time into account, this makes them very useful objects for the inference of steady-state properties of the underlying process.

The challenging feature of these observables are their fluctuations. Due to the time integrals, random properties of the whole trajectory $(\mathbf{x}_\tau)_{0\le\tau\le t}$ play into the fluctuations of $\overline{\rho_t}$ and $\overline{\mathbf{J}_t}$. Nevertheless, a simple observation regarding these statistics for long times $t \to \infty$

can be made from the basic principles introduced in Ch. 2. Rewriting the definition in Eq. (3.1.1) for time $t = Nt_1$ with $N \in \mathbb{N}$,

$$\begin{aligned}\overline{\rho_t} &= \frac{1}{N}\sum_{n=1}^{N}\frac{1}{t_1}\int_{(n-1)t_1}^{nt_1} U(\mathbf{x}_\tau)\mathrm{d}\tau \\ \overline{\mathbf{J}_t} &= \frac{1}{N}\sum_{n=1}^{N}\frac{1}{t_1}\int_{\tau=(n-1)t_1}^{\tau=nt_1} U(\mathbf{x}_\tau)\circ\mathrm{d}\mathbf{x}_\tau\,,\end{aligned} \tag{3.2.3}$$

we see that when t_1 becomes large enough for the individual terms in the sum to become effectively uncorrelated, then for $N \to \infty$, i.e., for long times $t \to \infty$[1], we obtain Gaussian fluctuations around the mean values in Eqs. (3.2.1) and (3.2.2) with variance $\mathrm{var}(\overline{\rho_t}), \mathrm{var}(\overline{\mathbf{J}_t}) \propto t^{-1}$, according to the central limit theorem, see Subsec. 2.2.4.

By swapping orders of averaging and stochastic integration, see Subsec. 2.1.4, we can even compute the second moment of $\overline{\rho_t}$ in terms of two-point joint densities $p(\mathbf{z},\tau;\mathbf{z}',\tau')$ as

$$\begin{aligned}\left\langle\overline{\rho_t}^2\right\rangle_\mathrm{s} &= \int_0^t \mathrm{d}\tau\int_0^t \mathrm{d}\tau'\,\langle U(\mathbf{z}_\tau)U(\mathbf{z}_{\tau'})\rangle_\mathrm{s} \\ &= \int_0^t \mathrm{d}\tau\int_0^t \mathrm{d}\tau'\int \mathrm{d}\mathbf{z}\int \mathrm{d}\mathbf{z}' p(\mathbf{z},\tau;\mathbf{z}',\tau')U(\mathbf{z})U(\mathbf{z}')\,.\end{aligned} \tag{3.2.4}$$

If $p(\mathbf{z},\tau;\mathbf{z}',\tau')$ is known, i.e., for steady-state dynamics this means that $p(\mathbf{z},t|\mathbf{z}')$ and $p_\mathrm{s}(\mathbf{z}')$ are known, the variance $\left\langle\overline{\rho_t}^2\right\rangle_\mathrm{s} - \langle\overline{\rho_t}\rangle_\mathrm{s}^2$ can be computed from Eqs. (3.2.1) and (3.2.4). Even though $p(\mathbf{z},\tau;\mathbf{z}',\tau')$ will usually not be known in applications, investigating theoretical results for such fluctuations can still lead to important insight, as we will illustrate in the course of this thesis.

Extending the approach from Eq. (3.2.4) to currents, as in $\langle\overline{\rho_t}\overline{\mathbf{J}_t}\rangle_\mathrm{s}$ or $\langle\overline{\mathbf{J}_t}^2\rangle_\mathrm{s}$, does not give immediate results since, as described in Subsec. 2.5.2, there are no basic calculation rules for averages of double integrals involving $\mathrm{d}W_{t_1}\mathrm{d}t_2$. To overcome this challenge and learn more about the statistics of time-averaged observables, there are different existing approaches available, which we address now.

3.2.2 Feynman-Kac theory

We here outline the Feynman-Kac approach to gain insight into the statistics of the time-integrated density $\rho_t \equiv \int_0^t U(\mathbf{x}_\tau)\mathrm{d}\tau$ following Refs. [33, 41]. On the one hand, this subsection introduces part of the technical background and state of the art. On the other

[1] Note that the quantitative meaning of *long times* depends not only on the considered dynamics, but also on the choice of f (see also Subsec. 3.2.3).

hand, this also serves as an introduction to the publication reproduced in Sec. 3.7 where we generalize this approach to time-integrated currents $\mathbf{J}_t$.

As mentioned before, ρ_t is a stochastic quantity, depending on the random motion $\mathbf{x}_\tau$ in the time interval $\tau \in [0,t]$. To approach the statistics of ρ_t, Mark Kac employed an analogy to quantum mechanical path integrals to arrive at the characteristic function or moment generating function (defined in general in Eq. (2.2.4)) for ρ_t [33,42]. Instead of repeating this approach, we stay closer to the methods developed in Ch. 2 and reproduce an equivalent approach from Ref. [41] that involves similar steps as the derivation of the Fokker-Planck equation in Subsec. 2.6.1

Consider overdamped Langevin dynamics

$$\mathrm{d}\mathbf{x}_t = \mathbf{F}(\mathbf{x}_t)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t\,. \tag{3.2.5}$$

Here and in the following sections, the drift is denoted by $\mathbf{F}$ (previously by $\mathbf{a}$), but recall that it only agrees with the force up to a proportionality constant, see, e.g., Eq. (1.3.2). Carrying out the same steps as in the derivation of the Fokker-Planck equation in Subsec. 2.6.1 (using Itô's lemma (2.5.14) and $\mathrm{d}\rho_t = U(\mathbf{x}_t)\mathrm{d}t$) but for a test function $f(\mathbf{x}_t,\rho_t)$ instead of only $f(\mathbf{x}_t)$ we obtain with $\mathbf{D} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T/2$ that [41] (here, ∇ only acts in the $\mathbf{x}_t$ components)

$$\begin{aligned}\mathrm{d}f(\mathbf{x}_t,\rho_t) =&[\mathbf{F}(\mathbf{x}_t)\cdot\nabla f(\mathbf{x}_t,\rho_t) + \nabla\cdot\mathbf{D}\nabla f(\mathbf{x}_t,\rho_t)]\mathrm{d}t\\ &+\nabla f(\mathbf{x}_t,\rho_t)\cdot\boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t + U(\mathbf{x}_t)\partial_\rho f(\mathbf{x}_t,\rho_t)\mathrm{d}t\,.\end{aligned} \tag{3.2.6}$$

Suppose the trajectory starts at a deterministic point $\mathbf{x}_0$ and denote by $\langle\cdots\rangle_{\mathbf{x}_0}$ the average starting from this point and by $Q_t(\mathbf{x},\rho|\mathbf{x}_0)$ the conditional density to find $\mathbf{x}_t = \mathbf{x}$ and $\rho_t = \rho$. Then, in complete analogy to Subsec. 2.6.1, i.e., by writing $\mathrm{d}\langle f(\mathbf{x}_t,\rho_t)\rangle/\mathrm{d}t$ either in terms of $\partial_t Q_t(\mathbf{x},\rho|\mathbf{x}_0)$, or in terms of $\mathrm{d}f(\mathbf{x}_t,\rho_t)$ using integration by parts, we obtain a Fokker-Planck equation for $Q_t(\mathbf{x},\rho|\mathbf{x}_0)$ as

$$\partial_t Q_t(\mathbf{x},\rho|\mathbf{x}_0) = [\hat{L}(\mathbf{x}) - U(\mathbf{x})\partial_\rho]Q_t(\mathbf{x},\rho|\mathbf{x}_0)\,, \tag{3.2.7}$$

where $\hat{L} = -\nabla_{\mathbf{x}}\cdot\mathbf{F}(\mathbf{x}) + \nabla_{\mathbf{x}}\cdot\mathbf{D}\nabla_{\mathbf{x}}$ is the familiar Fokker-Planck operator for $p(\mathbf{x},t|\mathbf{x}_0)$ as in Eq. (2.6.5). The usual Fokker-Planck dynamics Eq. (2.6.6) is recovered if ρ is integrated out, i.e., for $p(\mathbf{x},t|\mathbf{x}_0) = \int_{-\infty}^{\infty}\mathrm{d}\rho Q_t(\mathbf{x},\rho|\mathbf{x}_0)$. In analogy to the Fokker-Planck equation in Eq. (2.6.6), the operator $[\hat{L}(\mathbf{x}) - U(\mathbf{x})\partial_\rho]$ describes the time-evolution of

the conditional density. It is convenient to perform a Fourier transform[2] $\widetilde{Q}_t(\mathbf{x},\omega|\mathbf{x}_0) \equiv \int_{-\infty}^{\infty} \mathrm{d}\rho\, \mathrm{e}^{-\mathrm{i}\omega\rho} Q_t(\mathbf{x},\rho|\mathbf{x}_0)$ to obtain

$$\partial_t \widetilde{Q}_t(\mathbf{x},\omega|\mathbf{x}_0) = [\hat{L}(\mathbf{x}) - \mathrm{i}\omega U(\mathbf{x})]\widetilde{Q}_t(\mathbf{x},\omega|\mathbf{x}_0)\,. \tag{3.2.8}$$

Using the time-evolution in Eq. (3.2.8) with $\rho_0 = 0$ we may write a formal operator equation in analogy to Eq. (2.6.7),

$$\widetilde{Q}_t(\mathbf{x},\omega|\mathbf{x}_0) = \mathrm{e}^{[\hat{L}(\mathbf{x})-\mathrm{i}\omega U(\mathbf{x})]t}\delta(\mathbf{x}-\mathbf{x}_0)\,. \tag{3.2.9}$$

For now, we assume that $\mathbf{x}_0$ is sampled from the steady-state distribution $p_\mathrm{s}(\mathbf{x})$. The moment generating function of ρ_t can be expressed in terms of $Q_t(\mathbf{x},\rho|\mathbf{x}_0)$ as

$$\left\langle \mathrm{e}^{-\mathrm{i}\omega\rho_t}\right\rangle_\mathrm{s} = \int \mathrm{d}\mathbf{x}_0 p_\mathrm{s}(\mathbf{x}_0) \int \mathrm{d}\mathbf{x}\widetilde{Q}_t(\mathbf{x},\omega|\mathbf{x}_0)\,. \tag{3.2.10}$$

Plugging Eq. (3.2.9) into Eq. (3.2.10) we obtain that

$$\left\langle \mathrm{e}^{-\mathrm{i}\omega\rho_t}\right\rangle_\mathrm{s} = \int \mathrm{d}\mathbf{x} \mathrm{e}^{[\hat{L}(\mathbf{x})-\mathrm{i}\omega U(\mathbf{x})]t} p_\mathrm{s}(\mathbf{x})\,. \tag{3.2.11}$$

This is a version of the Feynman-Kac formula [42]. It relates the statistics of ρ_t, encoded in its moment generating function, to an operator expression.

To access the information contained in Eq. (3.2.11), we address the moments of ρ_t as follows. Expanding the left-hand side of Eq. (3.2.11) around $\omega = 0$ yields the expression $1 - \mathrm{i}\omega\left\langle\rho_t\right\rangle_\mathrm{s} - \omega^2\left\langle\rho_t^2\right\rangle_\mathrm{s}/2 + O(\omega^3)$ in terms of the moments. To expand the right-hand side of Eq. (3.2.11) we use the Dyson operator identity [41, Appendix B]

$$\begin{aligned}\mathrm{e}^{[\hat{L}(\mathbf{x})-\mathrm{i}\omega U(\mathbf{x})]t} =& 1 - \mathrm{i}\omega\int_0^t \mathrm{d}t_1 \mathrm{e}^{\hat{L}(\mathbf{x})(t-t_1)} U(\mathbf{x}) \mathrm{e}^{\hat{L}(\mathbf{x})t_1} \\ &- \omega^2 \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \mathrm{e}^{\hat{L}(\mathbf{x})(t-t_2)} U(\mathbf{x}) \mathrm{e}^{\hat{L}(\mathbf{x})(t_2-t_1)} U(\mathbf{x}) \mathrm{e}^{\hat{L}(\mathbf{x})t_1} + O(\omega^3)\,.\end{aligned} \tag{3.2.12}$$

[2] If ρ_t has non-negative support, i.e., if $U(\mathbf{x}) \geq 0$ for all $\mathbf{x}$, one would perform a Laplace transform instead [41].

Comparing the individual orders of ω on both sides of the expanded Eq. (3.2.11) yields the moments in terms of integrals over $p_{\mathrm{s}}(\mathbf{x})$ and $p(\mathbf{x},t|\mathbf{x}_0)$, i.e., using[3] $\int \mathrm{d}\mathbf{x} \mathrm{e}^{\hat{L}(\mathbf{x})t} U(\mathbf{x}) = \int \mathrm{d}\mathbf{x} U(\mathbf{x})$ and $\mathrm{e}^{\hat{L}(\mathbf{x})t} p_{\mathrm{s}}(\mathbf{x}) = p_{\mathrm{s}}(\mathbf{x})$, the mean reads (in agreement with Eq. (2.5.8))

$$\begin{aligned}\langle \rho_t \rangle_{\mathrm{s}} &= \int \mathrm{d}\mathbf{x} \int_0^t \mathrm{d}t_1 \mathrm{e}^{\hat{L}(\mathbf{x})(t-t_1)} U(\mathbf{x}) \mathrm{e}^{\hat{L}(\mathbf{x})t_1} p_{\mathrm{s}}(\mathbf{x}) \\ &= \int_0^t \mathrm{d}t_1 \int \mathrm{d}\mathbf{x} U(\mathbf{x}) p_{\mathrm{s}}(\mathbf{x}) \\ &= t \int \mathrm{d}\mathbf{x} U(\mathbf{x}) p_{\mathrm{s}}(\mathbf{x}) \,. \end{aligned} \tag{3.2.13}$$

Using $p(\mathbf{x},t|\mathbf{x}_0) = \mathrm{e}^{\hat{L}(\mathbf{x})t} \delta(\mathbf{x}-\mathbf{x}_0)$ (see Eq. (2.6.7)), that is

$$\int \mathrm{d}\mathbf{x}_1 U(\mathbf{x}_0) \mathrm{e}^{\hat{L}(\mathbf{x}_0)t} g(\mathbf{x}_0) = \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x}_0 U(\mathbf{x}) p(\mathbf{x},t|\mathbf{x}_0) g(\mathbf{x}_0) \,, \tag{3.2.14}$$

we obtain (for convenience relabeling $\mathbf{x} \to \mathbf{x}_0$ in Eq. (3.2.13)) from the ω^2-terms the second moment (in agreement with Eq. (3.2.4))

$$\left\langle \rho_t^2 \right\rangle_{\mathrm{s}} = 2 \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x}_0 U(\mathbf{x}) p(\mathbf{x}, t_2 - t_1|\mathbf{x}_0) U(\mathbf{x}_0) p_{\mathrm{s}}(\mathbf{x}_0) \,. \tag{3.2.15}$$

This shows how Feynman-Kac theory makes concrete statements about the fluctuations of time-integrated observables. Note that we already obtained these moments from a direct approach above (see Eqs. (2.5.8) and (3.2.4)). Advantages of the Feynman-Kac approach are that the Feynman-Kac formula (3.2.11) in principle contains much more information on the statistics, and that it provides a clear formalism that can be directly generalized to fluctuations of time-integrated currents and joint fluctuations of densities and currents, as we will show explicitly in Sec. 3.7. Compared to the stochastic-calculus approach that will be pursued in Secs. 3.4-3.6, the Feynman-Kac approach is, however, somewhat indirect and looses track of the direct interpretation in terms of stochastic trajectories.

3.2.3 Large deviation theory

Another large part of existing literature on time-averaged densities and currents is within the framework of *large deviation theory* [108–110], which in this context describes the statistics of ergodic systems on the longest time-scales. Although we will not use any of

[3] The first equation can, e.g., be seen from the divergence form $\hat{L} = \nabla \cdot [\cdots]$ which shows that all terms in the series expansion of $\mathrm{e}^{\hat{L}(\mathbf{x})t}$ except $(\hat{L}(\mathbf{x})t)^0 = 1$ vanish after integration by parts. The second equation follows from $\hat{L}(x) p_{\mathrm{s}}(\mathbf{x}) = 0$.

the results presented in this subsection explicitly, they are very relevant for the purpose of comparing our new results to the state of the art.

For a stochastic process a_t with time-dependent density $p(a,t)$ one defines a large deviation *rate function* as

$$I(a) \equiv -\lim_{t\to\infty} \frac{1}{t} \ln p(a,t)\,. \tag{3.2.16}$$

If this limit exists, then one says that a_t satisfies the large deviation principle [110] which is often written as a more suggestive version of Eq. (3.2.16) as

$$p(a,t) \asymp \exp[-tI(a)]\,. \tag{3.2.17}$$

Here, $\asymp$ is supposed to mean that asymptotically for $t\to\infty$ the statistics of $p(a,t)$ is (up to subexponential terms) completely characterized by $I(a)$. Typically [110], in particular for time-averaged observables [43], $I(a)$ is non-negative and convex with a single zero at the mean value, i.e., $I(\langle a_t\rangle) = 0$. Then, the large deviation principle in Eq. (3.2.17) implies that, as $t\to\infty$, the probability density concentrates around the mean value (which also becomes the most likely value) and fluctuations are exponentially suppressed. The central limit theorem is then typically [111] implied, with a limiting variance $\mathrm{var}(a_t) \to \sigma^2/t = 1/tI''(\langle a_t\rangle)$. The central limit theorem (see Eq. (2.2.14)) also motivates why the factor $1/t$ is included in the definition in Eq. (3.2.16). While the central limit theorem as stated in Eq. (2.2.14) asserts that the small deviations around the mean become Gaussian, the large deviation principle (if it exists) owes its name to the fact that it also applies for large deviations from the mean [43, 110].

For the large deviation theory of time-averaged densities and currents one typically [43–45] (although not always [112]) defines the *empirical* density and current with a delta function, i.e.,

$$\begin{aligned}\overline{\rho^{\mathrm{e}}}(\mathbf{x},t) &= \frac{1}{t}\int_0^t \delta(\mathbf{x}_\tau - \mathbf{x})\mathrm{d}\tau,\\ \overline{\mathbf{J}^{\mathrm{e}}}(\mathbf{x},t) &= \frac{1}{t}\int_0^t \delta(\mathbf{x}_\tau - \mathbf{x})\circ \mathrm{d}\mathbf{x}_\tau.\end{aligned} \tag{3.2.18}$$

This has the advantage that the mean values for steady-state dynamics are the steady-state density $\langle\overline{\rho^{\mathrm{e}}}(\mathbf{x},t)\rangle_{\mathrm{s}} = p_{\mathrm{s}}(\mathbf{x})$ and current $\langle\overline{\mathbf{J}^{\mathrm{e}}}(\mathbf{x},t)\rangle_{\mathrm{s}} = \mathbf{j}_{\mathrm{s}}(\mathbf{x})$, respectively (see Eqs. (3.2.1) and (3.2.2)). However, this definition has practical and mathematical disadvantages, as we point out below.

Instead of a rate function as in Eq. (3.2.16) one typically considers a rate *functional* depending on the fields $\overline{\rho^{\mathrm{e}}}(\mathbf{x},t), \overline{\mathbf{J}^{\mathrm{e}}}(\mathbf{x},t)$ for all $\mathbf{x}$ in the considered domain,

$$I[\rho(\cdot),\mathbf{J}(\cdot)] = -\lim_{t\to\infty}\frac{1}{t}\ln p[\rho(\cdot),\mathbf{J}(\cdot),t]\,. \tag{3.2.19}$$

A central result in this part of large deviation theory is the Donsker-Varadhan [46,113,114] large deviation principle for the empirical density of equilibrium steady-state Langevin dynamics $\mathrm{d}\mathbf{x}_t = \mathbf{F}(\mathbf{x}_t)+\sqrt{2D}\mathrm{d}\mathbf{W}_t$ where the rate functional as presented, e.g., in Ref. [45, Eq. (1.14)] obeys

$$I[\rho(\cdot)] = \frac{D}{4}\int \mathrm{d}\mathbf{x}\rho(\mathbf{x})\left[\nabla\ln\frac{\rho(\mathbf{x})}{p_{\mathrm{s}}(\mathbf{x})}\right]^2. \tag{3.2.20}$$

The large deviation functional for ρ is also known as *level 2* large deviation theory. Level 1 would correspond to fluctuations of, e.g., $\rho(\mathbf{x})$ in a single $\mathbf{x}$, which can be formally obtained from Eq. (3.2.20) via the *contraction principle* [110]. Level 3 large deviations refer to large deviations of the full trajectory $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$, and formally include level 1 and 2 large deviations via the contraction principle.

Another central result including the current $\mathbf{J}$ and valid even for non-equilibrium steady states is the large deviation principle for the joint fluctuations of ρ and $\mathbf{J}$ [34]

$$I[\rho(\cdot),\mathbf{J}(\cdot)] = \int \mathrm{d}\mathbf{x}\frac{[\mathbf{J}(\mathbf{x})-\mathbf{j}_{\mathrm{s}}(\mathbf{x})]^2}{4D\rho(\mathbf{x})}\,. \tag{3.2.21}$$

Since this result is between levels 2 and 3 of it is known as *level 2.5 large deviations* [44].

A great advantage of the results in Eqs. (3.2.20) and (3.2.21) is their simplicity in terms of appearing properties of the process since $p_{\mathrm{s}}(\mathbf{x})$, $\mathbf{j}_{\mathrm{s}}(\mathbf{x})$ and D are the only properties of the process that appear in the rate functionals. In particular, one does not require knowledge of $\mathbf{F}(\mathbf{x})$ or of $p(\mathbf{x},t|\mathbf{x}_0)$.

A serious disadvantage is the difficulty to relate it to measurements, or even to simulations. The limit $t\to\infty$ required for the approach to hold is of course not experimentally realizable and it is in practice hard to decide how large t has to be. Even more problematic is the delta function in the definition in Eq. (3.2.18). The delta function must be understood as a limit of small localization functions, and not only is this limit hard to take, but it is even not clear if or in which sense it converges. Moreover, the delta function also leads to an unfortunate interplay with the limit $t\to\infty$. In analogy to the application of the central limit theorem, see Eq. (3.2.3), an appropriate notion of long times scales should guarantee that the region where $U(\mathbf{x}_\tau)$ is non-zero is sufficiently often visited in the time $[0,t]$. Since the

empirical density and current in Eq. (3.2.18) are defined with a delta function, this notion of long times scales cannot be satisfied in $d \geq 2$-dimensional space since there the Langevin dynamics hits any specific point $\mathbf{x}$ with probability 0, i.e., $\mathbb{P}(\mathbf{x}_\tau = \mathbf{x} \text{ for some } \tau \in [0,t]) = 0$ (in one-dimensional space also $\mathbb{P}(x_t = x) = 0$ but $\mathbb{P}(x_\tau = x \text{ for some } \tau \in [0,t]) > 0$). We will revisit and discuss this and further issues with the definition with the delta function in Secs. 3.4 and 3.5.

3.2.4 A first application of thermodynamic uncertainty relations

Another concept addressing the fluctuations of time-integrated currents are *thermodynamic uncertainty relations* (TURs) [47,48] where, as presented in the Introduction in Eq. (1.8.1) for scalar J_t, the TUR for steady-state dynamics reads

$$\frac{\mathrm{var}(J_t)}{\langle J_t \rangle_{\mathrm{s}}^2} \Delta S_{\mathrm{tot}} \geq 2k_{\mathrm{B}} \,. \tag{3.2.22}$$

Unlike the Feynman-Kac and large deviation approaches presented before, TURs usually do not serve as a tool to learn about fluctuations of time-integrated observables. Instead, the TURs mainly serve as a tool to gain insight into entropy production and dissipation from knowledge about current fluctuations. In particular, the TUR connects dynamical fluctuations to thermodynamical properties of the system in a quantitative way, which will be used in the publications in Secs. 3.4 and 3.5. Note that the whole Ch. 4 will be devoted to TURs which is why, for now, we do not go any deeper into this topic.

3.3 Summary of results

Here, we briefly mention the results of the following four publications reproduced in the following Secs. 3.4-3.7. Since the results are presented in detail in the individual publications, and will again be discussed in the Conclusion in Ch. 7, this summary is kept as short as possible and should mainly serve as an overview to put the following sections into the context presented so far.

Sections 3.4 and 3.5 are a joint publication consisting of a concise letter (Sec. 3.4) and an accompanying, extended manuscript (Sec. 3.5). In this joint publication, we develop the theory to compute variances and correlations of time-integrated densities and currents based on stochastic calculus, and interpret the results in terms of spatial coarse graining. Connected to the above difficulties with the delta function (see Subsec. 3.2.3), we find that

spatial coarse graining is not only necessary for experimental reasons (limited precision and limited sampling) but also for mathematical and thermodynamical reasons. Without coarse graining, i.e., for a delta function as discussed in Subsec. 3.2.3, the variances of $\overline{\rho_t}$ and $\overline{\mathbf{J}_t}$ are found to diverge in $d \geq 2$-dimensional space, while the average values remain finite (see Eqs. (3.2.1) and (3.2.2)). Although there are different notions of convergence of random variables, and a diverging variance does not necessarily prevent convergence, the results and simulations indicate that the density and current in the limit $U(\mathbf{z}) \to \delta(\mathbf{z} - \mathbf{x})$ in $d \geq 2$-dimensional space converge in none of the notions of stochastic convergence, and that a limiting random variable with the correct mean does not exist.

The publication in Sec. 3.6 uses the stochastic-calculus approach to generalize and illustrate the results for the variances and correlations to Langevin systems beyond steady-state dynamics.

Finally, the publication in Sec. 3.7 focuses on methodology and employs Feynman-Kac theory (as presented above in Subsec. 3.2.2) and a related path integral approach to rederive the technical results of Secs. 3.4 and 3.5, and thereby puts these technical approaches on a common footing.

3.4 Mathematical, thermodynamical, and experimental necessity for coarse graining empirical densities and currents in continuous space (Phys. Rev. Lett. 2022)

This section is a slightly adapted version of the publication C. Dieball and A. Godec, "Mathematical, thermodynamical, and experimental necessity for coarse graining empirical densities and currents in continuous space", *Phys. Rev. Lett.* **129**, 140601 (2022).

My personal contribution as first author of the publication was in performing calculations and simulations, analyzing the results, and co-writing the manuscript.

Mathematical, thermodynamical, and experimental necessity for coarse graining empirical densities and currents in continuous space

Cai Dieball and Aljaž Godec

Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

Abstract

We present general results on fluctuations and spatial correlations of the coarse-grained empirical density and current of Markovian diffusion in equilibrium or non-equilibrium steady states on all timescales. We unravel a deep connection between current fluctuations and generalized time-reversal symmetry, providing new insight into time-averaged observables. We highlight the essential role of coarse graining in space from mathematical, thermodynamical, and experimental points of view. Spatial coarse graining is required to uncover salient features of currents that break detailed balance, and a thermodynamically "optimal" coarse graining ensures the most precise inference of dissipation. Defined without coarse graining, the fluctuations of empirical density and current are proven to diverge on all timescales in dimensions higher than one, which has far-reaching consequences for the central-limit regime in continuous space. We apply the results to examples of irreversible diffusion. Our findings provide new intuition about time-averaged observables and allow for a more efficient analysis of single-molecule experiments.

Phys. Rev. Lett. **129**, 140601 (2022)

3.4.1 Introduction

Single-molecule experiments [115–119] probe equilibrium and non-equilibrium (i.e., detailed balance violating) processes during relaxation [49, 73, 120–124] or in steady states [34, 47, 48, 106, 125–129] on the level of individual trajectories. These are typically analyzed by averaging along individual realizations, yielding random quantities with non-trivial statistics [41, 130]. Time-averaged observables, in particular generalized currents, are central to stochastic thermodynamics [48, 67, 131–134]. Such time-average statistical mechanics focuses on functionals of a trajectory $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$, in particular the empirical density (or occupation time [33, 36, 40, 42, 135–139]) $\overline{\rho_{\mathbf{x}}}(t)$ and current $\overline{\mathbf{J}_{\mathbf{x}}}(t)$ at a point $\mathbf{x}$. Necessary in the analysis of laboratory [115, 140] or computer [68] experiments with a finite spatial resolution, and useful for smoothing data *a posteriori* to improve statistics,

the density and current should be defined as spatial averages over a window $U^h_{\mathbf{x}}(\mathbf{x}')$ at $\mathbf{x}$ with coarse-graining scale h

$$\overline{\rho^U_{\mathbf{x}}}(t) \equiv \frac{1}{t}\int_0^t U^h_{\mathbf{x}}(\mathbf{x}_\tau)\mathrm{d}\tau$$
$$\overline{\mathbf{J}^U_{\mathbf{x}}}(t) \equiv \frac{1}{t}\int_{\tau=0}^{\tau=t} U^h_{\mathbf{x}}(\mathbf{x}_\tau)\circ \mathrm{d}\mathbf{x}_\tau, \tag{3.4.1}$$

where $\circ\,\mathrm{d}\mathbf{x}_\tau$ denotes the Stratonovich integral. These observables are illustrated in terms of sojourns of the window in Fig. 3.4.1a,b. Choosing the window $U^h_{\mathbf{x}}$ as a bin, the density and current observables appear as histograms along single trajectories over occupations of or displacements in the bin that fluctuate between different realizations (see Fig. 3.4.1c-e and accompanying extended paper [84] reproduced in Sec. 3.5). Aside from coarse graining, the integration over $U^h_{\mathbf{x}}(\mathbf{x}')$ may also represent a pathwise thermodynamic potential, e.g., heat dissipation (the force integrated along a stochastic path $\int_{\tau=0}^{\tau=t}\mathbf{F}(\mathbf{x}_\tau)\cdot\circ\mathrm{d}\mathbf{x}_\tau$ [131]) or generalized currents [70, 126, 133, 134]. Normalized windows, i.e., $\int U^h_{\mathbf{x}}(\mathbf{z})d\mathbf{z} = 1$, yield $\overline{\rho^U_{\mathbf{x}}}(t)$ and $\overline{\mathbf{J}^U_{\mathbf{x}}}(t)$ that are estimators of the probability density and current density, respectively. The usually defined empirical density $\overline{\rho_{\mathbf{x}}}(t)$ and current $\overline{\mathbf{J}_{\mathbf{x}}}(t)$ [34, 43–45, 107, 110, 141–144] correspond to no coarse graining, i.e., $U^{h=0}_{\mathbf{x}}(\mathbf{z})$ being Dirac's delta function $\delta(\mathbf{x}-\mathbf{z})$.

Reliably inferring from noisy trajectories whether a system obeys detailed balance, notwithstanding recent progress [68, 115, 116, 140, 145–149], remains challenging. *Quantifying* violations of detailed balance is a daunting task. One can quantify broken detailed balance through violations of the fluctuation dissipation theorem [127, 150, 151], which requires perturbing the system from the steady state. One can also check for a symmetry breaking of forward and backward transition-path times [147, 148], measure the entropy production [28, 104, 145, 149], or infer steady-state currents (see arrows in Fig. 3.4.1c) directly [115, 140], all of which require substantial statistics. However, single-molecule experiments often cannot reach ergodic times, have a finite resolution, and only allow for a limited number of repetitions. This leads to uncertainties in estimates of observables such as steady-state currents (see Figs. 3.4.1d-f). Notably, fluctuations of $\overline{\rho^U_{\mathbf{x}}}$ and $\overline{\mathbf{J}^U_{\mathbf{x}}}$ encode information about violations of detailed balance (even where the mean current or its components locally vanish; see Figs. 3.4.1d,e, which *a priori* is hard to interpret.

Current fluctuations have a noise floor—they are bounded from below by the "thermodynamic uncertainty relation" [47, 48, 126] which in turn allows for bounding dissipation in a system from below by current fluctuations [132, 152–154]. As we show in Fig. 3.4.1f (see

Phys. Rev. Lett. **129**, 140601 (2022)

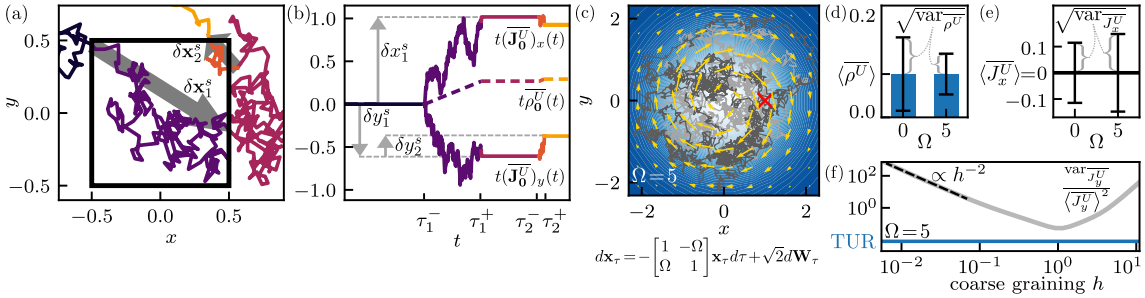

Figure 3.4.1: (a) Diffusive trajectory traversing an observation window $U_{\mathbf{0}}^h(x,y) = 1$ if $|x|, |y| \leq 1/2$ and $U_{\mathbf{0}}^h(x,y) = 0$ otherwise, with time running from dark to bright. Arrows denote contributions $\delta\mathbf{x}_i^s = (\delta x_i^s, \delta y_i^s)$ of the two sojourns in $U_{\mathbf{0}}^h$ between times τ_i^- and τ_i^+ (see Eq. (3.4.7) in Appendix I). (b) Corresponding $t\overline{\rho_{\mathbf{0}}^U}(t)$ and components of $t\overline{\mathbf{J}_{\mathbf{0}}^U}(t)$ from Eq. (3.4.1) as functions of t. (c) Two trajectories $(\mathbf{x}_\tau)$ (gray lines) of length $t = 5$ in confined rotational flow with $\Omega = 5$ (arrows depict the steady-state current $\mathbf{j}_s$). The red cross is the reference point $\mathbf{x}_R = (1,0)$ considered in (d)-(f). Coarse-grained density (d) and x current (e) for a Gaussian window $U_{\mathbf{x}_R}^h$ with $h = 0.3$. Fluctuations of $\overline{\rho_{\mathbf{x}}^U}$ and $\overline{J_{x,y}^U} \equiv (\overline{\mathbf{J}_{\mathbf{x}_R}^U})_{x,y}$ encode violations of detailed balance, even where $\overline{J_x^U}$ vanishes. (f) Squared relative error of $\overline{J_y^U}$ for $U_{\mathbf{x}_R}^h$ as a function of h (gray) bounded by the thermodynamic uncertainty relation (TUR; blue). A variance diverging as h^{-2} (dashed) as $h \to 0$ and vanishing mean for $h \gg 1$ allow for intermediate h optimizing the TUR-bound and thus the inferred dissipation.

Sec. 3.5 for a multiwell potential) the precision of inferring dissipation typically depends non-monotonically on the coarse-graining scale h—given a system, a point $\mathbf{x}$, and trajectory length t there exists a thermodynamically "optimal" coarse graining due to a diverging variance for $h \to 0$ and vanishing mean for large h. Moreover, $\overline{\rho_{\mathbf{x}}}$ and $\overline{\mathbf{J}_{\mathbf{x}}}$ without coarse graining turn out to be ill-defined.

In systems and on timescales where dynamics is reasonably described by a Markov jump process on a small state space, current fluctuations are well understood [44, 48, 60, 106, 155–168]. However, dynamics typically evolves in continuous space, and a continuous dynamics observed on a discrete space is *not* Markovian [169, 170] (see Supplemental Material (SM) in Subsec. 3.4.7 for a quantitative confirmation). An accurate Markov jump description may require too many states to be practical, and is known to fail when considering functionals as in Eq. (3.4.1) [170]. We therefore focus on continuous space, where, with exceptions [34,126,171,172], insight is limited to hydrodynamic scales [157,159,173] and large deviations [43–45, 107, 110, 141–144]. A comprehensive understanding of fluctuations and spatial correlations of density and current in continuous space remains elusive, and the interpretation of the typical definition without coarse graining in dimensions $d \geq 2$ apparently requires a revision, see below.

Phys. Rev. Lett. **129**, 140601 (2022)

Here, we provide general results on the empirical density and current in overdamped diffusive steady-state systems, revealing a mathematical, thermodynamical, and experimental necessity for spatial coarse graining. When defined in a point, fluctuations are proven to diverge in spatial dimensions above one, contradicting existing central-limit statements. We explain why a systematic variation of the coarse-graining scale provides deeper insight about the underlying dynamics and allows for improved inference of the system's thermodynamics. Exploiting a generalized time-reversal symmetry we provide intuition about fluctuating currents along individual trajectories. Non-vanishing density-current correlations are shown to unravel violations of detailed balance from short measurements. Our results allow for a more consistent and efficient analysis of experiments, and provide new insight into non-equilibrium steady states and their thermodynamics.

Phys. Rev. Lett. **129**, 140601 (2022)

3.4.2 Setup

We consider time-homogeneous overdamped Langevin dynamics [6, 16] in d-dimensional space evolving according to the stochastic differential equation $\mathrm{d}\mathbf{x}_\tau = \mathbf{F}(\mathbf{x}_\tau)\mathrm{d}\tau + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_\tau$, where $\mathrm{d}\mathbf{W}_\tau$ is the increment of a d-dimensional Wiener processes (i.e., white noise) with covariance $\langle \mathrm{d}W_{\tau,i}\mathrm{d}W_{\tau',j}\rangle = \delta(\tau-\tau')\delta_{ij}\mathrm{d}\tau\mathrm{d}\tau'$. The Fokker-Planck equation for the conditional probability density with initial condition $G(\mathbf{x},0|\mathbf{y}) = \delta(\mathbf{x}-\mathbf{y})$ reads $(\partial_t + \nabla_\mathbf{x}\cdot\hat{\mathbf{j}}_\mathbf{x})G(\mathbf{x},t|\mathbf{y}) = 0$ with current operator $\hat{\mathbf{j}}_\mathbf{x} \equiv \mathbf{F}(\mathbf{x}) - \mathbf{D}\nabla_\mathbf{x}$, where $\mathbf{D} \equiv \boldsymbol{\sigma}\boldsymbol{\sigma}^T/2$ is the positive definite diffusion matrix. All results directly generalize to multiplicative noise (see Sec. 3.5). The drift $\mathbf{F}(\mathbf{x})$ is assumed to be sufficiently smooth and confining to ensure the existence of a steady-state density $G(\mathbf{x}, t\to\infty|\mathbf{y}) = p_\mathrm{s}(\mathbf{x})$ and, if detailed balance is violated, a steady-state current $\mathbf{j}_\mathrm{s}(\mathbf{x}) \equiv \hat{\mathbf{j}}_\mathbf{x} p_\mathrm{s}(\mathbf{x}) \neq \mathbf{0}$ [6, 16].

3.4.3 Correlations and fluctuations from paths

To investigate the non-trivial statistics of the observables in Eq. (3.4.1) we now outline the derivation detailed in Sec. 3.5 of results for mean values, correlations and fluctuations assuming steady-state initial conditions. Let $\langle\cdot\rangle_\mathrm{s}$ denote the average over all paths $\{\mathbf{x}_\tau\}$ evolving from p_s. The mean values $\langle\overline{\rho^U_\mathbf{x}}(t)\rangle_\mathrm{s} = \int d\mathbf{z} U^h_\mathbf{x}(\mathbf{z})p_\mathrm{s}(\mathbf{z})$ and $\langle\overline{\mathbf{J}^U_\mathbf{x}}(t)\rangle_\mathrm{s} = \int d\mathbf{z} U^h_\mathbf{x}(\mathbf{z})\mathbf{j}_\mathrm{s}(\mathbf{z})$ (see Sec. 3.5) are time-independent estimators of the steady-state density and current coarse-grained over a window $U^h_\mathbf{x}$. In contrast to the mean values, covariances display a non-trivial time dependence and therefore contain salient features of the dynamics. We define the two-point steady-state covariance as

$$C^{\mathbf{xy}}_{AB}(t) \equiv \langle A_\mathbf{x}(t)B_\mathbf{y}(t)\rangle_\mathrm{s} - \langle A_\mathbf{x}(t)\rangle_\mathrm{s}\langle B_\mathbf{y}(t)\rangle_\mathrm{s}\,, \tag{3.4.2}$$

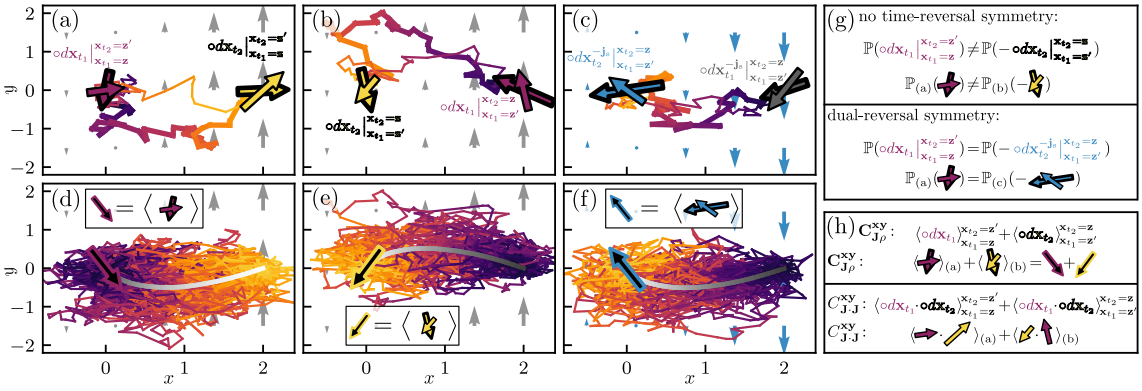

Figure 3.4.2: (a) Two sample trajectories in a shear flow $\mathbf{F}_{\rm sh}(\mathbf{x})$ (gray arrows) with Stratonovich displacements $\circ\mathrm{d}\mathbf{x}_t$ in the initial $\mathbf{x}_{t_1} = \mathbf{z}$ and final point $\mathbf{x}_{t_2} = \mathbf{z}'$ for fixed $t_1 < t_2$ depicted by purple and yellow arrows, respectively. Time is running from dark to bright. (b) Trajectories as in (a) but running from $\mathbf{x}_{t_1} = \mathbf{z}'$ to $\mathbf{x}_{t_2} = \mathbf{z}$. (c) As in (b) but with the inverted shear flow $-\mathbf{F}_{\rm sh}(\mathbf{x}')$ (blue background arrows) and initial and final increments depicted by gray and blue arrows. (d) Ensemble of paths from $\mathbf{x}_{t_1} = \mathbf{z}$ to $\mathbf{x}_{t_2} = \mathbf{z}'$ contributing to $P_{\mathbf{z}}(\mathbf{z}', t_2 - t_1)$. The average initial displacement $\langle \circ\mathrm{d}\mathbf{x}_t \rangle_{\mathbf{x}_{t_1}=\mathbf{z}}^{\mathbf{x}_{t_2}=\mathbf{z}'}$ is depicted by the black-purple arrow, and the mean path $\mathbf{z} \to \mathbf{z}'$ in time $t_2 - t_1$ by the gray gradient line. (e) As in (d) but corresponding to (b) instead of (a). (f) As in (e) but with the reversed shear flow as in (c). (g),(h) Since the shear flow breaks time-reversal symmetry, initial-point increments in (a) cannot be obtained by inverting final-point increments in (b). By dual-reversal symmetry initial-point increments follow from inverting the final-point increments in the inverted shear flow in (c), which explains initial point increments $\circ\mathrm{d}\mathbf{x}_{t_1}$ in current-density correlations and current (co)variances via the easier and more intuitive final point increments $\circ\mathrm{d}\mathbf{x}_{t_2}^{-\mathbf{j}_{\rm s}}$.

where A and B are either $\overline{\rho^U}$ or $\overline{\mathbf{J}^U}$, respectively. We refer to the case $A \neq B$ or $\mathbf{x} \neq \mathbf{y}$ as (linear) correlations and to $A = B$ with $\mathbf{x} = \mathbf{y}$ as fluctuations with the notation $\mathrm{var}_A^{\mathbf{x}}(t) \equiv C_{AA}^{\mathbf{xx}}(t)$. Recall that $\mathrm{var}_\rho^{\mathbf{x}}(t)$ and $\mathrm{var}_{\mathbf{J}}^{\mathbf{x}}(t)$ quantify (experimentally relevant) fluctuations of histograms along single trajectories (see Fig. 3.4.1d,e), and $\mathrm{var}_{\mathbf{J}}^{\mathbf{x}}(t)$ is at the heart of the thermodynamic uncertainty relation (see Fig. 3.4.1f). Moreover, $C_{\mathbf{J}\rho}^{\mathbf{xy}}(t)$ was recently found to play a vital role in stochastic thermodynamics [134]. All $C_{AB}^{\mathbf{xy}}(t)$ are easily inferred from data, but lack physical understanding. We now give $C_{AB}^{\mathbf{xy}}(t)$ a physical meaning in terms of the statistics of paths pinned at end points $\mathbf{z}$ and $\mathbf{z}'$ (see Fig. 3.4.2). Introduce $\langle\cdot\rangle_{\mathbf{x}_{t_1}=\mathbf{z}}^{\mathbf{x}_{t_2}=\mathbf{z}'} \equiv \langle\, \delta(\mathbf{x}_{t_1} - \mathbf{z})\delta(\mathbf{x}_{t_2} - \mathbf{z}')\,\cdot\rangle_{\rm s}$, the Stratonovich increment $\circ\mathrm{d}\mathbf{x}_\tau \equiv \mathbf{x}_{\tau+\mathrm{d}\tau/2} - \mathbf{x}_{\tau-\mathrm{d}\tau/2}$, and the operator

$$\hat{\mathcal{I}}_{\mathbf{xy}}^{t,U}[\cdot] \equiv \frac{1}{t^2}\int_0^t \mathrm{d}t_1 \int_{t_1}^t \mathrm{d}t_2 \int \mathrm{d}\mathbf{z} U_{\mathbf{x}}^h(\mathbf{z}) \int \mathrm{d}\mathbf{z}' U_{\mathbf{y}}^h(\mathbf{z}')[\cdot], \tag{3.4.3}$$

where $[\cdot]$ represents functions of $t_1, t_2, \mathbf{z}, \mathbf{z}'$ and without loss of generality we choose the convention $\int_{t_1}^t \mathrm{d}t_2 \delta(t_2 - t_1) = 1/2$. Upon plugging in mean values $\langle A_{\mathbf{x}} \rangle_{\rm s}$ and $\langle B_{\mathbf{y}} \rangle_{\rm s}$, the

Phys. Rev. Lett. **129**, 140601 (2022)

definition (3.4.2) becomes (see Sec. 3.5) $C^{\mathbf{xy}}_{\rho\rho}(t) = \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}[\Xi^{\mathbf{zz}'}_1 - 2p_{\rm s}(\mathbf{z})p_{\rm s}(\mathbf{z}')]$ for density-density correlations, $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}_\rho}(t) = \hat{\mathcal{I}}^{t,u}_{\mathbf{xy}}[\mathbf{\Xi}^{\mathbf{zz}'}_2 - 2\mathbf{j}_{\rm s}(\mathbf{z})p_{\rm s}(\mathbf{z}')]$ for current-density correlations, and[4] $C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) = \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}[\Xi^{\mathbf{zz}'}_3 - 2\mathbf{j}_{\rm s}(\mathbf{z})\cdot\mathbf{j}_{\rm s}(\mathbf{z}')]$ for current-current correlations, where we defined

$$\begin{aligned}
\Xi^{\mathbf{zz}'}_1 &\equiv \langle 1\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}} + \langle 1\rangle^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'} \\
\mathbf{\Xi}^{\mathbf{zz}'}_2 &\equiv \frac{\langle \circ\mathrm{d}\mathbf{x}_{t_1}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}}{\mathrm{d}t_1} + \frac{\langle \circ\mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'}}{\mathrm{d}t_2} \\
\Xi^{\mathbf{zz}'}_3 &\equiv \frac{\langle \circ\mathrm{d}\mathbf{x}_{t_1}\cdot\circ\mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}}{\mathrm{d}t_1\mathrm{d}t_2} + \frac{\langle \circ\mathrm{d}\mathbf{x}_{t_1}\cdot\circ\mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'}}{\mathrm{d}t_1\mathrm{d}t_2}\,.
\end{aligned} \tag{3.4.4}$$

Eqs. (3.4.3)-(3.4.4) tie $C^{\mathbf{xy}}_{AB}$ to properties of pinned paths, weighted by $U^h_{\mathbf{x}}(\mathbf{z}), U^h_{\mathbf{y}}(\mathbf{z}')$ and integrated over space and times $0 \le t_1 \le t_2 \le t$. In contrast to the somewhat better understood density-density covariance [41, 42, 174], current-density and current-current covariances involve (scalar products of) more subtle Stratonovich increments along pinned trajectories, explained graphically in Fig. 3.4.2 and further investigated in the following.

3.4.4 Correlations and fluctuations from two-point densities

To obtain quantitative results, we evaluate the averages $\langle\cdot\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}$ in terms of two-point functions $P_{\mathbf{z}}(\mathbf{z}', t_2-t_1) \equiv G(\mathbf{z}', t_2-t_1|\mathbf{z})p_{\rm s}(\mathbf{z})$. For density-density correlations $C^{\mathbf{xy}}_{\rho\rho}$, the result is readily obtained from Eq. (3.4.4) using $\langle 1\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}} = P_{\mathbf{z}}(\mathbf{z}', t_2-t_1)$. Conversely, Stratonovich increments, are difficult to understand and hard to evaluate, particularly initial-point increments $\circ\mathrm{d}\mathbf{x}_{t_1}$ because they are correlated with future events.

To gain intuition we examine a two-dimensional shear flow $\mathbf{F}_{\rm sh}(\mathbf{x}) = 2x\hat{\mathbf{y}}$ shown in Fig. 3.4.2, depicting initial-, $\circ\mathrm{d}\mathbf{x}_{t_1}$, and end-point, $\circ\mathrm{d}\mathbf{x}_{t_2}$, increments along forward (Fig. 3.4.2a) and time-reversed (Fig. 3.4.2b) pinned trajectories between times $t_1 < t_2$ and their ensemble averages (Fig. 3.4.2d,e). In the companion extended paper [84], see Sec. 3.5, we show that $\langle\circ\mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}} = \hat{\mathbf{j}}_{\mathbf{z}'}P_{\mathbf{z}}(\mathbf{z}', t_2)\mathrm{d}t_2$, i.e., mean displacements are given by the Fokker-Planck current as expected. Moreover, when detailed balance holds, time-reversal symmetry implies $\mathbb{P}(\circ\mathrm{d}\mathbf{x}_{t_1}|^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}) = \mathbb{P}(-\circ\mathrm{d}\mathbf{x}_{t_2}|^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'})$, whereas under broken detailed balance, e.g., due to the shear flow in Fig. 3.4.2, this ceases to hold. We may, however, employ a generalized time-reversal symmetry—the dual-reversal symmetry (see Sec. 3.5 and [133, 175, 176])—implying $\mathbb{P}(\circ\mathrm{d}\mathbf{x}_{t_1}|^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}) = \mathbb{P}(-\circ\mathrm{d}\mathbf{x}^{-\mathbf{j}_{\rm s}}_{t_2}|^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'})$ connecting ensembles with currents $\mathbf{j}_{\rm s}$ and $-\mathbf{j}_{\rm s}$ (see Fig. 3.4.2c,f,g). Via this generalized time-reversal symme-

[4] Fluctuations and correlations of $\overline{\mathbf{J}^U_{\mathbf{x}}}(t)$ are characterized by the $d\times d$ covariance matrix with elements $[\mathsf{C}^{\mathbf{xy}}_{\mathbf{JJ}}(t)]_{ik} = C^{\mathbf{xy}}_{\mathrm{J}_i\mathrm{J}_k}(t)$. We focus on the scalar case $C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) \equiv \mathrm{Tr}\mathsf{C}^{\mathbf{xy}}_{\mathbf{JJ}}(t)$. See [86] or Sec. 3.7 for the full covariance matrix.

Phys. Rev. Lett. **129**, 140601 (2022)

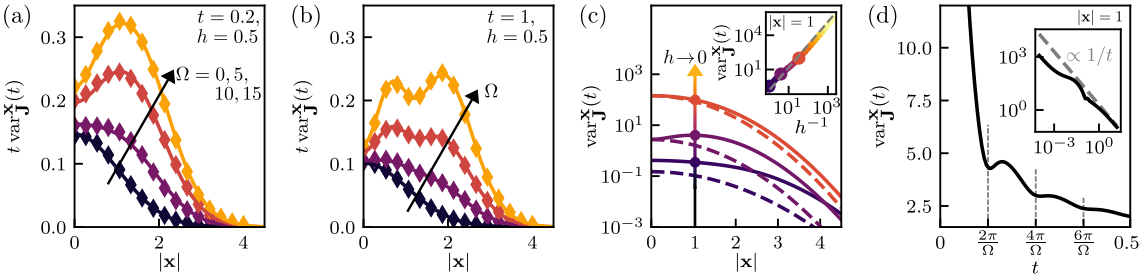

Figure 3.4.3: $t\,\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}$ as a function of the radius $|\mathbf{x}|$ in the harmonically confined rotational flow in Fig. 3.4.1c for increasing Ω with Gaussian $U^h_{\mathbf{x}}$ with width h at (a) $t = 0.2$ and (b) $t = 1$; lines depict Eq. (3.4.5) and symbols simulations, see SM in Subsec. 3.4.7. (c) $t\,\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}$ at $t = 1$ for $\Omega = 10$ (full lines) and equilibrium $\Omega = 0$ (dashed lines), for various h decreasing along the arrow. Inset: divergence of $\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}$ as $h \to 0$ at $|\mathbf{x}| = 1$; the dashed line depicts Eq. (3.4.6). Note the logarithmic scales. (d) $\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}$ as a function of t for very strong driving $\Omega = 50$; inset: (d) on logarithmic scales alongside the central-limit scaling $\propto t^{-1}$.

try we circumvent the correlation of $\circ\mathrm{d}\mathbf{x}_{t_1}$ with the future. To materialize this we isolate the irreversible drift in $\hat{\mathbf{j}}_{\mathbf{x}} = p_{\mathrm{s}}(\mathbf{x})^{-1}\mathbf{j}_{\mathrm{s}}(\mathbf{x}) - p_{\mathrm{s}}(\mathbf{x})\mathbf{D}\nabla_{\mathbf{x}}p_{\mathrm{s}}(\mathbf{x})^{-1}$, and introduce the dual current operator $\hat{\mathbf{j}}^{\ddagger}_{\mathbf{x}} \equiv -\hat{\mathbf{j}}^{-\mathbf{j}_{\mathrm{s}}}_{\mathbf{x}} = p_{\mathrm{s}}(\mathbf{x})^{-1}\mathbf{j}_{\mathrm{s}}(\mathbf{x}) + p_{\mathrm{s}}(\mathbf{x})\mathbf{D}\nabla_{\mathbf{x}}p_{\mathrm{s}}(\mathbf{x})^{-1}$, rendering all terms in Eq. (3.4.4) (illustrated in Fig. 3.4.2h) tractable, and ultimately leading to our main result

$$\begin{aligned}
\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}_\rho}(t) &= \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}[\hat{\mathbf{j}}_{\mathbf{z}}P_{\mathbf{z}'}(\mathbf{z},t') + \hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}P_{\mathbf{z}}(\mathbf{z}',t') - 2\mathbf{j}_{\mathrm{s}}(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}')] \\
C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) &= \frac{2\mathrm{Tr}\mathbf{D}}{t}\int \mathrm{d}\mathbf{z}U^h_{\mathbf{x}}(\mathbf{z})U^h_{\mathbf{y}}(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}) + \\
&\hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}[\hat{\mathbf{j}}_{\mathbf{z}}\cdot\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}'}P_{\mathbf{z}'}(\mathbf{z},t') + \hat{\mathbf{j}}_{\mathbf{z}'}\cdot\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}P_{\mathbf{z}}(\mathbf{z}',t') - 2\mathbf{j}_{\mathrm{s}}(\mathbf{z})\cdot\mathbf{j}_{\mathrm{s}}(\mathbf{z}')],
\end{aligned} \tag{3.4.5}$$

where the first term in $C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ arises from $t_1 = t_2$, see Sec. 3.5, and the operator $\hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}$ simplifies $t^{-2}\int_0^t \mathrm{d}t_1 \int_{t_1}^t \mathrm{d}t_2 \to t^{-1}\int_0^t \mathrm{d}t'(1 - t'/t)$ since Eq. (3.4.5) depends only on time differences $t' \equiv t_2 - t_1 \geq 0$. Notably, written in this simplified form, Eq. (3.4.5) establishes Green-Kubo relations [24, 177] connecting covariances $C^{\mathbf{xy}}_{AB}$ to time-integrals of generalized correlation functions.

Given the two-point function $P_{\mathbf{z}}(\mathbf{z}',t')$, Eq. (3.4.5) gives the correlation and fluctuations of observables defined in Eq. (3.4.1). In practice, $P_{\mathbf{z}}(\mathbf{z}',t')$ may not necessarily be available. However, the theoretical result Eq. (3.4.5) nevertheless allows us to draw several conclusions, in particular by considering special cases and limits. At equilibrium $\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}} = -\hat{\mathbf{j}}_{\mathbf{z}}$, implying $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}_\rho}(t) = 0$. A non-zero $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}_\rho}(t)$ at any time t is thus a conclusive signature of broken detailed balance. Moreover, at equilibrium $C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ does not vanish although $\langle\overline{\mathbf{J}^U_{\mathbf{x}}}\rangle_{\mathrm{s}} = 0$. When $\mathbf{j}_{\mathrm{s}} \neq \mathbf{0}$, $\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) \equiv C^{\mathbf{xx}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ may display maxima where $P_{\mathrm{s}}(\mathbf{x})$ has none (see Fig. 3.4.3a-c), and an oscillatory time dependence due to circulating currents (see Fig. 3.4.3d), both signaling non-equilibrium. For a more detailed discussion of Eq. (3.4.5) see Sec. 3.5.

Phys. Rev. Lett. **129**, 140601 (2022)

3.4.5 Necessity of coarse graining

Of particular interest is the dependence of fluctuations on the coarse-graining length scale h (see Fig. 3.4.1f, Fig. 3.4.3c and Sec. 3.5). Importantly, the limits $h \to \infty$ and $h \to 0$ are generally accessible from Eq. (3.4.5) independent of the detailed dynamics (see Sec. 3.5). The limit $h \to 0$ with $U^h_{\mathbf{x}}(\mathbf{z}) \to \delta(\mathbf{x}-\mathbf{z})$ corresponds to no coarse graining, i.e., the observables Eq. (3.4.1) are evaluated in a single point $\mathbf{z}$. In this limit, the variance and covariance of $\overline{\rho^U_{\mathbf{x}}}$ and $\overline{\mathbf{J}^U_{\mathbf{x}}}$ for $d \geq 2$ and any t behave as (see Sec. 3.5)

$$\begin{aligned}
\mathrm{var}^{\mathbf{x}}_{\rho}(t) &\overset{h\to 0}{\simeq} \frac{k p_{\mathrm{s}}(\mathbf{x})}{t} \times \begin{cases} \frac{h^{2-d}}{d-2} & \text{for } d > 2 \\ -\ln h & \text{for } d = 2 \end{cases} \\
\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t) &\overset{h\to 0}{\simeq} \mathbf{j}_{\mathrm{s}}(\mathbf{x}) \mathrm{var}^{\mathbf{x}}_{\rho}(t)/2p_{\mathrm{s}}(\mathbf{x}) \\
\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) &\overset{h\to 0}{=} \frac{k' p_{\mathrm{s}}(\mathbf{x})}{t}(d-1)h^{-d} + \mathcal{O}(t^{-1})\mathcal{O}(h^{1-d}),
\end{aligned} \tag{3.4.6}$$

where $\simeq$ denotes asymptotic equality, and k, k' are constants depending on $\mathbf{D}$ and $U_{\mathbf{x}}$, see Sec. 3.5. Therefore, taking $U^h_{\mathbf{x}}(\mathbf{z}) \overset{h\to 0}{\longrightarrow} \delta(\mathbf{x}-\mathbf{z})$ as implicitly assumed in [41, 43–45, 107, 110, 141–144] we find for $d \geq 2$ that $\mathrm{var}^{\mathbf{x}}_{\rho,\mathbf{J}}(t), \mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t)$ diverge for all t (see Fig. 3.4.3c). Eq. (3.4.6) also applies to Markov-jump processes defined on a grid with spacing $h \to 0$; for details and an example, see SM in Subsec. 3.4.7. The divergence can be understood intuitively (see also Sec. 3.5), e.g., based on the following argument.

Note that the probability that a point $\mathbf{z}$ is hit by the trajectory $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$, i.e., that there is a $\tau \in [0,t]$ such that $\mathbf{x}_\tau = \mathbf{z}$, delicately depends on the spatial dimensionality d. This probability is positive for $d = 1$ but zero in higher-dimensional space. That is, $\mathbb{P}(\exists\tau \in (0,t] \colon \mathbf{x}_\tau = \mathbf{z}) = 0$ for diffusion in $d \geq 2$ [84, 178]. Mean values remain finite in the limit $h \to 0$, namely $\langle \overline{\rho_{\mathbf{x}}}(t) \rangle_{\mathrm{s}} = p_{\mathrm{s}}(\mathbf{x})$ and $\langle \overline{\mathbf{J}_{\mathbf{x}}}(t) \rangle_{\mathrm{s}} = \mathbf{j}_{\mathrm{s}}(\mathbf{x})$ in agreement with existing literature [34, 43–45, 107, 143, 144, 172]. Since the probability to hit the point $\mathbf{z}$ is approaching zero as $h \to 0$, this implies that the mean is precisely balanced by the infinite contribution of the delta function $U^h_{\mathbf{x}}(\mathbf{z}) \to \delta(\mathbf{x}-\mathbf{z})$, as in $\langle \delta(\mathbf{x}_\tau - \mathbf{x}) \rangle_{\mathrm{s}} = p_{\mathrm{s}}(\mathbf{x})$. Loosely speaking, here "$0 \times \infty$" is finite. One may therefore expect diverging second (and higher) moments when $h \to 0$, as this argument extends to[5] "$0 \times \infty^2 = \infty$". The argument is not limited to overdamped motion but seems to extend to a larger class of stochastic dynamics, such as underdamped diffusion and experimental data on anomalous intracellular transport [91] shown in Fig. 3.5.9 in Sec. 3.5.

[5] For a concrete example following this reasoning, consider random variables X_n where $X_n = n$ with probability $1/n$ and $X_n = 0$ otherwise. Then $\langle X_n \rangle = n/n = 1$ for all n but $\langle X_n^2 \rangle = n^2/n \to \infty$ as $n \to \infty$.

Phys. Rev. Lett. **129**, 140601 (2022)

We hypothesize that not only the moments diverge, but that the density and current cannot even be consistently defined for $h = 0$. Moreover, the limits $h \to 0$ and $t \to \infty$ do *not* commute. This has important consequences for the central-limit regime, i.e., statistics on longest timescales (see Appendix II and Sec. 3.5). Some coarse graining $h > 0$ is therefore necessary for mathematical consistency and anticipated central-limit properties.

Notably, for small windows, Eq. (3.4.6) implies that fluctuations (unlike correlations) carry no information about steady-state currents $\mathbf{j}_\mathrm{s}(\mathbf{x})$ and thus violations of detailed balance and thermodynamic properties such as the system's dissipation. In this limit, fluctuations reflect only Brownian, thermal currents that are invariant with respect to $\mathbf{j}_\mathrm{s}(\mathbf{x})$—systems with equal $p_\mathrm{s}(\mathbf{x})$ and $\mathbf{D}$ display identical fluctuations (see Eq. (3.4.6) and Fig. 3.4.3c). Recall that the dissipation can be inferred from current fluctuations via the thermodynamic uncertainty relation [47, 48, 132]. We now see that only an intermediate coarse graining, such as the "optimum" in Fig. 3.4.1f, allows to infer dissipation from fluctuations. Moreover, spatial features of steady-state currents (see Fig. 3.4.3c) are only revealed with coarse graining. Some coarse graining $h > 0$ is thus necessary to infer thermodynamic properties. In addition, divergent fluctuations make it impossible to accurately infer densities and currents without coarse graining from experiments. Experiments also nominally have a finite spatial resolution. Thus, coarse graining is also experimentally necessary.

Phys. Rev. Lett. **129**, 140601 (2022)

3.4.6 Conclusion

Leveraging Itô calculus and generalized time-reversal symmetry, we were able to provide elusive physical intuition about fluctuations and correlations of empirical densities and currents that are central to stochastic thermodynamics. We established the so far overlooked necessity for spatial coarse graining—it is required to ensure mathematically well-defined observables and the validity of central-limit statements in dimensions $d \geq 2$, to improve the accuracy of inferring thermodynamic properties (e.g., dissipation) from fluctuations and to uncover salient features of non-equilibrium steady-state currents without inferring these individually [179–181], and is unavoidable in the analysis of experimental data with a finite resolution. Non-vanishing current-density correlations were shown to be a conclusive indicator of broken detailed balance, and may improve the accuracy of inferring invariant densities [182] and dissipation far from equilibrium [134]. Our results allow for generalizations to nonstationary initial conditions or non-ergodic dynamics, which will be addressed in forthcoming publications.

Acknowledgments

Financial support from Studienstiftung des Deutschen Volkes (to C. D.) and the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2 (to A. G.) is gratefully acknowledged.

Appendix I: Density and current from sojourns

In general, the density and current functionals measure the ($U^h_{\mathbf{x}}$-weighted) time spent and displacement accumulated in the window $U^h_{\mathbf{x}}$ averaged over time. Specifically, when $U^h_{\mathbf{x}}$ is the indicator function, $U^h_{\mathbf{x}}(\mathbf{z}) = h^{-d}\mathbb{1}_{\Omega_{\mathbf{x}}}(\mathbf{z})$, of a region $\Omega_{\mathbf{x}}$ centered at $\mathbf{x}$ with volume h^d, we can write this illustratively in terms of the sojourns of the window as follows. Letting the times of entering and exiting said window be τ_i^- and τ_i^+, respectively, $t\overline{\rho^U_{\mathbf{x}}}(t)$ corresponds to the sum of sojourn times, $\tau_i^s = \tau_i^+ - \tau_i^-$, and $t\overline{\mathbf{J}^U_{\mathbf{x}}}(t)$ the sum of vectors $\delta\mathbf{x}_i^s$ between entrance $\mathbf{x}_{\tau_i^-}$ and exit $\mathbf{x}_{\tau_i^+}$ points, that is,

$$\begin{aligned} t\overline{\rho^U_{\mathbf{x}}}(t) &= \frac{1}{h^d}\sum_{i\le N_t}(\tau_i^+ - \tau_i^-) \equiv \frac{1}{h^d}\sum_{i\le N_t}\tau_i^s \\ t\,\overline{\mathbf{J}^U_{\mathbf{x}}}(t) &= \frac{1}{h^d}\sum_{i\le N_t}(\mathbf{x}_{\tau_i^+} - \mathbf{x}_{\tau_i^-}) \equiv \frac{1}{h^d}\sum_{i\le N_t}\delta\mathbf{x}_i^s, \end{aligned} \tag{3.4.7}$$

where N_t is the number of visits of the window. Note that N_t is almost surely either ∞ or 0, but the sum converges. The points $\mathbf{x}_0$ or $\mathbf{x}_t$ may lie within $U^h_{\mathbf{x}}$ for which we set $\mathbf{x}_{\tau_1^-} = \mathbf{x}_0$ and/or $\mathbf{x}_{\tau^+_{N_t}} \equiv \mathbf{x}_t$. As a result of correlations between $\mathbf{x}_{\tau_i^-}$ and τ_i^s as well as $\mathbf{x}_{\tau_i^+}$ and $\mathbf{x}_{\tau_{i+1}^-}$, $t\overline{\rho^U}$ and $t\,\overline{\mathbf{J}^U}$ are in general *not* renewal processes. A realization of $\mathbf{x}_\tau$ in Fig. 3.4.1a,b provides intuition about Eq. (3.4.7).

Appendix II: Central-limit regime

Since the observables defined in Eq. (3.4.1) involve time-averages, their statistics on the longest timescales is expected to be governed by the central limit theorem. Indeed, for non-zero h or in spatial dimension $d = 1$ (in both cases we obtained finite variances) on timescales t that are very large compared to all timescales in the system, different parts of a trajectory (e.g., the sojourns in Fig. 3.4.1a and Eq. (3.4.7)) become sufficiently uncorrelated such that the central limit theorem implies Gaussian statistics. However, the diverging variance for $h \to 0$ for $d \ge 2$ prevents Gaussian central-limit statistics on all timescales for the empirical density and current defined with a delta-function (i.e., without coarse graining). Since the diverging part of the variance in Eq. (3.4.6) has the dominant central-limit scaling $\propto t^{-1}$, the asymptotic variance $\sigma_A^2 \overset{t\to\infty}{=} t\mathrm{var}^{\mathbf{x}}_A(t)$ (where $A_{\mathbf{x}}(t)$ denotes $\overline{\rho^U_{\mathbf{x}}}(t)$

Phys. Rev. Lett. **129**, 140601 (2022)

or $\overline{\mathbf{J}_{\mathbf{x}}^{U}(t)}$) also diverges as $h \to 0$. This implies that taking $t \to \infty$ first and then $h \to 0$ also does *not* yield finite variances. Moreover, note that the longest timescale in the system becomes the recurrence time, which diverges as $h \to 0$. We hypothesize that a limiting distribution of $A_{\mathbf{x}}(t)$ only exists as a scaling limit where $h \to 0$ and $t \to \infty$ simultaneously in some d-dependent manner, see Sec. 3.5.

The central-limit regime is generally contained in the framework of large deviation theory [43, 110, 183]. Due to the divergent variance σ_A^2 and the resulting breakdown of Gaussian central-limit statistics, any large deviation principle for empirical densities and currents without coarse graining that predicts finite variances ceases to hold in $d \geq 2$.

Phys. Rev. Lett. **129**, 140601 (2022)

3.4.7 Supplemental Material

In this Supplementary Material (SM) we present an explicit numerical confirmation that a discretely observed diffusion is *not* a Markov process.Moreover, we apply the results for the limit of small window sizes to current fluctuations in a Markov jump process in discretized space. Finally, we list the information that is necessary to reproduce all simulations and analytical results shown in the figures presented above. Further details and derivations related to statements above can be found in accompanying extended manuscript [84], see Sec. 3.5.

Quantification of non-Markovianity of diffusion observed in discrete space (i.e., on a grid)

In this section, we provide a quantitative example for the statement made above that "a continuous dynamics observed on a discrete space is *not* Markovian". This in particular demonstrates that a Markov-jump description with a limited number of states in general *cannot* accurately describe a diffusion process in continuous space. A Markov-jump description may be accurate in systems with a timescale separation (e.g., as a result of high energy barriers separating minima) on timescales sufficiently larger than the slowest relaxation (e.g., much larger than the longest relaxation time in the minima). However, in general an accurate Markov-jump representation requires too many states to describe diffusive dynamics, and in particular to accurately describe functionals of paths [170]. The example we provide here is the Ornstein-Uhlenbeck process with a rotational flow, see Fig. 3.4.1c. To quantify the non-Markovianity we discretize the dynamics on a finite grid and quantify violations of the Chapman-Kolmogorov equation, as described in Ref. [96].

We consider the Ornstein-Uhlenbeck process in two-dimensional continuous space, i.e., Eq. (3.4.17) with $r = D = 1$ and $\Omega = 3$. Then we divide the area $[-5, 5] \times [-5, 5]$ into

$N \times N$ squares S_i that we label by an index $i \in \{1, \dots N^2\}$. We define the steady-state occupation of the state i (i.e., of the area/square S_i) from the continuous-space steady-state density $p_{\rm s}(\mathbf{x}_0)$ (see Eq. (3.4.19)) as

$$p_{\rm s}(i_0) \equiv \int_{x_0 \in S_{i_0}} \mathrm{d}^2 x_0 p_{\rm s}(\mathbf{x}_0), \tag{3.4.8}$$

and, for any time t, from the continuous-space propagator (i.e., conditional density) $G(\mathbf{x}, t|\mathbf{x}_0)$ (see Eq. (3.4.20)) we define the propagator of the discrete space observation as

$$G(i, t|i_0) \equiv \frac{1}{p_{\rm s}(i_0)} \int_{x \in S_i} \mathrm{d}^2 x \int_{x_0 \in S_{i_0}} \mathrm{d}^2 x_0 G(\mathbf{x}, t|\mathbf{x}_0) p_{\rm s}(\mathbf{x}_0). \tag{3.4.9}$$

Moreover, we define [96]

$$G^{\rm CK}_{t'}(i, t|i_0) \equiv \sum_{j=1}^{N^2} G(i, t - t'|j) G(j, t'|i_0). \tag{3.4.10}$$

Note that by the Chapman-Kolmogorov equation a Markov process would obey $G^{\rm CK}_{t'}(i, t|i_0) = G(i, t|i_0)$ for all i, i_0, t, t'. To quantify non-Markovianity we compute the Kullback-Leibler divergence between G and $G^{\rm CK}$,

$$\mathcal{D}^{\rm CK}_{\rm KL}(t', t, i_0) \equiv \sum_{i=1}^{N^2} G(i, t|i_0) \ln \left[\frac{G(i, t|i_0)}{G^{\rm CK}_{t'}(i, t|i_0)} \right]. \tag{3.4.11}$$

The results for this example are shown in Fig. 3.4.4a. Whenever $\mathcal{D}^{\rm CK}_{\rm KL} > 0$ the process is non-Markovian. However, the exact value of $\mathcal{D}^{\rm CK}_{\rm KL} \neq 0$ does not have a direct interpretation. To gain some intuition about the actual value, in Fig. 3.4.4b we normalize by the Kullback-Leibler divergence of G and $p_{\rm s}$,

$$\mathcal{D}^{\rm ps}_{\rm KL}(t, i_0) \equiv \sum_{i=1}^{N^2} G(i, t|i_0) \ln \left[\frac{G(i, t|i_0)}{p_{\rm s}(i)} \right]. \tag{3.4.12}$$

The rationale is that the actual value (at least) should not directly depend on how far the actual dynamics is displaced from the steady state, i.e., if the dynamics is closer to the steady state the same value of $\mathcal{D}^{\rm CK}_{\rm KL}$ should be interpreted as stronger violation of Markovianity as compared to when it is farther away, since the actual dynamics changes less in magnitude in the former case.

Phys. Rev. Lett. **129**, 140601 (2022)

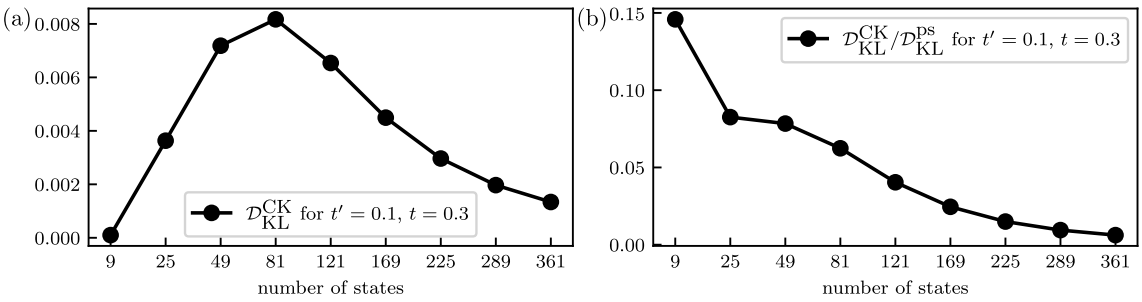

Figure 3.4.4: Quantification of non-Markovianity (as described in the text) in a discrete-space observation with $N \times N$ states for $N = 3, 5, 7, \ldots, 19$. The integrals in Eq. (3.4.9) were evaluated numerically. (a) The value of $\mathcal{D}_{\mathrm{KL}}^{\mathrm{CK}}(t', t, i_0)$, for i_0 denoting the index of the square containing the point $(0, 0)$, obviously decreases for large numbers of states according to expectations. For very small numbers of states, where very few details of the dynamics are observed, the non-Markovianity also appears to be smaller. (b) To gain intuition about the values of $\mathcal{D}_{\mathrm{KL}}^{\mathrm{CK}}(t', t, i_0)$ we normalized it by $\mathcal{D}_{\mathrm{KL}}^{\mathrm{ps}}(t, i_0)$, i.e., a value of 1 means that the Kullback-Leibler divergence of G and G^{CK} equals the divergence of G and G^{ps}. Since the Kullback-Leibler divergence is *not* a metric (it is not symmetric and does not satisfy the triangle inequality), one should be careful when interpreting its values *quantitatively*.

Phys. Rev. Lett. **129**, 140601 (2022)

As expected, the extent of the violation of Markovianity depends on the grid-size. It reduces for sufficiently small grids (large N), i.e., the dynamics become effectively Markovian on shorter time-scales, as well as for large "ignorant" discrete observations (small N), where all probability flows are averaged over. Both limits are intuitive—non-Markovianity arises because there is no timescale separation ensuring local-equilibrium. That is, the direction and rate of leaving a discretely observed state depends on the previous state and the precise location of entering the binned discrete state. Notably (and obviously), the dynamics in the limit of infinite number of grid points, i.e., in the continuum limit, is exactly the diffusion process and thus Markovian. Moreover, the attenuation of non-Markovianity for very "ignorant" grids (i.e., small number of states) as a result of spatial averaging over large regions is easiest understood by realizing that the dynamics with one state is (trivially) Markovian and stationary at all times, the dynamics with two states slightly less so, etc.

A diffusion observed on a grid is thus *not* a Markov jump process and one in general requires many states for an accurate discrete-state Markov-jump representation, which is typically not experimentally feasible. Moreover, one actually needs to parameterize the Markov state model with such a large number of states, which is even less feasible. In contrast, evaluating empirical densities and currents in finite windows assuming an underlying continuous-space diffusion—as carried out in the present work—is *not* constrained to small windows *nor* does it require any parameterization of a discrete-state model for its interpretation.

Implications for fluctuations in Markov jump processes on a grid

If one approximates $d \geq 2$-dimensional continuous-space dynamics by a Markov jump process on a grid, the fact that the Markov-jump description becomes asymptotically accurate for the number of states $N \to \infty$ implies that for large N, correlations and fluctuations of densities and currents will be governed by the limits for $h \to 0$ in Eq. (3.4.6). Corresponding to the problems with large deviation theory for $h \to 0$ discussed in the Appendix, central-limit large deviations in dimensions $d \geq 2$ exist only for finite grids (corresponding to $h > 0$). In contrast, in the continuum limit where the number of states tends to infinity (see, e.g., [132]; corresponding to $h \to 0$) the recurrence time to visit a state diverges, which would require $t > \infty$ for validity of the large deviation principle. In particular, a finite relaxation timescale (where the relaxation time is the inverse of the smallest non-zero eigenvalue of the generator of the dynamics) that remains finite in the continuum limit alone does not guarantee the validity of the large deviation principle.

We now give a specific example for the validity of the limits for $h \to 0$ in Eq. (3.4.6). A state at position $\mathbf{x}$ of a Markov jump process on a grid with spacing h can asymptotically be interpreted as corresponding to a window function $U^h_{\mathbf{x}}$ that is the normalized indicator function of a square with spacing h around $\mathbf{x}$, e.g., in two-dimensional space $U^h_{\mathbf{x}}(x', y') = h^{-2}\mathbb{1}_{|x-x'|,|y-y'|\leq h/2}$. With this correspondence, applied to a two-dimensional example with constant isotropic diffusion, the current fluctuations for a small grid-spacing h should according to Eq. (3.4.6) (see accompanying paper for prefactor $2D$) be governed by

$$\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) \overset{h\to 0}{=} \frac{2D}{t} p_{\mathrm{s}}(\mathbf{x})(d-1)h^{-2} + \mathcal{O}(t^{-1})\mathcal{O}(h^{-1}). \tag{3.4.13}$$

Rates $k^{j\to j'}_{i\to i'}$ (denoting x-indices by i and y-indices by j) for a discretized process originating from continuous dynamics are not unique, but can e.g., be obtained following [184]. Via the pseudo potentials (we need to use *pseudo* potentials since we consider a non-equilibrium process [184])

$$\begin{aligned} \tilde{U}(x,y) &= \frac{r}{2D}x^2 - \frac{\Omega}{D}xy \\ \tilde{V}(x,y) &= \frac{r}{2D}y^2 + \frac{\Omega}{D}xy, \end{aligned} \tag{3.4.14}$$

we obtain rates

$$k^{j\to j}_{i\to i\pm 1} = \frac{D}{h^2}\exp\left(-\frac{1}{2}[\tilde{U}(x_{i\pm 1}, y_j) - \tilde{U}(x_i, y_j)]\right)$$

Phys. Rev. Lett. **129**, 140601 (2022)

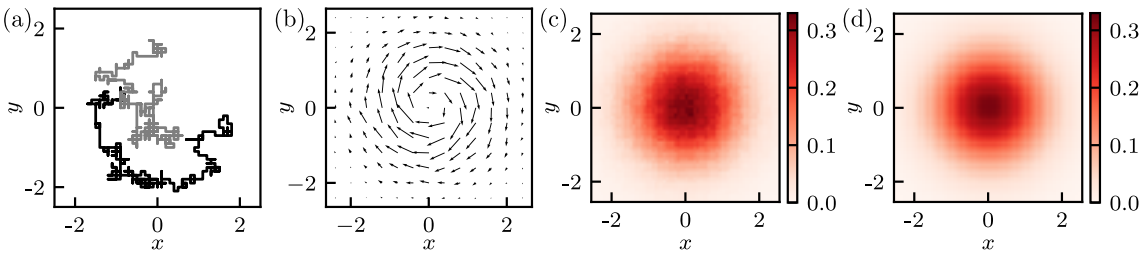

Figure 3.4.5: (a) Two sample trajectories of the Ornstein-Uhlenbeck process discretized as described in the text with time step $\mathrm{d}t = 0.001$ and total time $T = 1$, starting in steady-state initial conditions. (b) Mean current as defined in the text obtained from a simulation of 10,000 trajectories such as the ones in (a). For visibility, arrows are only drawn at every fourth x and y value contained in the grid. (c) Variance of the current at individual grid points obtained from the same simulation as in (b). (d) Eq. (3.4.16) evaluated at individual grid points. The qualitative and quantitative agreement with (c) shows that Eq. (3.4.16), which was derived in continuous space, has direct consequences for Markov jump processes on grids with small h.

$$k_{i\to i}^{j\to j\pm1} = \frac{D}{h^2}\exp\left(-\frac{1}{2}[\tilde{V}(x_i, y_{j\pm1}) - \tilde{V}(x_i, y_j)]\right). \tag{3.4.15}$$

Note that all other rates (i.e., to non-neighboring states) vanish in this construction.

Currents are now defined as transition counts on the edges of the grid. To compare to continuous-space time-averaged currents, we define the x-component of a current $\overline{\mathbf{J}_\mathbf{x}}$ at a grid point $\mathbf{x}$ as the net number of transitions on the edge to the right of $\mathbf{x}$ and the y-component as the net number of transitions on the edge above $\mathbf{x}$. Note that this current reflects probabilities of transitions and not probability densities which differs by the current density by a factor of h^2, i.e., by the area corresponding to of a state. With this definition, the limit Eq. (3.4.13) for $h \to 0$ becomes

$$\mathrm{var}_\mathbf{J}^\mathbf{x}(t) \overset{h\to0}{=} \frac{2D}{t} p_\mathrm{s}(\mathbf{x})(d-1) + \mathcal{O}(t^{-1})\mathcal{O}(h^1). \tag{3.4.16}$$

Note that one could instead equivalently define the current with the normalization h^{-2} to obtain densities and use the limit in Eq. (3.4.13).

Fig. 3.4.5 illustrates the validity of the limit Eq. (3.4.16) for a discretized Ornstein-Uhlenbeck process, see Eq. (3.4.17) with $r = D = 1$, $\Omega = 3$. The process on the area $[-5, 5]^2$ is discretized into a $101 \times 101 = 10201$-state Markov jump process, i.e., the grid spacing is $h = 0.1$. The quantitative agreement of Fig. 3.4.5c and d illustrates that the limit current fluctuations on the jump process are indeed governed by Eq. (3.4.16). Note that the t^{-1} scaling of the h^{-2} term in Eq. (3.4.16) ensures that for $h \to 0$ current fluctuations are governed by this equation for all time-scales.

Phys. Rev. Lett. **129**, 140601 (2022)

Numerical and analytical evaluation used for the figures

This section gives further parameters and all details necessary to reproduce Figs. 3.4.1-3.4.3.

Analytical results for the two-dimensional Ornstein-Uhlenbeck process

For the numerical and analytical results shown in Figs. 3.4.1 and 3.4.3, we use the two-dimensional Ornstein-Uhlenbeck process given by the Langevin equation

$$\mathrm{d}\mathbf{x}_t = \mathbf{F}_{\mathrm{rot}}(\mathbf{x}_t)\mathrm{d}t + \sqrt{2D}\mathrm{d}\mathbf{W}_t, \tag{3.4.17}$$

with drift field $\mathbf{F}_{\mathrm{rot}}(\mathbf{x}) = -\Theta\mathbf{x}$ where $\Theta = \begin{bmatrix} r & -\Omega \\ \Omega & r \end{bmatrix}$, $r > 0$. The drift part splits into

$$\mathbf{F}_{\mathrm{rot}}(\mathbf{x}) = -D\{\nabla\phi(\mathbf{x})\} + \mathbf{j}_{\mathrm{s}}(\mathbf{x})/p_{\mathrm{s}}(\mathbf{x}), \tag{3.4.18}$$

with potential $\phi(\mathbf{x}) = \frac{r}{2D}\mathbf{x}^T\mathbf{x}$ and steady-state density and current

$$\begin{aligned} p_{\mathrm{s}}(\mathbf{x}) &= \frac{r}{2\pi D}\mathrm{e}^{-r\frac{\mathbf{x}^T\mathbf{x}}{2D}} \\ \mathbf{j}_{\mathrm{s}}(\mathbf{x}) &= (\mathbf{F} - D\nabla_{\mathbf{x}})p_{\mathrm{s}}(\mathbf{x}) = \Omega p_{\mathrm{s}}(\mathbf{x}) \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix}. \end{aligned} \tag{3.4.19}$$

A straightforward left-right decomposition [6] gives the propagator/two-point function

$$\begin{aligned} G(\mathbf{x}, t'|\mathbf{x}_0) &= \frac{r}{2\pi D(1-\mathrm{e}^{-2rt'})} \exp\left[\frac{-r\left(\mathbf{x} - \mathrm{e}^{-rt'}\begin{bmatrix} \cos(\Omega t') & \sin(\Omega t') \\ -\sin(\Omega t') & \cos(\Omega t') \end{bmatrix}\mathbf{x}_0\right)^2}{2D(1-\mathrm{e}^{-2rt'})}\right] \\ P_{\mathbf{x}_0}(\mathbf{x}, t') &\equiv G(\mathbf{x}, t'|\mathbf{x}_0)p_{\mathrm{s}}(\mathbf{x}_0). \end{aligned} \tag{3.4.20}$$

We then analytically solve the necessary Gaussian integrals for Gaussian window functions

$$U^h_{\mathbf{x}}(\mathbf{z}) = (2\pi h^2)^{-d/2}\exp\left[-\frac{(\mathbf{z}-\mathbf{x})^2}{2h^2}\right], \tag{3.4.21}$$

and numerically solve the remaining t'-integral. This enables a very fast and stable (even for very small coarse graining where numerical spatial integrals would eventually fail) computation of the second moments shown in Figures 2 and 3. The analytical integrals were performed with the Python-based computer algebra system SymPy [185]. To give an ex-

Phys. Rev. Lett. **129**, 140601 (2022)

ample, we now show the computation of one of the terms in the current variance result in Eq. (3.4.5) (other terms similarly).

We start e.g., with the spatial integrals

$$\int \mathrm{d}^2x \int \mathrm{d}^2x_0 V(\mathbf{x})U(\mathbf{x}_0)j_\mathrm{s}^2(\mathbf{x})j_\mathrm{s}^2(\mathbf{x}_0)G(\mathbf{x},t|\mathbf{x}_0)p_\mathrm{s}(\mathbf{x}_0), \tag{3.4.22}$$

where we set $\mathbf{x}=(x_1,x_2), \mathbf{x}_0=(x_3,x_4)$ such that $j_\mathrm{s}^2(\mathbf{x})=-\Omega x_1$ and $j_\mathrm{s}^2(\mathbf{x}_0)=-\Omega x_3$ and we use constants $\{c_i\}$ to write

$$\begin{aligned}
V(\mathbf{x}) &= \frac{c_1}{\pi}\mathrm{e}^{-c_1\left((x_1-y_1)^2+(x_2-y_2)^2\right)}\\
U(\mathbf{x}_0) &= \frac{c_2}{\pi}\mathrm{e}^{-c_2\left((x_3-y_3)^2+(x_4-y_4)^2\right)}\\
p_\mathrm{s}(\mathbf{x}_0) &= \frac{c_3}{\pi}\mathrm{e}^{-c_3\left(x_3^2+x_4^2\right)}\\
G(\mathbf{x},t|\mathbf{x}_0) &= \frac{c_4}{\pi}\mathrm{e}^{-c_4\left((c_5x_3+c_6x_4-x_1)^2+(c_5x_4-c_6x_3-x_2)^2\right)}.
\end{aligned} \tag{3.4.23}$$

Integrating from $-\infty$ to ∞ over x_3 and x_4 gives (Gaussian integrals with $c_1,c_2,c_3,c_4>0$)

$$\begin{aligned}
&\int \mathrm{d}^2x_0 V(\mathbf{x})U(\mathbf{x}_0)j_\mathrm{s}^2(\mathbf{x})j_\mathrm{s}^2(\mathbf{x}_0)G(\mathbf{x},t|\mathbf{x}_0)p_\mathrm{s}(\mathbf{x}_0)\\
&\quad=\frac{\Omega^2c_1c_2c_3c_4x_1\left(c_2y_3+c_4c_5x_1-c_4c_6x_2\right)}{\pi^3\left(c_2+c_3+c_4c_5^2+c_4c_6^2\right)^2}\\
&\qquad\times\mathrm{e}^{\frac{c_2^2y_4^2+2c_2c_4y_4(c_5x_2y_4+c_6x_1)+c_4^2(c_5x_2+c_6x_1)^2+(c_2y_3+c_4c_5x_1-c_4c_6x_2)^2}{c_2+c_3+c_4c_5^2+c_4c_6^2}}\\
&\qquad\times\mathrm{e}^{-c_1x_1^2+2c_1x_1y_1-c_1x_2^2+2c_1x_2y_2-c_1(y_1^2+y_2^2)-c_2(y_3^2+y_4^2)-c_4(x_1^2+x_2^2)}.
\end{aligned} \tag{3.4.24}$$

To integrate over x_1 and x_2, we simply use ($a_4,a_5>0$)

$$\begin{aligned}
&\int_{-\infty}^{\infty}\mathrm{d}x_1\int_{-\infty}^{\infty}\mathrm{d}x_2\left(a_1x_1+a_2x_1^2+a_3x_1x_2\right)\mathrm{e}^{-a_4x_1^2-a_5x_2^2+a_6x_1+a_7x_2+a_8}\\
&\quad=\frac{\pi\left(2a_1a_4a_5a_6+a_2a_5\left(2a_4+a_6^2\right)+a_3a_4a_6a_7\right)\mathrm{e}^{\frac{4a_5a_8+a_7^2}{4a_5}+\frac{a_6^2}{4a_4}}}{4a_4^{\frac{5}{2}}a_5^{\frac{3}{2}}}.
\end{aligned} \tag{3.4.25}$$

Equations for the $\{a_i\}$ in terms of the $\{c_i\}$ can be read off Eq. (3.4.24) and the $\{c_i\}$ contain all parameter dependencies of the process, including the t'. The t'-integration is then performed numerically.

Phys. Rev. Lett. **129**, 140601 (2022)

Details and simulation parameters for Figs. 3.4.1-3.4.3

The process in Fig. 3.4.1a,b is simulated as a free two-dimensional Brownian motion. Numerical Stratonovich integration gives the empirical density and current. The times shown are $\tau_1^- = 1.14$, $\tau_1^+ = 3.83$, $\tau_2^- = 6.54$, $\tau_2^+ = 6.80$.

The mean and variance in the histograms in Fig. 1d,e and the relative error in Fig. 3.4.1f are obtained analytically as described above. The TUR-bound is given by $\frac{2}{\sigma t}$ where $\sigma = \frac{2\Omega^2}{r}$ is the dissipation in the steady state of the Ornstein-Uhlenbeck process (3.4.17) (see Sec. 3.5). For $\Omega = 5, r = 1, t = 5$, we obtain that the TUR-bound shown in Fig. 3.4.1f is at 0.008.

The process in Fig. 3.4.2 is the shear flow with $\mathbf{F}(x, y) = 2x\hat{\mathbf{y}}$ and $\mathbf{D} = \mathbf{1}$ from $(0, 0)$ to $(2, 0)$ in total time $t_2 - t_1 = 1$. It is simulated with time step size $\mathrm{d}t = 0.02$ as Brownian bridge in x-direction (exactly hits 2 after time 1) and then pick trajectories that hit $y_{\text{final}} = 0$ with deviation less than 0.02. Time-reversed and dual reversed trajectories are similarly from $(2, 0)$ to $(0, 0)$ with same or inverted shear. For each transition around $11,000 - 12,000$ trajectories were considered. Arrows are in the direction of the first/last step in discretized time.

The trajectory in Fig. 3.4.2e is sampled from an Ornstein-Uhlenbeck process Eq. (3.4.17) with $\Omega = 3, r = D = 1$ and total time $t \approx 37$.

The simulations in Fig. 3.4.3 are performed with time step $\mathrm{d}t = 10^{-4}$ and 8192 repetitions for 3a and 4096 repetitions for 3b. All simulations are performed by discretizing Eq. (3.4.17) and sampling the initial point $\mathbf{x}_0$ from the steady-state distribution $p_{\mathrm{s}}(\mathbf{x})$. Additional parameters in 3c are $h = 1, 0.25, 0.03$ from dark to bright.

Phys. Rev. Lett. **129**, 140601 (2022)

3.5 Coarse graining empirical densities and currents in continuous-space steady states (Phys. Rev. Res. 2022)

This section is a slightly adapted version of the publication C. Dieball and A. Godec, "Coarse graining empirical densities and currents in continuous-space steady states", *Phys. Rev. Res.* **4**, 033243 (2022), accompanying the Letter reproduced in Sec. 3.4.

My personal contribution as first author of the publication was in performing calculations and simulations, analyzing the results, and co-writing the manuscript.

Coarse graining empirical densities and currents in continuous-space steady states

Cai Dieball and Aljaž Godec

Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

Abstract

We present the conceptual and technical background required to describe and understand the correlations and fluctuations of the empirical density and current of steady-state diffusion processes on all timescales — observables central to statistical mechanics and thermodynamics on the level of individual trajectories. We focus on the important and non-trivial effect of a spatial coarse graining. Making use of a generalized time-reversal symmetry, we provide deeper insight about the physical meaning of fluctuations of the coarse-grained empirical density and current, and explain why a systematic variation of the coarse-graining scale offers an efficient method to infer bounds on a system's dissipation. Moreover, we discuss emerging symmetries in the statistics of the empirical density and current, and the statistics in the central-limit regime. More broadly, our work promotes the application of stochastic calculus as a powerful direct alternative to Feynman-Kac theory and path-integral methods.

Phys. Rev. Res. 4, 033243 (2022)

3.5.1 Introduction

A non-vanishing probability current [34, 44, 48, 68, 106, 126, 140, 157, 159–163, 166, 167, 172, 173] and entropy production [28, 54, 73, 104, 120, 122, 145, 186–188] are the hallmarks of non-equilibrium, manifested as transients during relaxation [49, 73, 120–124] or in non-equilibrium, current-carrying steady states [34, 47, 48, 106, 125, 127]. Genuinely irreversible, detailed balance violating dynamics emerge in the presence of non-conservative forces (e.g., shear or rotational flow) [189–192] or active driving in living matter fueled by ATP hydrolysis [32, 115, 116, 140, 146, 193–196]. Such systems are typically small and "soft" and thus subject to large thermal fluctuations. Single-molecule [115–119] and particle-tracking [130] experiments probe dynamical processes on the level of individual, stochastic trajectories. These are typically analyzed within the framework of "time-average statistical mechanics" [41, 48, 67, 130–134], i.e., by averaging along individual finite realizations yielding random quantities with non-trivial statistics.

Ergodic steady states are characterized by the (invariant) steady-state density $p_{\mathrm{s}}(\mathbf{x})$ and a steady-state probability current $\mathbf{j}_{\mathrm{s}}(\mathbf{x})$ in systems with a broken detailed balance. One can equivalently infer $p_{\mathrm{s}}(\mathbf{x})$ and $\mathbf{j}_{\mathrm{s}}(\mathbf{x})$ from an ensemble of statistically independent trajectories of an ergodic process, or from an individual but very long (i.e., ergodically long[6]) trajectory. To infer $p_{\mathrm{s}}(\mathbf{x})$ and $\mathbf{j}_{\mathrm{s}}(\mathbf{x})$ from individual sample paths one uses estimators that are called the *empirical density* and *empirical current*, respectively, defined as

$$\begin{aligned}\overline{\rho_{\mathbf{x}}^{U}}(t) &\equiv \frac{1}{t}\int_0^t U_{\mathbf{x}}^h(\mathbf{x}_\tau)\mathrm{d}\tau,\\ \overline{\mathbf{J}_{\mathbf{x}}^{U}}(t) &\equiv \frac{1}{t}\int_{\tau=0}^{\tau=t} U_{\mathbf{x}}^h(\mathbf{x}_\tau)\circ \mathrm{d}\mathbf{x}_\tau,\end{aligned} \tag{3.5.1}$$

where $U_{\mathbf{x}}^h(\mathbf{z})$ is a "window function" around a point $\mathbf{x}$ with a characteristic scale h (see Sec. 3.4) and $\circ\,\mathrm{d}\mathbf{x}_\tau$ denotes the Stratonovich integral, which both will be specified more precisely below. Notably, the Stratonovich integration $\circ\mathrm{d}\mathbf{x}_\tau$ in Eq. (3.5.1) is the correct way to make sense of the expression "$\dot{\mathbf{x}}_\tau \mathrm{d}\tau$", which is ill-defined since, for any τ with probability one, $|\dot{\mathbf{x}}_\tau| = \infty$ for overdamped Langevin dynamics [178]. Because $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$ is random, $\overline{\rho_{\mathbf{x}}^{U}}(t)$ and $\overline{\mathbf{J}_{\mathbf{x}}^{U}}(t)$ are fluctuating quantities. Notably, the empirical density and current are typically defined with a δ function, i.e., with $U_{\mathbf{x}}^{h\to 0}(\mathbf{z}) = \delta(\mathbf{x}-\mathbf{z})$ [34,43–45,107,110,141–144] . For a variety of reasons detailed below and in the companion Letter (see Sec. 3.4), we here define $U_{\mathbf{x}}^h$ with a finite length scale $h > 0$, such that $\overline{\rho_{\mathbf{x}}^{U}}(t)$ measures the time spent in the region $U_{\mathbf{x}}^h$ around $\mathbf{x}$ and $\overline{\mathbf{J}_{\mathbf{x}}^{U}}(t)$ the displacements in the region $U_{\mathbf{x}}^h$ around $\mathbf{x}$. Such a definition is in line with that of generalized currents in stochastic thermodynamics [48, 67, 131, 132] except that we here consider vector-valued currents. Important recent results on such generalized currents (however, without the notion of coarse graining) may be found in Refs. [70, 72, 126, 133, 134].

The fluctuations of $\overline{\rho_{\mathbf{x}}^{U}}(t)$ and $\overline{\mathbf{J}_{\mathbf{x}}^{U}}(t)$ may be interpreted as variances of fluctuating histograms. Namely, after "binning" into (hyper)volumes around points $\mathbf{x}$ (or in our language the coarse-graining around $\mathbf{x}$), often carried out on a grid, each individual trajectory yields a random histogram of occupation fractions or displacements. That is, the height of bins in the histogram reflects the time spent or displacement in said bin accumulated over all visits of the trajectory until time t for $\overline{\rho_{\mathbf{x}}^{U}}(t)$ and $\overline{\mathbf{J}_{\mathbf{x}}^{U}}(t)$, respectively, and is a fluctuating quantity due to the stochasticity of trajectories. The variance of these fluctuations quantifies the inference uncertainty. In Fig. 3.5.1 we show such histograms inferred from individual tra-

[6] An ergodic timescale is longer than any correlation time in the system.

Phys. Rev. Res. 4, 033243 (2022)

jectories of a two-dimensional harmonically confined overdamped diffusion in a rotational flow

$$\mathrm{d}\mathbf{x}_t = -\begin{bmatrix} 1 & -\Omega \\ \Omega & 1 \end{bmatrix} \mathbf{x}\mathrm{d}t + \sqrt{2}\mathrm{d}\mathbf{W}_t, \tag{3.5.2}$$

with Gaussian window

$$U^h_{\mathbf{x}}(\mathbf{z}) = \frac{1}{2\pi h^2} \exp\left[-\frac{(\mathbf{z}-\mathbf{x})^2}{2h^2}\right]. \tag{3.5.3}$$

For this process and window function we analytically solved all spatial integrals entering the results derived below (see Sec. 3.4), and numerically evaluated one remaining time integral.

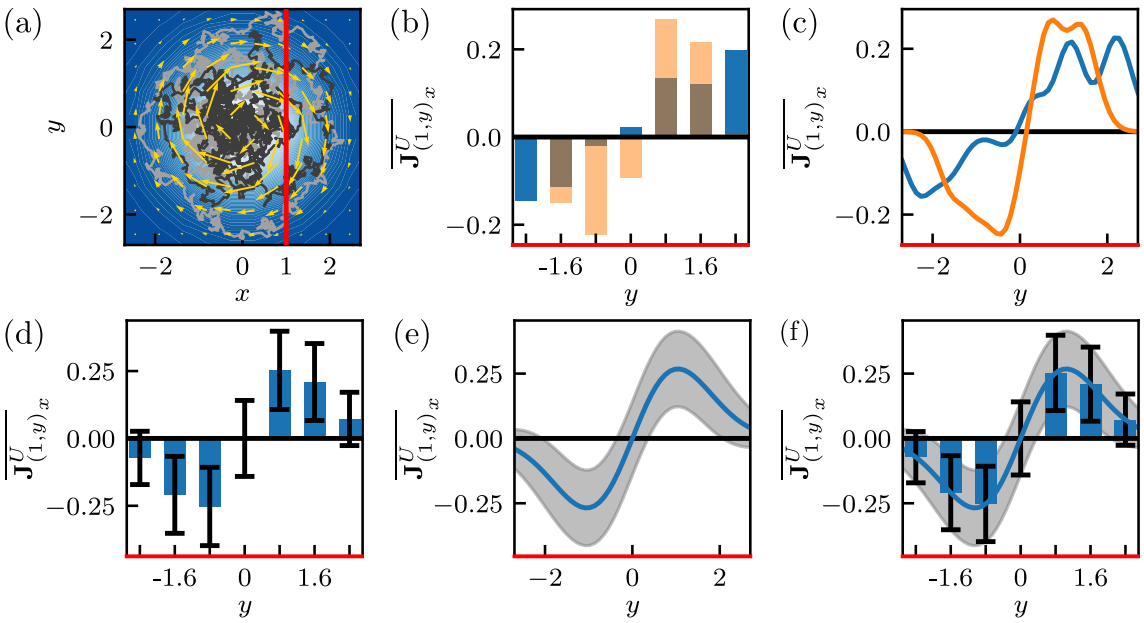

Figure 3.5.1: (a) Two trajectories (gray) with length $t = 5$ in harmonically confined rotational flow [Eq. (3.5.2)] with $\Omega = 5$. The steady-state density and current are depicted by the color gradient and yellow arrows, respectively. (b) Height of bins depicts the time-averaged x component of the current with Gaussian coarse-graining window [Eq. (3.5.3)] with $h = 0.3$ evaluated for several points on the red line in (a) for the two trajectories in (a). This corresponds to time averaging all local displacements (weighted by $U^h_{\mathbf{x}}$) within a single trajectory. (c) As in (b) but for the continuum of points on the red line in (a). This can be considered as the x component of the current smoothened over a scale h. (d) Mean value $\langle A \rangle_{\mathrm{s}}$ and standard deviation $\sqrt{\mathrm{var(A)}}$ of $A = \overline{\mathbf{J}^U_{(1,y)}}_x$ obtained from our result [Eq. (3.5.52)]. This represents the statistics of many histograms as in (b). (e) As in (d) but for continuous y as in (c). (f) Overlaying (d) and (e) shows that the histogram picture is fully contained in the continuous coarse-graining procedure.

Phys. Rev. Res. 4, 033243 (2022)

The interpretation of the coarse graining captured in or induced by $U_{\mathbf{x}}^h$ in Eq. (3.5.1) is flexible; it can represent a projection or a "generalized current" [48, 67, 70, 72, 126, 131–134] or may be thought of as a spatial smoothing of the empirical current and density as shown in Figs. 3.5.1c,e and 3.5.2, also for the case of a finite experimental resolution. Our main focus here is the smoothing aspect in the context of uncertainty of $p_{\mathrm{s}}(\mathbf{x}), \mathbf{j}_{\mathrm{s}}(\mathbf{x})$ and steady-state dissipation from individual trajectories. Note that some form of coarse graining or smoothing is in fact required in order for the quantities in Eq. (3.5.1) to be well-defined (see Sec. 3.4). A suitable smoothing decreases the uncertainty of the estimate and, if varied over sufficiently many h and $\mathbf{x}$ [see also Fig. 3.5.1c,e) instead of simply "binning," one does not necessarily lose information (as compared to input data). Moreover, a systematic variation of the scale h may reveal more information about $\overline{\rho_{\mathbf{x}}^U}(t)$ and $\overline{\mathbf{J}_{\mathbf{x}}^U}(t)$. The same reasoning is found to apply to generalized thermodynamic currents and allows for an improved inference of dissipation (see Sec. 3.4 and below).

The present work is an extended exposé of the conceptual and technical background that is required to understand and materialize the above observations. It accompanies the Letter reproduced in Sec. 3.4. Several additional explanations, illustrations and applications are given here.

The article is structured as follows. In Sec. 3.5.2 we lay out the theoretical background on stochastic differential equations in the Itô, Stratonovich and anti-Itô interpretations and the corresponding equations for the probability densities. We furthermore decompose the drift and steady-state current into conservative and nonconservative (i.e., irreversible) contributions and introduce dissipation. In Sec. 3.5.3 we prove a generalize time-reversal symmetry called "dual-reversal symmetry". In Sec. 3.5.4 we derive our main results for the steady-state (co)variances of $\overline{\rho_{\mathbf{x}}^U}(t)$ and $\overline{\mathbf{J}_{\mathbf{x}}^U}(t)$ and interpret them in terms of initial- and end-point currents and increments. We then use these results to explicitly evaluate the limit $h \to 0$ of no coarse graining in Sec. 3.5.5, where we find that fluctuations diverge in $(d \geq 2)$-dimensional space. In Sec. 3.5.6 we use current fluctuations to infer steady-state dissipation via the thermodynamic uncertainty relation (TUR) [47, 126] with an emphasis on the importance of the coarse-graining scale h. In particular, we demonstrate and explain the existence of a thermodynamically optimal coarse graining. In Sec. 3.5.7 we discuss symmetries obeyed by the (co)variances and explain how the results simplify in thermodynamic equilibrium, and in Sec. 3.5.8 we present a continuity equation for coarse grained empirical densities and currents. In Sec. 3.5.9 we present asymptotic results for short and long trajectories and give results for the central-limit regime. We conclude with an outlook beyond overdamped dynamics in Sec. 3.5.10 by considering underdamped sys-

Phys. Rev. Res. 4, 033243 (2022)

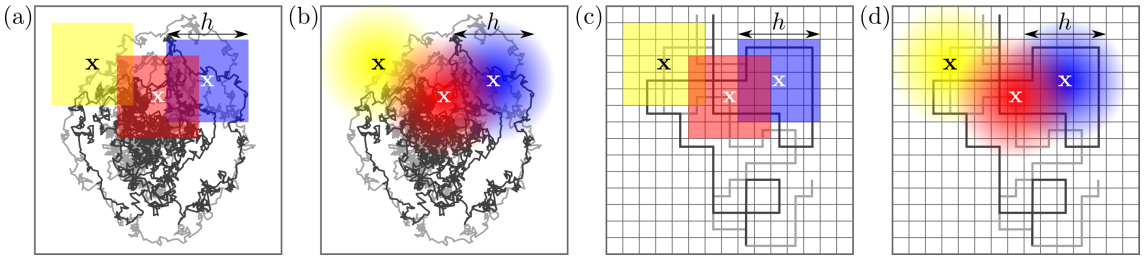

Figure 3.5.2: (a) Coarse-graining windows (colors) in the form of an indicator function of a rectangle centered at different points $\mathbf{x}$ with coarse-graining scale h. For each $\mathbf{x}$ and h, each trajectory (gray lines) gives rise to one value for the (coarse grained) time-averaged density and current. Note that the choice of $\mathbf{x}$ and h is flexible, such that the windows may overlap. (b) Same as (a) but with Gaussian coarse-graining windows. (c, d) Coarse-graining windows in the case of trajectory data with a finite experimental resolution (grid, gray trajectories). The coarse graining scale h should be chosen large compared to the resolution to obtain reliable approximations of the (coarse-grained) densities and currents.

tems as well as experimental data derived from particle-tracking experiments in biological cells, and with a summary and perspectives for the future.

3.5.2 Theory

Setup: Overdamped Langevin dynamics

In this section we provide background on the equations of motion for the coordinate $\mathbf{x}_\tau$ highlighting the differences between the Itô, Stratonovich, and anti-Itô interpretations, and for their corresponding conditional probability density functions of a transition $\mathbf{x}_0 \to \mathbf{x}$.

We consider time-homogeneous (i.e., coefficients do not explicitly depend on time) overdamped Langevin dynamics in d-dimensional space with (possibly) multiplicative noise [6, 16] described by the thermodynamically consistent [104, 169] anti-Itô (or Hänggi-Klimontovich [197, 198]) stochastic differential equation[7]

$$\mathrm{d}\mathbf{x}_\tau = \mathbf{F}(\mathbf{x}_\tau)\mathrm{d}\tau + \boldsymbol{\sigma}(\mathbf{x}_\tau) \circledast \mathrm{d}\mathbf{W}_\tau, \tag{3.5.4}$$

where $\mathrm{d}\mathbf{W}_\tau$ is the increment of a d-dimensional Wiener processes (i.e., white noise) with zero mean and covariance $\langle \mathrm{d}W_{\tau,i}\mathrm{d}W_{\tau',j}\rangle = \delta(\tau-\tau')\delta_{ij}\mathrm{d}\tau\mathrm{d}\tau'$. The noise amplitude is related to the diffusion coefficient via $\mathbf{D}(\mathbf{x}) \equiv \boldsymbol{\sigma}(\mathbf{x})\boldsymbol{\sigma}(\mathbf{x})^T/2$. We assume the drift field $\mathbf{F}(\mathbf{x})$ to be smooth and sufficiently confining, such that the anti-Itô (end-point) convention $\circledast\mathrm{d}\mathbf{W}_\tau = \mathbf{W}_\tau - \mathbf{W}_{\tau-d\tau}$ guarantees the existence of a steady-state probability density $p_\mathrm{s}(\mathbf{x}) = \mathrm{e}^{-\phi(\mathbf{x})}$

[7] For more precise statements in the case of multiplicative noise in multidimensional space see Sec. 2.8, in particular the additional condition in Eq. (2.8.6).

Phys. Rev. Res. 4, 033243 (2022)

and steady-state current $\mathbf{j}_{\mathrm{s}}(\mathbf{x})$, and yields the thermodynamically consistent Boltzmann-Gibbs (equilibrium) statistics when $\mathbf{D}(\mathbf{x})^{-1}\mathbf{F}(\mathbf{x}) = -\nabla\phi(\mathbf{x})$ is a potential force. The anti-Itô equation (3.5.4) can equivalently be rewritten as an Itô equation with an adapted drift as[8],

$$\mathrm{d}\mathbf{x}_t = \left[\mathbf{F}(\mathbf{x}_t) + \left\{\nabla^T\mathbf{D}\right\}(\mathbf{x}_t)\right]\mathrm{d}t + \sqrt{2\mathbf{D}(\mathbf{x}_t)}\mathrm{d}\mathbf{W}_t, \tag{3.5.5}$$

where the brackets $\{\cdot\}$ throughout denote that the differential operator only acts within the bracket.

At this point, several remarks are in order. First, the anti-Itô interpretation of the stochastic differential equation (3.5.4) as well as the Stratonovich integral in Eq. (3.5.1) are both required for thermodynamic consistency. Second, there is no difference between the interpretations of Eq. (3.5.4) if $\mathbf{D}(\mathbf{x}) = \mathbf{D}$ is a constant matrix; i.e., the convention only matters for multiplicative noise. However, even in this case the Stratonovich integral in Eq. (3.5.1) is required for thermodynamic consistency of the empirical current and to use it as an estimator of $\mathbf{j}_{\mathrm{s}}(\mathbf{x})$.

The Fokker-Planck equation for the conditional probability density $G(\mathbf{x}, t|\mathbf{y})$ to be at a point $\mathbf{x}$ at time t after starting at $\mathbf{y}$ that corresponds to Eqs. (3.5.4) and (3.5.5) reads

$$\begin{aligned}\partial_t G(\mathbf{x}, t|\mathbf{y}) &= [-\nabla_{\mathbf{x}}\cdot\mathbf{F}(\mathbf{x}) + \nabla_{\mathbf{x}}^T\mathbf{D}(\mathbf{x})\nabla_{\mathbf{x}}]G(\mathbf{x}, t|\mathbf{y}) \\ &\equiv L(\mathbf{x})G(\mathbf{x}, t|\mathbf{y}),\end{aligned} \tag{3.5.6}$$

which satisfies a continuity equation $(\partial_t + \nabla_{\mathbf{x}}\cdot\hat{\mathbf{j}}_{\mathbf{x}})G(\mathbf{x}, t|\mathbf{y}) = 0$, where

$$\hat{\mathbf{j}}_{\mathbf{x}} \equiv \mathbf{F}(\mathbf{x}) - \mathbf{D}(\mathbf{x})\nabla_{\mathbf{x}}. \tag{3.5.7}$$

Decomposing of the drift $\mathbf{F}(\mathbf{x})$ into reversible $\mathbf{F}^{\mathrm{rev}}(\mathbf{x}) = -\mathbf{D}(\mathbf{x})\{\nabla\phi\}(\mathbf{x})$ and irreversible $\mathbf{F}^{\mathrm{irrev}}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{F}^{\mathrm{rev}}(\mathbf{x})$ parts translates to a decomposition of $\hat{\mathbf{j}}_{\mathbf{x}}$ into a gradient part $\hat{\mathbf{j}}^g_{\mathbf{x}}$ and steady-state-current contributions, namely $\hat{\mathbf{j}}_{\mathbf{x}} = \mathbf{F}^{\mathrm{irrev}}(\mathbf{x}) + \mathbf{F}^{\mathrm{rev}}(\mathbf{x}) - \mathbf{D}(\mathbf{x})\nabla_{\mathbf{x}}$. This is rewritten using

$$\begin{aligned}\hat{\mathbf{j}}^g_{\mathbf{x}} &\equiv \mathbf{F}^{\mathrm{rev}}(\mathbf{x}) - \mathbf{D}(\mathbf{x})\nabla \\ &= \mathbf{D}(\mathbf{x})\left\{\nabla\log(p_{\mathrm{s}}(\mathbf{x})\right\} - \mathbf{D}(\mathbf{x})\nabla \\ &= \mathbf{D}(\mathbf{x})p_{\mathrm{s}}^{-1}(\mathbf{x})\left\{\nabla p_{\mathrm{s}}(\mathbf{x})\right\} - \mathbf{D}(\mathbf{x})\nabla \\ &= -\mathbf{D}(\mathbf{x})\left[p_{\mathrm{s}}(\mathbf{x})\{\nabla p_{\mathrm{s}}^{-1}(\mathbf{x})\} - \nabla\right]\end{aligned}$$

[8] For more precise statements in the case of multiplicative noise in multidimensional space see Sec. 2.8, in particular the additional condition in Eq. (2.8.6).

Phys. Rev. Res. 4, 033243 (2022)

$$= -\mathbf{D}(\mathbf{x})p_{\rm s}(\mathbf{x})\nabla p_{\rm s}^{-1}(\mathbf{x}), \tag{3.5.8}$$

where we have used that $\{\nabla p_{\rm s}(\mathbf{x})^{-1}\}=-p_{\rm s}^{-2}(\mathbf{x})\{\nabla p_{\rm s}\}(\mathbf{x})$ implies $\{\nabla p_{\rm s}\}(\mathbf{x})=-p_{\rm s}^{2}(\mathbf{x})\{\nabla p_{\rm s}^{-1}\}(\mathbf{x}))$. Therefore we have $\hat{\mathbf{j}}^{g}_{\mathbf{x}}p_{\rm s}(\mathbf{x}) = 0$, such that the definition of the steady-state current $\mathbf{j}_{\rm s}(\mathbf{x}) \equiv \hat{\mathbf{j}}_{\mathbf{x}}p_{\rm s}(\mathbf{x})$ with $\hat{\mathbf{j}}_{\mathbf{x}} = \hat{\mathbf{j}}^{g}_{\mathbf{x}} + \mathbf{F}^{\rm irrev}(\mathbf{x})$ implies $\mathbf{F}^{\rm irrev}(\mathbf{x}) = p_{\rm s}^{-1}(\mathbf{x})\mathbf{j}_{\rm s}(\mathbf{x})$ and we obtain

$$\begin{aligned}\hat{\mathbf{j}}_{\mathbf{x}} &= \hat{\mathbf{j}}^{g}_{\mathbf{x}} + p_{\rm s}^{-1}(\mathbf{x})\mathbf{j}_{\rm s}(\mathbf{x}) \\ &= -p_{\rm s}(\mathbf{x})\mathbf{D}(\mathbf{x})\nabla_{\mathbf{x}}p_{\rm s}^{-1}(\mathbf{x}) + p_{\rm s}^{-1}(\mathbf{x})\mathbf{j}_{\rm s}(\mathbf{x}).\end{aligned} \tag{3.5.9}$$

Moreover, note that the steady-state two-point density $P_{\mathbf{y}}(\mathbf{x},t) \equiv G(\mathbf{x},t|\mathbf{y})p_{\rm s}(\mathbf{y})$ also satisfies the same Fokker-Planck equation as $G(\mathbf{x},t|\mathbf{y})$.

Finally, if the process is irreversible, i.e., $\mathbf{F}^{\rm irrev}(\mathbf{x}) \neq \mathbf{0}$ the steady state is dissipative with an average total entropy production rate $\dot{\Sigma}$ given by [28, 105]

$$\begin{aligned}\dot{\Sigma} &= \int {\rm d}\mathbf{x}\mathbf{F}^{\rm irrev}(\mathbf{x})\cdot\mathbf{D}^{-1}(\mathbf{x})\mathbf{F}^{\rm irrev}(\mathbf{x})p_{\rm s}(\mathbf{x}) \\ &= \int {\rm d}\mathbf{x}\frac{\mathbf{j}_{\rm s}^{T}(\mathbf{x})}{p_{\rm s}(\mathbf{x})}\mathbf{D}^{-1}(\mathbf{x})\mathbf{j}_{\rm s}(\mathbf{x}),\end{aligned} \tag{3.5.10}$$

which can be obtained as the mean value of a sum over steady-state expectations of the respective i-th component of $\overline{\mathbf{J}^{U_i}_{\mathbf{x}}(t)}$ in Eq. (3.5.1) with $U_i = (\mathbf{F}^{\rm irrev}(\mathbf{x})^{T}\mathbf{D}^{-1}(\mathbf{x}))_i$.

Note that by adopting the Itô or Stratonovich conventions instead of the anti-Itô convention in Eq. (3.5.4) one obtains a different Fokker-Planck equation with a different steady-state density. In particular, $L^{\rm Ito}(\mathbf{x}) = -\nabla_{\mathbf{x}}\cdot\mathbf{F}(\mathbf{x}) + \sum_{i,j=1}^{d}\partial_i\partial_j D_{ij}(\mathbf{x})$ and $L^{\rm Strato}(\mathbf{x}) = L(\mathbf{x})/2 + L^{\rm Ito}(\mathbf{x})/2 = -\nabla_{\mathbf{x}}\cdot\mathbf{F}(\mathbf{x}) + \sum_{i,j=1}^{d}\partial_i\sqrt{D_{ij}(\mathbf{x})}\partial_j\sqrt{D_{ij}(\mathbf{x})}$ and the respective steady-state densities $p_{\rm s}^{\rm Ito}(\mathbf{x})$ and $p_{\rm s}^{\rm Strato}(\mathbf{x})$ depend explicitly on $\mathbf{D}(\mathbf{x})$ and are therefore in general not thermodynamically consistent since the steady state deviates from Gibbs-Boltzmann statistics [e.g., in dimension one we have $p_{\rm s}^{\rm Ito}(x) \propto p_{\rm s}^{\text{anti-Ito}}(x)/D(x)$ and $p_{\rm s}^{\rm Strato}(x) \propto p_{\rm s}^{\text{anti-Ito}}(x)/\sqrt{D(x)}$, respectively, where the deviation from $p_{\rm s}^{\text{anti-Ito}}(x)$ cannot be absorbed in the normalization if $D(x)$ depends on x].

3.5.3 Generalized time-reversal symmetry

It will later prove useful to take into account a form of generalized time-reversal symmetry obeyed by Eq. (3.5.4) called "continuous time reversal" or "dual-reversal symmetry" [133, 176]. Analogous generalized symmetries were also found in deterministic systems (see, e.g., Ref. [199]). Generalized time-reversal symmetry relates forward dynamics

Phys. Rev. Res. 4, 033243 (2022)

in non-equilibrium steady states to time-reversed dynamics in an ensemble with inverted irreversible steady-state current, i.e., in an ensemble with $\mathbf{F}^{\rm irrev} \to -\mathbf{F}^{\rm irrev}$ or equivalently $\mathbf{j}_{\rm s} \to -\mathbf{j}_{\rm s}$. The dual-reversal symmetry for the two-point probability densities states that

$$G(\mathbf{x},t|\mathbf{y})p_{\rm s}(\mathbf{y}) = G^{-\mathbf{j}_{\rm s}}(\mathbf{y},t|\mathbf{x})p_{\rm s}(\mathbf{x}), \tag{3.5.11}$$

or equivalently $G^{-\mathbf{j}_{\rm s}}(\mathbf{x},t|\mathbf{y})p_{\rm s}(\mathbf{y}) = G(\mathbf{y},t|\mathbf{x})p_{\rm s}(\mathbf{x})$ where $G^{-\mathbf{j}_{\rm s}}(\mathbf{y},t|\mathbf{x})$ is the conditional probability density of the process with drift $\mathbf{F}^{-\mathbf{j}_{\rm s}}(\mathbf{x}) \equiv \mathbf{F}^{\rm rev}(\mathbf{x}) - \mathbf{F}^{\rm irrev}(\mathbf{x})$ instead of $\mathbf{F}(\mathbf{x}) = \mathbf{F}^{\rm rev}(\mathbf{x}) + \mathbf{F}^{\rm irrev}(\mathbf{x})$. At equilibrium, i.e., $\mathbf{j}_{\rm s}(\mathbf{x}) = \mathbf{0}$ (for all $\mathbf{x}$), this symmetry simplifies to the well-known time-reversal symmetry called the "detailed balance" condition for two-point densities. We here provide an original and intuitive proof of Eq. (3.5.11) that proceeds entirely in continuous space and time, based on the decomposition of currents in Eq. (3.5.9). The Fokker-Planck operator $L(\mathbf{x}) = -\nabla_{\mathbf{x}} \cdot \hat{\mathbf{j}}_{\mathbf{x}}$, using the decomposition in Eq. (3.5.9) and multiplying by $p_{\rm s}$ from the right-hand side, reads

$$L(\mathbf{x})p_{\rm s}(\mathbf{x}) = -\nabla_{\mathbf{x}} \cdot \mathbf{j}_{\rm s}(\mathbf{x}) + \nabla_{\mathbf{x}}^T p_{\rm s}(\mathbf{x})\mathbf{D}(\mathbf{x})\nabla_{\mathbf{x}}. \tag{3.5.12}$$

Taking the adjoint gives (since $\mathbf{D} = \mathbf{D}^T$)

$$\begin{aligned} p_{\rm s}(\mathbf{x})L^\dagger(\mathbf{x}) &= [L(\mathbf{x})p_{\rm s}(\mathbf{x})]^\dagger \\ &= \mathbf{j}_{\rm s}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} + \nabla_{\mathbf{x}}^T p_{\rm s}(\mathbf{x})\mathbf{D}(\mathbf{x})\nabla_{\mathbf{x}}. \end{aligned} \tag{3.5.13}$$

Since for the steady state density $Lp_{\rm s} = 0$, $\mathbf{j}_{\rm s}$ is divergence free $\{\nabla_{\mathbf{x}} \cdot \mathbf{j}_{\rm s}(\mathbf{x})\} = 0$ and we have $\nabla_{\mathbf{x}} \cdot \mathbf{j}_{\rm s}(\mathbf{x}) = \mathbf{j}_{\rm s}(\mathbf{x}) \cdot \nabla_{\mathbf{x}}$. Thus we see the symmetry under inversion $\mathbf{j}_{\rm s} \to -\mathbf{j}_{\rm s}$

$$p_{\rm s}(\mathbf{x})L^\dagger(\mathbf{x}) = L^{-\mathbf{j}_{\rm s}}(\mathbf{x})p_{\rm s}(\mathbf{x}). \tag{3.5.14}$$

Under detailed balance $\mathbf{j}_{\rm s} = \mathbf{0}$, i.e., $L^{-\mathbf{j}_{\rm s}} = L$, and $p_{\rm s}(\mathbf{x})L^\dagger(\mathbf{x}) = L(\mathbf{x})p_{\rm s}(\mathbf{x})$ which implies the time-reversal symmetry $G(\mathbf{x},t|\mathbf{y})p_{\rm s}(\mathbf{y}) = G(\mathbf{y},t|\mathbf{x})p_{\rm s}(\mathbf{x})$ [16, 17, 178]. Equation (3.5.14) implies for all integers $n \geq 1$ that $p_{\rm s}(\mathbf{x})[L^\dagger(\mathbf{x})]^n = [L(\mathbf{x})^{-\mathbf{j}_{\rm s}}]^n(\mathbf{x})p_{\rm s}(\mathbf{x})$, and consequently for all $t \geq 0$ that $p_{\rm s}(\mathbf{x})\exp[L^\dagger(\mathbf{x})t] = \exp[L^{-\mathbf{j}_{\rm s}}(\mathbf{x})t]p_{\rm s}(\mathbf{x})$. Applying this operator equation to the initial condition $\delta(\mathbf{y}-\mathbf{x})$ and using $p_{\rm s}(\mathbf{x})\delta(\mathbf{y}-\mathbf{x}) = p_{\rm s}(\mathbf{y})\delta(\mathbf{y}-\mathbf{x})$ as well as that $L^\dagger$ propagates the initial condition as $G(\mathbf{y},t|\mathbf{x}) = \exp[L^\dagger(\mathbf{x})t]\delta(\mathbf{y}-\mathbf{x})$ while $L^{-\mathbf{j}_{\rm s}}$ propagates the final point in the ensemble with $\mathbf{j}_{\rm s}$ inverted as $G^{-\mathbf{j}_{\rm s}}(\mathbf{x},t|\mathbf{y}) = \exp[L^{-\mathbf{j}_{\rm s}}(\mathbf{x})t]\delta(\mathbf{y}-\mathbf{x})$, we obtain the dual-reversal symmetry in Eq. (3.5.11). This generalized time-reversal symmetry relates the dynamics in the time-reversed ensemble to the propagation in the ensemble with reversed current, or equivalently, the forward dynamics to the propagation with concurrent time and $\mathbf{j}_{\rm s}$-reversal. While at equilibrium (i.e.,

Phys. Rev. Res. 4, 033243 (2022)

under detailed balance, $\mathbf{j}_{\rm s} = \mathbf{0}$) the forward dynamics is indistinguishable from the time-reversed dynamics, the statement in Eq. (3.5.11) (if generalized to all paths; see, e.g., Ref. [133]) means that forward dynamics (with $\mathbf{j}_{\rm s}$) is indistinguishable from backwards or time-reversed dynamics with reversed $\mathbf{j}_{\rm s} \to -\mathbf{j}_{\rm s}$ [i.e., $\mathbf{j}_{\rm s}(\mathbf{x}) \to -\mathbf{j}_{\rm s}(\mathbf{x})$ at all $\mathbf{x}$]. We will later use this dual-reversal symmetry to understand the fluctuations of observables that involve (time-integrated) currents in non-equilibrium steady states.

3.5.4 Derivation of the main results, initial- and final-point currents and their application to density-current correlations

Mean empirical density and current

Although the time-averaged density and current defined in Eq. (3.5.1) are functionals with complicated statistics, their mean values can be readily computed. Throughout the paper we will assume steady-state initial conditions, i.e., initial conditions drawn from $p_{\rm s}(\mathbf{x}')$, denoted by $\langle\cdot\rangle_{\rm s}$. This renders mean values time independent, and we have (see also Ref. [106])

$$\begin{aligned}\left\langle \overline{\rho^U_{\mathbf{x}}}(t)\right\rangle_{\rm s} &= \frac{1}{t}\int_0^t \mathrm{d}\tau \left\langle U^h_{\mathbf{x}}(\mathbf{x}_\tau)\right\rangle_{\rm s}\\ &= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{z} U^h_{\mathbf{x}}(\mathbf{z}) p_{\rm s}(\mathbf{z})\\ &= \int \mathrm{d}\mathbf{z} U^h_{\mathbf{x}}(\mathbf{z}) p_{\rm s}(\mathbf{z}),\end{aligned} \tag{3.5.15}$$

and by rewriting the Stratonovich integration $\circ\mathrm{d}\mathbf{x}_\tau$ in terms of Itô integration as $U^h_{\mathbf{x}}(\mathbf{x}_\tau)\circ \mathrm{d}\mathbf{x}_\tau = U^h_{\mathbf{x}}(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau + \frac{1}{2}\mathrm{d}U^h_{\mathbf{x}}(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau$, where $\mathrm{d}\mathbf{x}_\tau \mathrm{d}\mathbf{x}_\tau^T/2 = \mathbf{D}(\mathbf{x}_\tau)\mathrm{d}\tau$ and thus $\mathrm{d}U^h_{\mathbf{x}}(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau/2 = \mathbf{D}(\mathbf{x}_\tau)\{\nabla U^h_{\mathbf{x}}\}(\mathbf{x}_\tau)\mathrm{d}\tau$,

$$\begin{aligned}&\left\langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t)\right\rangle_{\rm s} = \frac{1}{t}\int_0^t \left\langle U^h_{\mathbf{x}}(\mathbf{x}_\tau)\circ \mathrm{d}\mathbf{x}_\tau\right\rangle_{\rm s}\\ &= \frac{1}{t}\int_{\tau=0}^{\tau=t} \left\langle U^h_{\mathbf{x}}(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau\right\rangle_{\rm s} + \frac{1}{t}\int_{\tau=0}^{\tau=t}\frac{1}{2}\left\langle \mathrm{d}U^h_{\mathbf{x}}(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau\right\rangle_{\rm s}\\ &= \frac{1}{t}\int_0^t \mathrm{d}\tau\int \mathrm{d}\mathbf{z} p_{\rm s}(\mathbf{z})\big[U^h_{\mathbf{x}}(\mathbf{z})\mathbf{F}(\mathbf{z}) + \left\{\nabla^T_{\mathbf{z}}\mathbf{D}(\mathbf{z})\right\}U^h_{\mathbf{x}}(\mathbf{z})\\ &+ \mathbf{D}(\mathbf{z})\left\{\nabla_{\mathbf{z}}U^h_{\mathbf{x}}(\mathbf{z})\right\}\big] + \frac{1}{t}\int_{\tau=0}^{\tau=t}\left\langle U^h_{\mathbf{x}}(\mathbf{x}_\tau)\sqrt{2\mathbf{D}(\mathbf{x}_\tau)}\mathrm{d}\mathbf{W}_\tau\right\rangle_{\rm s}.\end{aligned} \tag{3.5.16}$$

Phys. Rev. Res. 4, 033243 (2022)

Note that the mean value involving $\mathrm{d}\mathbf{W}_\tau$ vanishes since this Itô-noise increment has zero mean and is uncorrelated with functions of $\mathbf{x}_\tau$, i.e., $\langle f(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau\rangle = \langle f(\mathbf{x}_\tau)\rangle\langle \mathrm{d}\mathbf{W}_\tau\rangle = 0$. Integrating by parts and using that $\mathbf{D}(\mathbf{z}) = \mathbf{D}^T(\mathbf{z})$ is symmetric we get

$$\begin{aligned}\left\langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t)\right\rangle_{\mathrm{s}} &= \int \mathrm{d}\mathbf{z} p_{\mathrm{s}}(\mathbf{z}) \left[U^h_{\mathbf{x}}(\mathbf{z})\mathbf{F}(\mathbf{z}) + \nabla^T_{\mathbf{z}}\mathbf{D}(\mathbf{z})U^h_{\mathbf{x}}(\mathbf{z})\right] \\ &= \int \mathrm{d}\mathbf{z} U^h_{\mathbf{x}}(\mathbf{z}) \left[\mathbf{F}(\mathbf{z}) - \mathbf{D}(\mathbf{z})\nabla_{\mathbf{z}}\right] p_{\mathrm{s}}(\mathbf{z}) \\ &= \int \mathrm{d}\mathbf{z} U^h_{\mathbf{x}}(\mathbf{z})\hat{\mathbf{j}}_{\mathbf{z}} p_{\mathrm{s}}(\mathbf{z}) = \int \mathrm{d}\mathbf{z} U^h_{\mathbf{x}}(\mathbf{z})\mathbf{j}_{\mathrm{s}}(\mathbf{z}). \end{aligned} \tag{3.5.17}$$

Note that if we had defined Eq. (3.5.1) with an Itô integral instead of the Stratonovich, we would miss the $\mathbf{D}(\mathbf{z})\nabla_{\mathbf{z}}$ term and would *not* get $\hat{\mathbf{j}}_{\mathbf{z}}$ and thus $\mathbf{j}_{\mathrm{s}}$, not even for additive noise. The Stratonovich integral is therefore required for consistency.

The interpretation of the steady-state mean values in Eqs. (3.5.15) and (3.5.17) is immediate — the mean time-averaged density and current are (at least for positive normalized windows) the steady-state density p_{s} and current $\mathbf{j}_{\mathrm{s}}$ averaged over the coarse-graining window function $U^h_{\mathbf{x}}$.

(Co)variances of empirical density and current

Since fluctuations [41, 47, 48, 67, 126, 130–133] (and correlations [134]) play a crucial role in time-average statistical mechanics and stochastic thermodynamics, we discuss (co)variances of coarse-grained time-averaged densities and currents (recall the interpretation of the variance within the "fluctuating histogram" picture in Fig. 3.5.1).

To keep the notation tractable, we introduce the integral operator

$$\hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}[\,\cdot\,] \equiv \frac{1}{t^2}\int_0^t \mathrm{d}t_1 \int_{t_1}^t \mathrm{d}t_2 \int \mathrm{d}\mathbf{z} U^h_{\mathbf{x}}(\mathbf{z}) \int \mathrm{d}\mathbf{z}' U^h_{\mathbf{y}}(\mathbf{z}')[\,\cdot\,], \tag{3.5.18}$$

with the convention $\int_{t_1}^t \mathrm{d}t_2 \delta(t_2 - t_1) = 1/2$. Note that other conventions would only change the appearance of intermediate steps but not the final result. We define the two-point steady-state covariance according to Sec. 3.4 as

$$C^{\mathbf{xy}}_{AB}(t) \equiv \langle A_{\mathbf{x}}(t)B_{\mathbf{y}}(t)\rangle_{\mathrm{s}} - \langle A_{\mathbf{x}}(t)\rangle_{\mathrm{s}}\langle B_{\mathbf{y}}(t)\rangle_{\mathrm{s}}\,, \tag{3.5.19}$$

where A and B are henceforth either $\overline{\rho^U}$ or $\overline{\mathbf{J}^U}$, respectively. We refer to the case when $A \neq B$ or $\mathbf{x} \neq \mathbf{y}$ as (linear) "correlations" and to the case $A = B$ with $\mathbf{x} = \mathbf{y}$ as "fluctuations" whereby we adopt the convention $\mathrm{var}^{\mathbf{x}}_A(t) \equiv C^{\mathbf{xx}}_{AA}(t)$. Note that for simplicity and enhanced readability we only assume coarse-graining windows $U^h_{\mathbf{x}}$ and $U^h_{\mathbf{y}}$ where the shape is fixed

Phys. Rev. Res. 4, 033243 (2022)

but the center points $\mathbf{x}, \mathbf{y}$ may differ. All results equivalently hold for window functions whose shape and h differ as well.

We now address correlations $C^{\mathbf{xy}}_{\rho\rho}$ of the coarse-grained time-averaged density at points $\mathbf{x}$ and $\mathbf{y}$, which correspond to the density variance when $\mathbf{x} = \mathbf{y}$. To do so, first consider the (mixed) second moment

$$t^2 \left\langle \overline{\rho^U_{\mathbf{x}}}(t) \overline{\rho^U_{\mathbf{y}}}(t) \right\rangle_{\mathrm{s}} = \int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' \left\langle U^h_{\mathbf{x}}(\mathbf{x}_\tau) U^h_{\mathbf{y}}(\mathbf{x}_{\tau'}) \right\rangle_{\mathrm{s}} . \tag{3.5.20}$$

The expectation value corresponds to an integration over the two-point probability density to have $\mathbf{x}_\tau = \mathbf{z}$ and $\mathbf{x}_{\tau'} = \mathbf{z}'$ given by the two-point function $P_{\mathbf{z}}(\mathbf{z}', \tau' - \tau) \equiv G(\mathbf{z}', \tau' - \tau|\mathbf{z}) p_{\mathrm{s}}(\mathbf{z})$ for $\tau' > \tau$ and $P_{\mathbf{z}'}(\mathbf{z}, \tau - \tau')$ for $\tau' < \tau$. We relabel the times τ, τ' as $t_1 < t_2$ and use the integral operator in Eq. (3.5.18) to obtain

$$\left\langle \overline{\rho^U_{\mathbf{x}}}(t) \overline{\rho^U_{\mathbf{y}}}(t) \right\rangle_{\mathrm{s}} = \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}} \left[P_{\mathbf{z}}(\mathbf{z}', t_2 - t_1) + P_{\mathbf{z}'}(\mathbf{z}, t_2 - t_1) \right] . \tag{3.5.21}$$

Since the argument only depends on time differences $t' = t_2 - t_1 \geq 0$ the integral operator Eq. (3.5.18) simplifies to

$$\hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}[\,\cdot\,] \equiv \frac{1}{t} \int_0^t \mathrm{d}t' \left(1 - \frac{t'}{t}\right) \int \mathrm{d}\mathbf{z} U^h_{\mathbf{x}}(\mathbf{z}) \int \mathrm{d}\mathbf{z}' U^h_{\mathbf{y}}(\mathbf{z}')[\,\cdot\,]. \tag{3.5.22}$$

To obtain the correlation we subtract the mean values [see Eq. (3.5.15)] which [noting that $(1/t)\int_0^t \mathrm{d}t'(1 - t'/t) = 1/2$] gives

$$C^{\mathbf{xy}}_{\rho\rho}(t) = \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}} \left[P_{\mathbf{z}}(\mathbf{z}', t') + P_{\mathbf{z}'}(\mathbf{z}, t') - 2 p_{\mathrm{s}}(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}') \right] , \tag{3.5.23}$$

which has been derived before [41, 135]. Equation (3.5.23) simplifies further for $\mathbf{x} = \mathbf{y}$ as well as under detailed balance and is also symmetric under $\mathbf{j}_{\mathrm{s}} \to -\mathbf{j}_{\mathrm{s}}$, all of which will be discussed in Sec. 3.5.7.

The interpretation of Eq. (3.5.23) (see also Ref. [41]) is that all paths from $\mathbf{z}$ to $\mathbf{z}'$ (i.e., from $U^h_{\mathbf{x}}$ to $U^h_{\mathbf{y}}$) and vice versa from $\mathbf{z}'$ to $\mathbf{z}$, in time $t' = t_2 - t_1$ contribute according to their correlation to $C^{\mathbf{xy}}_{\rho\rho}(t)$. These contributions are integrated over all possible time differences and pairs of points within $U^h_{\mathbf{x}}$ and $U^h_{\mathbf{y}}$, respectively.

We now explore the important effect of coarse graining over the windows $U^h_{\mathbf{x}}$ for the inference of $p_{\mathrm{s}}(\mathbf{x})$ from noisy individual trajectories. If one wants to reliably infer the (coarse-grained) steady-state density from $\overline{\rho^U_{\mathbf{x}}}(t)$ the relative error $\mathrm{var}_\rho / \langle \overline{\rho^U_{\mathbf{x}}}(t) \rangle^2$ should be small. We have shown that $\lim_{h\to 0} \mathrm{var}_\rho / \langle \overline{\rho^U_{\mathbf{x}}}(t) \rangle^2 = \infty$ (see Sec. 3.4) and Fig. 3.5.3 (blue line) demonstrates that $\mathrm{var}_\rho / \langle \overline{\rho^U_{\mathbf{x}}}(t) \rangle^2$ decreases with increasing h. However, such a decrease

Phys. Rev. Res. 4, 033243 (2022)

does not guarantee an improved inference. Namely, as $h \to \infty$ the time spent in the region around $\mathbf{x}$ tends to t and $U^h_{\mathbf{x}}$ becomes constant on a large region and hence $\overline{\rho^U_{\mathbf{x}}}(t) \to U^h_{\mathbf{x}}(\mathbf{x})$ which contains no information about $p_{\mathrm{s}}(\mathbf{x})$. Therefore, to reliably infer that $\overline{\rho^U_{\mathbf{x}}}$ significantly deviates from $U^h_{\mathbf{x}}(\mathbf{x})$, we must also consider the relative error of $[\overline{\rho^U_{\mathbf{x}}} - U^h_{\mathbf{x}}(\mathbf{x})]$ depicted in Fig. 3.5.3 (orange line). There exists an "optimal coarse graining" where the uncertainty of simultaneously inferring $\overline{\rho^U_{\mathbf{x}}}$ and $\overline{\rho^U_{\mathbf{x}}} - U^h_{\mathbf{x}}(\mathbf{x})$ is minimal (minimum of the solid lines in Fig. 3.5.3) which represents the most reliable and informative estimate of $\overline{\rho^U_{\mathbf{x}}}$. In Sec. 3.5.6 we will turn to an analogous "optimal coarse graining" with respect to current variances and a system's dissipation.

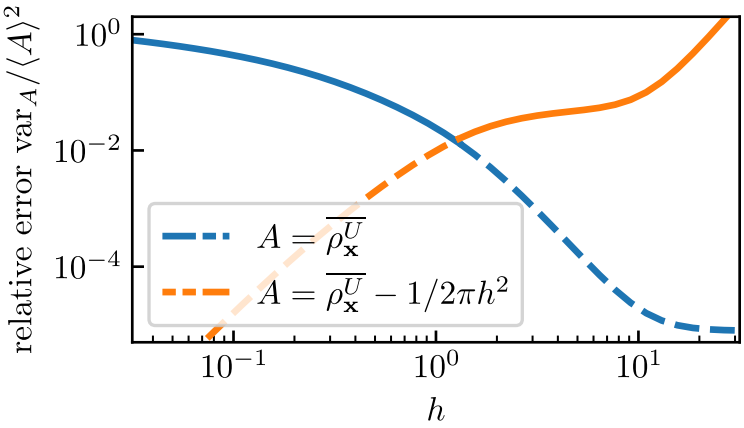

Figure 3.5.3: Relative error of $\overline{\rho^U_{\mathbf{x}}}(t)$ (blue line) compared to the relative error of $[\overline{\rho^U_{\mathbf{x}}}(t) - U^h_{\mathbf{x}}(\mathbf{x})]$ (orange line) as a function of the coarsening scale h for the rotational flow in Eq. (3.5.2) with $\Omega = 3$ for time $t = 10$ with a Gaussian window function Eq. (3.5.3) around $\mathbf{x} = (1, 0)^T$ with width h, i.e., $U^h_{\mathbf{x}}(\mathbf{x}) = (2\pi h^2)^{-1}$. The intersection point of blue and orange lines at $h \approx 1.3$ yields an "optimal coarse graining" where the maximum of the two lines (solid line) is minimal, whereas the maximum of the relative errors diverges as $h \to \infty$ since $\langle A\rangle_{\mathrm{s}} \to 0$ and diverges logarithmically for $h \to 0$ [as we will see in Eq. (3.5.51)].

We now consider coarse-grained time-averaged currents. To compute the correlation of the current at a point $\mathbf{x}$ and the density at $\mathbf{y}$ we need to consider

$$t^2 \left\langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t)\overline{\rho^U_{\mathbf{y}}}(t)\right\rangle_{\mathrm{s}} = \int_0^t \mathrm{d}\tau' \int_{\tau=0}^{\tau=t} \left\langle U^h_{\mathbf{x}}(\mathbf{x}_\tau)U^h_{\mathbf{y}}(\mathbf{x}_{\tau'}) \circ \mathrm{d}\mathbf{x}_\tau \right\rangle_{\mathrm{s}}. \tag{3.5.24}$$

Relabeling with $t_1 \le t_2$, introducing the notation

$$\langle \cdots \rangle^{\mathbf{x}_{t_1}=\mathbf{z}}_{\mathbf{x}_{t_2}=\mathbf{z}'} \equiv \left\langle \delta(\mathbf{x}_{t_1} - \mathbf{z})\delta(\mathbf{x}_{t_2} - \mathbf{z}') \cdots \right\rangle_{\mathrm{s}}, \tag{3.5.25}$$

and considering the Stratonovich increments

$$\circ \mathrm{d}\mathbf{x}_\tau \equiv \mathbf{x}_{\tau+\mathrm{d}\tau/2} - \mathbf{x}_{\tau-\mathrm{d}\tau/2}, \tag{3.5.26}$$

Phys. Rev. Res. 4, 033243 (2022)

and subtracting the mean values (3.5.15) and (3.5.17), we can write the correlation as

$$\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t) = \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}\left[\frac{\langle \circ \mathrm{d}\mathbf{x}_{t_1}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}}{\mathrm{d}t_1} + \frac{\langle \circ \mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'}}{\mathrm{d}t_2} - 2\mathbf{j}_{\mathrm{s}}(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}')\right]. \tag{3.5.27}$$

Equation (3.5.27) is harder to compute and more difficult to interpret as compared to $C^{\mathbf{xy}}_{\rho\rho}(t)$ [see Eq. (3.5.23)]. The quantities involving Stratonovich increments characterize the mean initial- and final displacements of "pinned" paths of duration $t_2 - t_1$ conditioned on the initial and final points $\mathbf{z}, \mathbf{z}'$ or $\mathbf{z}', \mathbf{z}$, respectively. Note that $\mathbf{z}$ always denotes the point where the increment occurs. Via the integral operator in Eq. (3.5.18) or Eq. (3.5.22) the $\mathbf{z}$ variable is integrated over $U^h_{\mathbf{x}}(\mathbf{z})$; i.e., in $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t)$ the variable $\mathbf{z}$ corresponds to the window at $\mathbf{x}$ where the (coarse-grained) current is evaluated. Therefore, correlations between a current and a density depend on integrals over conditioned initial-point increments at a point $\mathbf{z}$ at time t_1, and conditioned final-point increments, also at $\mathbf{z}$, at time $t_2 > t_1$. We define the increments divided by $\mathrm{d}t_i$ to be the "initial- and final-point currents,"

$$\begin{aligned}\mathbf{j}_{\mathrm{in}}(\mathbf{z}', t_2 - t_1; \mathbf{z}) &\equiv \frac{\langle \circ \mathrm{d}\mathbf{x}_{t_1}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}}{\mathrm{d}t_1},\\ \mathbf{j}_{\mathrm{fi}}(\mathbf{z}, t_2 - t_1; \mathbf{z}') &\equiv \frac{\langle \circ \mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'}}{\mathrm{d}t_2}.\end{aligned} \tag{3.5.28}$$

In order to understand the correlation in Eq. (3.5.27) we must therefore understand initial- and final-point currents. This is *a priori* not easy, since initial-point currents involve both, spatial increments at t_1 and probabilities of reaching a final point at time $t_2 > t_1$, which involves non-trivial correlations — a given displacement affects (and thus correlates with) the probability to reach the final point. We will derive a statement ("Lemma") in the next section that solves all mathematical difficulties related to this issue, without resorting to Feynman-Kac and path-integral methods as in Ref. [86] or Sec. 3.7. Then we will make intuitive sense of the result by exploiting the dual-reversal symmetry in Eq. (3.5.11).

Before doing so, we also consider the scalar current-current covariance $C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ [note that the complete fluctuations and correlations of $\overline{\mathbf{J}^U_{\mathbf{x}}}(t)$ are characterized by the $d\times d$ covariance matrix with elements $[\mathsf{C}^{\mathbf{xy}}_{\mathbf{JJ}}(t)]_{ik} = C^{\mathbf{xy}}_{\mathrm{J}_i\mathrm{J}_k}(t)$; here we focus on the scalar case $C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) \equiv \mathrm{Tr}\mathsf{C}^{\mathbf{xy}}_{\mathbf{JJ}}(t)$]. Notably, almost all results remain completely equivalent for other elements of the covariance matrix, scalar products simply have a slightly more intuitive geometrical

Phys. Rev. Res. 4, 033243 (2022)

interpretation and notation. Writing down the definition and using the notations as in the steps towards Eq. (3.5.27) we immediately arrive at

$$C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) = \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}\left[\frac{\langle \circ\mathrm{d}\mathbf{x}_{t_1}\cdot\circ\mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}}{\mathrm{d}t_1\mathrm{d}t_2} + \frac{\langle \circ\mathrm{d}\mathbf{x}_{t_1}\cdot\circ\mathrm{d}\mathbf{x}_{t_2}\rangle^{\mathbf{x}_{t_2}=\mathbf{z}}_{\mathbf{x}_{t_1}=\mathbf{z}'}}{\mathrm{d}t_1\mathrm{d}t_2} - 2\mathbf{j}_{\mathrm{s}}(\mathbf{z})\cdot\mathbf{j}_{\mathrm{s}}(\mathbf{z}')\right], \quad (3.5.29)$$

which is similar to the correlation in Eq. (3.5.27) but involves an average over scalar products of initial- and final-point increments along individual trajectories "pinned" at initial and end points. We will return to Eq. (3.5.29) and solve for these increments in Sec. 3.5.4 upon having explained the density-current correlation.

Lemma

To be able to treat expressions involving the increments correlated with future positions, we need a technical lemma that will turn out to be very powerful and central to all calculations. The required statement can also be obtained from the more general concept of Doob conditioning [104, 107, 133, 200], but here we provide a direct proof. Consider an Itô noise increment $\sqrt{2\mathbf{D}(\mathbf{x}_\tau)}\mathrm{d}\mathbf{W}_\tau$ [or equivalently $\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau$] with $\mathrm{d}\mathbf{W}_\tau = \mathbf{W}_{\tau+\mathrm{d}\tau} - \mathbf{W}_\tau$. In the following we will need to compute the expected values involving expressions like

$$\star = \left\langle \left[\sqrt{2\mathbf{D}(\mathbf{x}_\tau)}\mathrm{d}\mathbf{W}_\tau\right]_k U(\mathbf{x}_\tau)V(\mathbf{x}_{\tau'})\right\rangle_{\mathrm{s}}, \quad (3.5.30)$$

where $U(\mathbf{x}')$ and $V(\mathbf{x}')$ are arbitrary differentiable, square integrable functions, the subsubscript k denotes the kth component, and the subsubscript s denotes that the process evolves from $p_{\mathrm{s}}(\mathbf{x}')$. Correlations of $\mathrm{d}\mathbf{W}_\tau = \mathbf{W}_{\tau+\mathrm{d}\tau} - \mathbf{W}_\tau$ with any function of $\mathbf{x}_{\tau'}$ at a time $\tau' \leq \tau$ vanish by construction of the Wiener process (it has nominally independent increments). However, correlations with functions at $\tau' > \tau$ are non-trivial.

Note that given an initial point $\mathbf{x}_0 = \mathbf{z}$ and setting $\sqrt{2\mathbf{D}(\mathbf{z})}\mathrm{d}\mathbf{W}_\mathbf{0} = \boldsymbol{\varepsilon}$, the Itô/Langevin equation in Eq. (3.5.5) predicts a displacement $\mathrm{d}\mathbf{x}_0(\mathbf{z},\boldsymbol{\varepsilon}) = [\mathbf{F}(\mathbf{z}) + \nabla^T_\mathbf{z}\mathbf{D}(\mathbf{z})]\mathrm{d}t' + \boldsymbol{\varepsilon}$. With this we can write the expectation in Eq. (3.5.30) for $\tau = 0 < t' = \tau'$ as ε_k integrated over the probability to be at points $\mathbf{z}, \mathbf{z}+\mathrm{d}\mathbf{x}_0(\mathbf{z},\boldsymbol{\varepsilon}), \mathbf{z}'$ at times $0, \mathrm{d}t', t'$, i.e.,

$$\begin{aligned}\star = &\int \mathrm{d}\mathbf{z}\int \mathrm{d}\mathbf{z}' U(\mathbf{z})V(\mathbf{z}')\\ &\times \int \mathrm{d}\boldsymbol{\varepsilon}\,\mathbb{P}(\boldsymbol{\varepsilon})\varepsilon_k G(\mathbf{z}', t'-\mathrm{d}t'|\mathbf{z}+\mathrm{d}\mathbf{x}_0(\mathbf{z},\boldsymbol{\varepsilon}))p_{\mathrm{s}}(\mathbf{z}),\end{aligned} \quad (3.5.31)$$

where the probability $\mathbb{P}(\boldsymbol{\varepsilon})$ of $\sqrt{2\mathbf{D}(\mathbf{z})}\mathrm{d}\mathbf{W}_\mathbf{0} = \boldsymbol{\varepsilon}$ is given by a Gaussian distribution with zero mean and covariance matrix $2\mathbf{D}(\mathbf{z})\mathrm{d}t'$. Since this distribution is symmetric around $\mathbf{0}$,

Phys. Rev. Res. 4, 033243 (2022)

only terms with even powers of the components of $\boldsymbol{\varepsilon}$ survive the $\mathbb{P}(\boldsymbol{\varepsilon})$ integration. Noting that for $\mathrm{d}t' \to 0$ we have $G(\mathbf{z}', t' - \mathrm{d}t'|\mathbf{z} + \mathrm{d}\mathbf{x}_0(\mathbf{z}, \boldsymbol{\varepsilon})) \to [1 + \mathrm{d}\mathbf{x}_0(\mathbf{z}, \boldsymbol{\varepsilon}) \cdot \nabla_{\mathbf{z}}] G(\mathbf{z}', t'|\mathbf{z})$, we see that the only even power of the components of $\boldsymbol{\varepsilon}$ in $\varepsilon_k G(\cdots)$ gives

$$\star = \int \mathrm{d}\mathbf{z} \int \mathrm{d}\mathbf{z}' U(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}) V(\mathbf{z}') \int \mathrm{d}\boldsymbol{\varepsilon}\, \mathbb{P}(\boldsymbol{\varepsilon}) \varepsilon_k \boldsymbol{\varepsilon} \cdot \nabla_{\mathbf{z}} G(\mathbf{z}', t'|\mathbf{z}), \tag{3.5.32}$$

which using $\int \mathrm{d}\boldsymbol{\varepsilon} \mathbb{P}(\boldsymbol{\varepsilon}) \varepsilon_k \varepsilon_j = 2 D_{kj}(\mathbf{z}) \mathrm{d}t'$ yields the result

$$\star = \int \mathrm{d}\mathbf{z} \int \mathrm{d}\mathbf{z}' U(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}) V(\mathbf{z}') \left[2\mathbf{D}(\mathbf{z}) \nabla_{\mathbf{z}} G(\mathbf{z}', t'|\mathbf{z}) \right]_k \mathrm{d}t'. \tag{3.5.33}$$

Rewritten in terms of $P_{\mathbf{z}}(\mathbf{z}', t') \equiv G(\mathbf{z}', t'|\mathbf{z}) p_{\mathrm{s}}(\mathbf{z})$ and $\hat{\mathbf{j}}^g_{\mathbf{z}} \equiv -p_{\mathrm{s}}(\mathbf{z}) \mathbf{D}(\mathbf{z}) \nabla_{\mathbf{z}} p_{\mathrm{s}}^{-1}(\mathbf{z})$ we have $\hat{\mathbf{j}}^g_{\mathbf{z}} P_{\mathbf{z}}(\mathbf{z}', t') = -p_{\mathrm{s}}(\mathbf{z}) \mathbf{D}(\mathbf{z}) \nabla_{\mathbf{z}} G(\mathbf{z}', t'|\mathbf{z})$, and thus

$$\star = -2 \int \mathrm{d}\mathbf{z} \int \mathrm{d}\mathbf{z}' U(\mathbf{z}) V(\mathbf{z}') \left[\hat{\mathbf{j}}^g_{\mathbf{z}} \right]_k P_{\mathbf{z}}(\mathbf{z}', t') \mathrm{d}t'. \tag{3.5.34}$$

Motivated by the dual-reversal symmetry and the anticipated applications we define the dual-reversed current operator by inverting $\hat{\mathbf{j}}$ and concurrently inverting $\mathbf{j}_{\mathrm{s}} \to -\mathbf{j}_{\mathrm{s}}$, i.e.,

$$\begin{aligned} \hat{\mathbf{j}}^{\dagger}_{\mathbf{x}} \equiv -\hat{\mathbf{j}}^{-\mathbf{j}_{\mathrm{s}}}_{\mathbf{x}} &= -\left[\hat{\mathbf{j}}^g_{\mathbf{x}} - p_{\mathrm{s}}^{-1}(\mathbf{x}) \mathbf{j}_{\mathrm{s}}(\mathbf{x}) \right] \\ &= p_{\mathrm{s}}(\mathbf{x}) \mathbf{D}(\mathbf{x}) \nabla_{\mathbf{x}} p_{\mathrm{s}}^{-1}(\mathbf{x}) + p_{\mathrm{s}}^{-1}(\mathbf{x}) \mathbf{j}_{\mathrm{s}}(\mathbf{x}). \end{aligned} \tag{3.5.35}$$

Since $\hat{\mathbf{j}}^{\dagger}_{\mathbf{z}} - \hat{\mathbf{j}}_{\mathbf{z}} = -2\hat{\mathbf{j}}^g_{\mathbf{z}}$ we can rewrite Eqs. (3.5.33) and (3.5.34) as

$$\star = \int \mathrm{d}\mathbf{z} \int \mathrm{d}\mathbf{z}' U(\mathbf{z}) V(\mathbf{z}') \left(\hat{\mathbf{j}}^{\dagger}_{\mathbf{z}} - \hat{\mathbf{j}}_{\mathbf{z}} \right)_k P_{\mathbf{z}}(\mathbf{z}', t') \mathrm{d}t', \tag{3.5.36}$$

which will turn out to be the crucial part of the following calculations and will allow for an intuitive interpretation of the results in terms of dual-reversed dynamics.

Application of the lemma to initial- and final-point currents

In order to quantify and understand the density-current correlation expression in Eq. (3.5.27), we now turn back to the initial- and final-point currents, recalling the definitions in Eq. (3.5.28). These observables characterize the mean initial- and final displacements of "pinned" paths of duration $t_2 - t_1$ conditioned on the respective initial and final points $\mathbf{z}, \mathbf{z}'$ or $\mathbf{z}', \mathbf{z}$. The fact that both are currents in $\mathbf{z}$ justifies the name "initial- and final-point current." Such objects turn out to play a crucial role in the evaluation and understanding of correlations of densities and currents [see Eq. (3.5.27)]. The computation of

Phys. Rev. Res. 4, 033243 (2022)

current variances in fact involves the expectation of scalar products of such displacements [see Eq. (3.5.29)], but we first focus on simple displacements.

Final-point currents can be computed by subsubstituting for $\circ\mathrm{d}\mathbf{x}_\tau$ and integrating by parts as in Eq. (3.5.17),

$$\begin{aligned}\frac{\langle\circ\mathrm{d}\mathbf{x}_{t_2}\rangle_{\mathbf{x}_{t_1}=\mathbf{z}'}^{\mathbf{x}_{t_2}=\mathbf{z}}}{\mathrm{d}t_2} =& \int \mathrm{d}\mathbf{z}_1 \int \mathrm{d}\mathbf{z}_2 \delta(\mathbf{z}_1-\mathbf{z}')\delta(\mathbf{z}_2-\mathbf{z})P_{\mathbf{z}_1}(\mathbf{z}_2,t_2-t_1)[\mathbf{F}(\mathbf{z}_2)+\nabla_{\mathbf{z}_2}^T\mathbf{D}(\mathbf{z}_2)]\\ =&[\mathbf{F}(\mathbf{z})-\mathbf{D}(\mathbf{z})\nabla_{\mathbf{z}}]P_{\mathbf{z}'}(\mathbf{z},t_2-t_1)\\ =&\hat{\mathbf{j}}_{\mathbf{z}}P_{\mathbf{z}'}(\mathbf{z},t_2-t_1),\end{aligned} \tag{3.5.37}$$

where the Itô term involving $\mathrm{d}\mathbf{W}_{t_2}$ vanishes whereas the Stratonovich correction term survives. Therefore, the final-point current is obtained from the two-point density and current operator, both appearing in the Fokker-Planck equation [recall that $(\partial_t+\nabla_{\mathbf{x}}\cdot\hat{\mathbf{j}}_{\mathbf{x}})P_{\mathbf{y}}(\mathbf{x},t)=0$]

$$\mathbf{j}_{\mathrm{fi}}(\mathbf{z},t_2-t_1;\mathbf{z}')=\hat{\mathbf{j}}_{\mathbf{z}}P_{\mathbf{z}'}(\mathbf{z},t_2-t_1). \tag{3.5.38}$$

For the initial-point current analogous computations yield an Itô increment as a correction

$$\mathbf{j}_{\mathrm{in}}(\mathbf{z}',t_2-t_1;\mathbf{z})=\hat{\mathbf{j}}_{\mathbf{z}}P_{\mathbf{z}}(\mathbf{z}',t_2-t_1)+\langle\sqrt{2\mathbf{D}(\mathbf{x}_{t_1})}\mathrm{d}\mathbf{W}_{t_1}\rangle_{\mathbf{x}_{t_1}=\mathbf{z}}^{\mathbf{x}_{t_2}=\mathbf{z}'}. \tag{3.5.39}$$

Note that the latter Itô increment also appears in the calculations in Eqs. (3.5.17) and (3.5.37), but its mean vanishes since it involves end-point increments $\mathrm{d}\mathbf{W}_{t_2}$ (note t_2 and not t_1), which are by construction uncorrelated with the evolution up to time t_2. The correction term here does *not* vanish since the increment at time t_1 is correlated with the probability to reach $\mathbf{z}'$ at time t_2. Therefore, this expectation is non-trivial, but fortunately we solved this problem with the Lemma derived in Eqs. (3.5.30)–(3.5.36).

When U and V in Eq. (3.5.36) tend to a Dirac delta function (which is mathematically not problematic since we later integrate over $\mathbf{z},\mathbf{z}'$), we obtain

$$\langle\sqrt{2\mathbf{D}(\mathbf{x}_{t_1})}\mathrm{d}\mathbf{W}_{t_1}\rangle_{\mathbf{x}_{t_1}=\mathbf{z}}^{\mathbf{x}_{t_2}=\mathbf{z}'}=\left(\hat{\mathbf{j}}_{\mathbf{z}}^{\dagger}-\hat{\mathbf{j}}_{\mathbf{z}}\right)P_{\mathbf{z}}(\mathbf{z}',t_2-t_1), \tag{3.5.40}$$

which gives, recalling Eq. (3.5.35),

$$\mathbf{j}_{\mathrm{in}}(\mathbf{z}',t_2-t_1;\mathbf{z})=\hat{\mathbf{j}}_{\mathbf{z}}^{\dagger}P_{\mathbf{z}}(\mathbf{z}',t_2-t_1). \tag{3.5.41}$$

Note that $\mathbf{j}_{\mathrm{in}}(\mathbf{y},t;\mathbf{x},0)=-\mathbf{j}_{\mathrm{fi}}^{-\mathbf{j}_{\mathrm{s}}}(\mathbf{x},t;\mathbf{y},0)$ in agreement with dual-reversal symmetry.

Phys. Rev. Res. 4, 033243 (2022)

To better understand these currents and their symmetry, we require some intuition about the generalized time-reversal symmetry (i.e., the dual-reversal symmetry), which we gain on the basis of a simple overdamped shear flow in Fig. 3.5.4. Consider an isotropic diffusion with additive noise in a shear flow $\mathrm{d}\mathbf{x}_\tau = \mathbf{F}_{\mathrm{sh}}(\mathbf{x}_\tau)\mathrm{d}\tau + \sqrt{2}\mathrm{d}\mathbf{W}_\tau$ with $\mathbf{F}_{\mathrm{sh}}((x,y)^T) = (0,2x)^T$ (see gray arrows in Fig. 3.5.4a-c). For simplicity, we here only consider shear flow in a flat potential, such that strictly speaking a steady-state density p_{s} does not exist. The existence of p_{s} is in fact not necessary for the discussion in this section, nor to connect this example to a genuine non-equilibrium steady state. One may equally consider the shear flow to be confined in a box that is large enough to allow to neglect boundary effects at times before t and yet would yield flat p_{s} as $t \to \infty$. The drift of the unconfined shear flow is purely irreversible, i.e., $\mathbf{F}_{\mathrm{sh}}^{\mathrm{rev}}(\mathbf{x}') = \mathbf{0}$. Thus, inverting the irreversible part completely inverts the drift $\mathbf{F}_{\mathrm{sh}}^{-\mathbf{j}_{\mathrm{s}}}(\mathbf{x}') = -\mathbf{F}_{\mathrm{sh}}(\mathbf{x}')$ (see blue arrows in Fig. 3.5.4a,d). The initial-point current (purple arrow in Fig. 3.5.4b) is difficult to understand, since it correlates with the constraint to reach the end point after time t'. In the case of detailed balance, the time-reversal symmetry would allow to obtain this initial-point current as the inverted final-point current (yellow arrow in Fig. 3.5.4c). However, since detailed balance is broken by the shear flow this does not suffice. Instead, one has to consider the final-point current for the dynamics with the inverted irreversible drift (blue arrow in Fig. 3.5.4d). According to $\mathbf{j}_{\mathrm{in}}(\mathbf{y},t;\mathbf{x},0) = -\mathbf{j}_{\mathrm{fi}}^{-\mathbf{j}_{\mathrm{s}}}(\mathbf{x},t;\mathbf{y},0)$ and as can be seen in Fig. 3.5.4a, this allows to obtain the cumbersome initial-point current (yellow) as the inverted final-point current (blue).

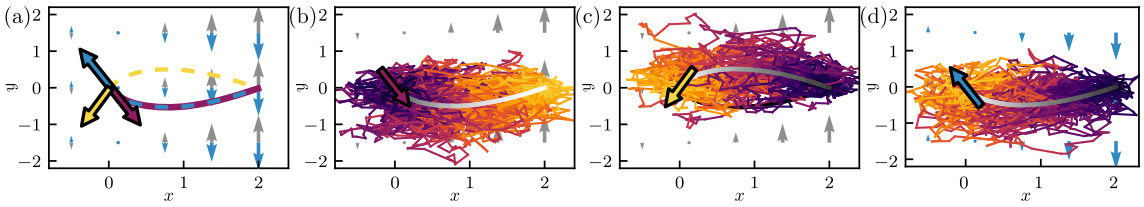

Figure 3.5.4: (a) Shear drift (gray background arrows) and inverted shear drift (blue background arrows) as described in the text, and currents and paths from (b) to (d) shown in purple, yellow, and blue. We see that the purple arrow equals the inverted blue arrow, and the purple line overlaps with the blue dashed line, as implied by Eq. (3.5.43). (b) Simulated trajectories in the shear flow (gray background arrows) from $\mathbf{z} = (0,0)^T$ to $\mathbf{z}' = (2,0)^T$ in time $t' = 1$ with time always running from dark to bright. The initial-point current, i.e., the initial-point increment averaged over all trajectories, is depicted by the purple arrow and the mean paths (averaged over all trajectories) by the gray curve. (c) As in (b) but from $\mathbf{z}' = (2,0)^T$ to $\mathbf{z} = (0,0)^T$ and final-point current depicted by a yellow arrow. (d) As in (c) but with the inverted shear flow depicted by blue arrows in the background.

In addition to the initial- and final-point currents, we also depict in Fig. 3.5.4 the mean "pinned" paths. In Fig. 3.5.4a we see that the forward and dual-reversed paths (purple

Phys. Rev. Res. 4, 033243 (2022)

and blue dashed lines) overlap. This can also be seen from the dual-reversal symmetry in Eq. (3.5.11).

To prove the equality of mean paths consider $0 < \tau < t'$ where $t' = t_2 - t_2 > 0$. The (non-random) point $\boldsymbol{\mu}(\tau) \equiv \langle \mathbf{x}_{t_1+\tau} \rangle_{\mathbf{x}_{t_1}=\mathbf{z}}^{\mathbf{x}_{t_2}=\mathbf{z}'}$ on the mean path $\mathbf{z} \to \mathbf{z}'$ is given by an integral over all possible intermediate points $\boldsymbol{\mu}(\tau) = \mathbf{x}$ weighted by $G(\mathbf{z}', t' - \tau|\mathbf{x})G(\mathbf{x}, \tau|\mathbf{z})/G(\mathbf{z}', t'|\mathbf{z})$ (since $\mathbf{x}_\tau$ is a Markov process) which gives the Chapman-Kolmogorov-like equation

$$G(\mathbf{z}', t'|\mathbf{z})\boldsymbol{\mu}(\tau) = \int \mathrm{d}\mathbf{x} G(\mathbf{z}', t' - \tau|\mathbf{x})G(\mathbf{x}, \tau|\mathbf{z})\mathbf{x}. \tag{3.5.42}$$

The corresponding point on the mean dual-reversed path $\boldsymbol{\mu}^\ddagger(\tau) \equiv \langle \mathbf{x}_{t_2-\tau}^{-\mathbf{j}_\mathrm{s}} \rangle_{\mathbf{x}_{t_1}=\mathbf{z}'}^{\mathbf{x}_{t_2}=\mathbf{z}}$ from $\mathbf{z}'$ to $\mathbf{z}$ with reversed steady-state current $\mathbf{j}_\mathrm{s} \to -\mathbf{j}_\mathrm{s}$ is given by [using three times the dual-reversal in Eq. (3.5.11)]

$$\begin{aligned} G^{-\mathbf{j}_\mathrm{s}}(\mathbf{z}, t'|\mathbf{z}')\boldsymbol{\mu}^\ddagger(t' - \tau) &= \int \mathrm{d}\mathbf{x} G^{-\mathbf{j}_\mathrm{s}}(\mathbf{z}, \tau|\mathbf{x})G^{-\mathbf{j}_\mathrm{s}}(\mathbf{x}, t' - \tau|\mathbf{z}')\mathbf{x} \\ &= \int \mathrm{d}\mathbf{x} G(\mathbf{x}, \tau|\mathbf{z}) \frac{p_\mathrm{s}(\mathbf{z})}{p_\mathrm{s}(\mathbf{x})} G(\mathbf{z}', t' - \tau|\mathbf{x}) \frac{p_\mathrm{s}(\mathbf{x})}{p_\mathrm{s}(\mathbf{z}')}\mathbf{x} \\ &= \frac{p_\mathrm{s}(\mathbf{z})}{p_\mathrm{s}(\mathbf{z}')} G(\mathbf{z}', t'|\mathbf{z})\boldsymbol{\mu}(\tau) \\ &= G^{-\mathbf{j}_\mathrm{s}}(\mathbf{z}, t'|\mathbf{z}')\boldsymbol{\mu}(\tau), \end{aligned} \tag{3.5.43}$$

which implies $\boldsymbol{\mu}(\tau) = \boldsymbol{\mu}^\ddagger(t' - \tau)$ for all $t_1 < \tau < t_2$, so the mean paths indeed agree (but run in opposite directions), which completes the proof that the blue and purple paths in Fig. 3.5.4a overlap.

Current-density correlation

With the definitions (3.5.28) and $t' = t_2 - t_1 > 0$ we have [recall the simplification of $\hat{\mathcal{I}}_{\mathbf{xy}}^{t,U}$ in Eq. (3.5.22)]

$$\mathbf{C}_{\mathbf{J}\rho}^{\mathbf{xy}}(t) = \hat{\mathcal{I}}_{\mathbf{xy}}^{t,U}[\mathbf{j}_\mathrm{fi}(\mathbf{z}, t'; \mathbf{z}') + \mathbf{j}_\mathrm{in}(\mathbf{z}', t'; \mathbf{z}) - 2\mathbf{j}_\mathrm{s}(\mathbf{z})p_\mathrm{s}(\mathbf{z}')]. \tag{3.5.44}$$

As we have shown in Eqs. (3.5.38) and (3.5.41) the initial- and final-point currents can be expressed in terms of the current operators, yielding

$$\mathbf{C}_{\mathbf{J}\rho}^{\mathbf{xy}}(t) = \hat{\mathcal{I}}_{\mathbf{xy}}^{t,U}[\hat{\mathbf{j}}_\mathbf{z} P_{\mathbf{z}'}(\mathbf{z}, t') + \hat{\mathbf{j}}_\mathbf{z}^\ddagger P_\mathbf{z}(\mathbf{z}', t') - 2\mathbf{j}_\mathrm{s}(\mathbf{z})p_\mathrm{s}(\mathbf{z}')], \tag{3.5.45}$$

which allows to explicitly calculate $\mathbf{C}_{\mathbf{J}\rho}^{\mathbf{xy}}(t)$ if $P_{\mathbf{z}'}(\mathbf{z}, t')$ is known. An analogous result for the scalar current variance was very recently obtained in Ref. [133] but did not establish

Phys. Rev. Res. 4, 033243 (2022)

a connection to current operators and dual-reversal symmetry and did not consider coarse graining or multidimensional continuous-space examples. The current-density correlation $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t)$ can be interpreted analogous to $\mathbf{C}^{\mathbf{xy}}_{\rho\rho}(t)$ as follows.

All possible paths between points $\mathbf{z}, \mathbf{z}'$ in time $0 < t' \leq t$ contribute, weighted by their corresponding probability, to this correlation. The difference with respect to density correlations $\mathbf{C}^{\mathbf{xy}}_{\rho\rho}(t)$ is that now currents at position $\mathbf{z}$ are correlated with probabilities to be at the point $\mathbf{z}'$. For paths $\mathbf{z}' \to \mathbf{z}$ the displacement is obtained from the familiar current operator $\mathbf{j}_{\text{fi}} = \hat{\mathbf{j}}_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z}, t')$. Paths from $\mathbf{z} \to \mathbf{z}'$ are mathematically more involved (and somewhat harder to understand), but can be understood intuitively with the dual-reversal symmetry (see also Fig. 3.5.4). More precisely, they can be understood and calculated in terms of the dual-reversed current operator $\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}} \equiv -\hat{\mathbf{j}}^{-\mathbf{j}_{\text{s}}}_{\mathbf{z}}$.

A direct observation that follows from the result in Eq. (3.5.45) is that at equilibrium (i.e., under detailed balance), we have $\mathbf{j}_{\text{s}} = \mathbf{0}$, $\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}} = -\hat{\mathbf{j}}_{\mathbf{z}}$ and $P_{\mathbf{z}}(\mathbf{z}', t') = P_{\mathbf{z}'}(\mathbf{z}, t')$ and thus $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t) = \mathbf{0}$ for all window functions and all points $\mathbf{x}, \mathbf{y}$. The correlation $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t)$ can also be utilized to improve the TUR, as recently shown in Ref. [134]. The result in Eq. (3.5.45) thus allows to inspect and understand more deeply this improved TUR.

An explicit example of the correlation result in Eq. (3.5.45) for $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t)$ is shown in Fig. 3.5.5. In line with the previous arguing, $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t)$ can be understood as a vector with initial- and final-point contributions, $C^{\mathbf{xy}}_{\mathbf{J}\rho} = \mathbf{C}_{\text{in}} + \mathbf{C}_{\text{fi}}$, where $\mathbf{C}_{\text{in}} \equiv \hat{\mathcal{I}}^{t}_{\mathbf{xy}}[\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}} P_{\mathbf{z}}(\mathbf{z}', t') - \mathbf{j}_{\text{s}}(\mathbf{z}) p_{\text{s}}(\mathbf{z}')]$. In the Appendix we shown that for $\mathbf{x} = \mathbf{y}$ in the limit $h \to 0$ of small windows the results for the correlation simplify $\mathbf{C}_{\text{in}}(t) \simeq [2\mathbf{j}_{\text{s}}(\mathbf{x})/p_{\text{s}}(\mathbf{x}) - \mathbf{F}(\mathbf{x})]\text{var}^{\mathbf{x}}_{\rho}(t)/4$ and $\mathbf{C}_{\text{fi}}(t) \simeq \mathbf{F}(\mathbf{x})\text{var}^{\mathbf{x}}_{\rho}(t)/4$, implying $\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t) \simeq \mathbf{j}_{\text{s}}(\mathbf{x})\text{var}^{\mathbf{x}}_{\rho}(t)/2p_{\text{s}}(\mathbf{x})$. Since $\mathbf{F} = \mathbf{F}^{\text{rev}} + \mathbf{j}_{\text{s}}/p_{\text{s}}$ and thus $2\mathbf{j}_{\text{s}}(\mathbf{x})/p_{\text{s}}(\mathbf{x}) - \mathbf{F}(\mathbf{x}) = -\mathbf{F}^{-\mathbf{j}_{\text{s}}}(\mathbf{x})$, the above implies that for $\mathbf{x} = \mathbf{y}$ and small windows h we have $-\mathbf{C}_{\text{in}} = \mathbf{C}^{-\mathbf{j}_{\text{s}}}_{\text{fi}}$ and $\mathbf{C}_{\text{fi}}$ points along $\mathbf{F}(\mathbf{x})$ that is tangent to the mean trajectory $[\mu]$ at $\mathbf{x}$, while $\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t)$ points in the $\mathbf{j}_{\text{s}}(\mathbf{x})$ direction (see Fig. 3.5.5b). For longer times t and/or larger h, the direction of $\mathbf{C}_{\text{fi}}$ changes but $-\mathbf{C}_{\text{in}} = \mathbf{C}^{-\mathbf{j}_{\text{s}}}_{\text{fi}}$ still holds (see Fig. 3.5.5c) since the symmetry $\mathbf{j}_{\text{in}}(\mathbf{y}, t; \mathbf{x}, 0) = -\mathbf{j}^{-\mathbf{j}_{\text{s}}}_{\text{fi}}(\mathbf{x}, t; \mathbf{y}, 0)$ can be applied in the integrands. Conversely, the two-point correlation $C^{\mathbf{xy}}_{\mathbf{J}\rho}$ *need not* point along $\mathbf{j}_{\text{s}}(\mathbf{x})$ (Fig. 3.5.5d). In fact, its direction changes over time (see inset of Fig. 3.5.5d). Notably, results for $\mathbf{x} \neq \mathbf{y}$ akin to Fig. 3.5.5d may provide deeper insight into barrier-crossing problems on the level of individual trajectories in the absence of detailed balance.

Current (co)variance

Recall that the current (co)variance in Eq. (3.5.29) involves scalar products of initial- and final-point increments $\langle \circ \mathrm{d}\mathbf{x}_{t_1} \cdot \circ \mathrm{d}\mathbf{x}_{t_2} \rangle^{\mathbf{x}_{t_2} = \mathbf{z}'}_{\mathbf{x}_{t_1} = \mathbf{z}}$, which *cannot* be easily interpreted as scalar

Phys. Rev. Res. 4, 033243 (2022)

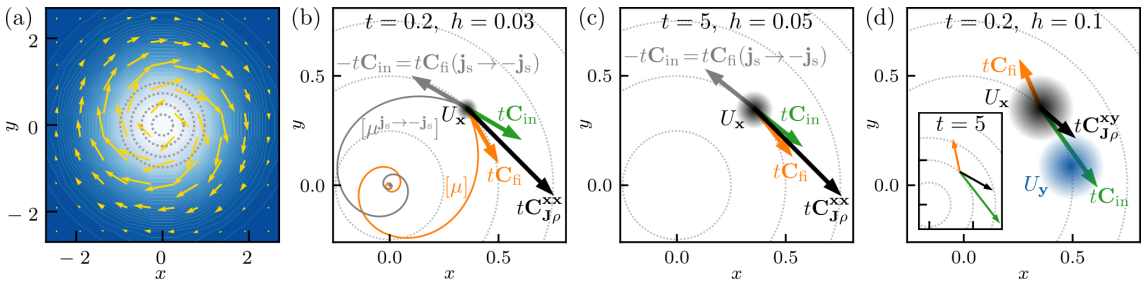

Figure 3.5.5: (a) Illustration of the steady-state density (color gradient) and current (arrows) of the two-dimensional rotational flow in Eq. (3.5.2) with $\Omega = 3$. Gray dotted lines in (a)–(d) are circles with radii $0.25, 0.5, 0.75, 1$. (b, c) Single-point $\mathbf{x} = \mathbf{y}$ and (d) two-point time-accumulated correlation $t\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}$ at $t = 0.2$ and $t = 5$ (black arrow), with final point $\mathbf{C}_{\rm fi} \equiv \hat{\mathcal{I}}^t_{\mathbf{xy}}[\hat{\mathbf{j}}_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z}, t') - \mathbf{j}_{\rm s}(\mathbf{z}) p_{\rm s}(\mathbf{z}')]$ (orange) and initial-point $\mathbf{C}_{\rm in}$ (green) contribution, such that $\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho} = \mathbf{C}_{\rm in} + \mathbf{C}_{\rm fi}$. $\mathbf{C}_{\rm fi}(\mathbf{j}_{\rm s} \to -\mathbf{j}_{\rm s})$ (gray) is the current-reversed final-point contribution which agrees with the inverted initial-point contribution $-\mathbf{C}_{\rm in}$. Solid lines in (b) are the mean trajectory $[\mu] \equiv \langle \mathbf{x}_{\tau \geq 0} \rangle_{\mathbf{x}_0 = \mathbf{x}}$ (orange) and its current reverse $[\mu^{\mathbf{j}_{\rm s}\to-\mathbf{j}_{\rm s}}]$ (gray). $U_{\mathbf{x},\mathbf{y}}$ (shaded circles) is a Gaussian at $\mathbf{x}, \mathbf{y}$ with width h [see Eq. (3.5.3)].

Phys. Rev. Res. 4, 033243 (2022)

products of currents. They are *not* the scalar products of initial- and final-point currents, since $\langle \circ \mathrm{d}\mathbf{x}_{t_1} \cdot \circ \mathrm{d}\mathbf{x}_{t_2} \rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}} \neq \langle \circ \mathrm{d}\mathbf{x}_{t_1} \rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}} \cdot \langle \circ \mathrm{d}\mathbf{x}_{t_2} \rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}$. Rather they correspond to the scalar product of the initial- and final-point increment *along the same trajectory* and only then they become averaged over all trajectories from $\mathbf{z}$ to $\mathbf{z}'$ (see also Fig. 3.4.2 Sec. 3.4). For $t_1 < t_2$ these are computed equivalently to Eqs. (3.5.37)–(3.5.41) based on the Lemma (3.5.36) as

$$\frac{\langle \circ \mathrm{d}\mathbf{x}_{t_1} \cdot \circ \mathrm{d}\mathbf{x}_{t_2} \rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}}}{\mathrm{d}t_1 \mathrm{d}t_2} = \hat{\mathbf{j}}^{\dagger}_{\mathbf{z}} \cdot \hat{\mathbf{j}}_{\mathbf{z}'} P_{\mathbf{z}}(\mathbf{z}', t'). \tag{3.5.46}$$

However, according to the convention $\int_{t_1}^{t} \mathrm{d}t_2 \delta(t_2 - t_1) = 1/2$ in Eq. (3.5.18), we also need to consider the case $t_1 = t_2$, i.e., $t' = 0$, which did not contribute for $C_{\rho\rho}$ and $C_{\mathbf{J}\rho}$. In the case $t_1 = t_2$ [recall the definition in Eq. (3.5.25)],

$$\begin{aligned} \langle \circ \mathrm{d}\mathbf{x}_{t_1} \cdot \circ \mathrm{d}\mathbf{x}_{t_2} \rangle^{\mathbf{x}_{t_2}=\mathbf{z}'}_{\mathbf{x}_{t_1}=\mathbf{z}} &\equiv \big\langle \delta(\mathbf{x}_{t_1} - \mathbf{z}) \circ \mathrm{d}\mathbf{x}_{t_1} \cdot \delta(\mathbf{x}_{t_2} - \mathbf{z}') \circ \mathrm{d}\mathbf{x}_{t_2} \big\rangle_{\rm s} \\ &\overset{t_1=t_2}{=} \Big\langle \delta(\mathbf{x}_{t_1} - \mathbf{z}) \sqrt{2\mathbf{D}(\mathbf{x}_{t_1})} \mathrm{d}\mathbf{W}_{t_1} \cdot \delta(\mathbf{x}_{t_2} - \mathbf{z}') \sqrt{2\mathbf{D}(\mathbf{x}_{t_2})} \mathrm{d}\mathbf{W}_{t_2} \Big\rangle, \end{aligned} \tag{3.5.47}$$

where we used that for $t_1 = t_2$ the only term surviving is $\mathrm{d}\mathbf{W}^2_{t_1}$ (and not $\mathrm{d}\mathbf{W}_{t_1}\mathrm{d}t_1$ and $\mathrm{d}t^2_1$, which is why such terms only enter in current-current expressions but not in current-

density or density-density correlations). Using $P_{\mathbf{z}}(\mathbf{z}', t'=0) = \delta(\mathbf{z}-\mathbf{z}')p_{\mathrm{s}}(\mathbf{z})$ we find for $t_1 = t_2$

$$\begin{aligned}\langle \circ \mathrm{d}\mathbf{x}_{t_1} \cdot \circ \mathrm{d}\mathbf{x}_{t_2} \rangle_{\mathbf{x}_{t_1}=\mathbf{z}}^{\mathbf{x}_{t_2}=\mathbf{z}'} &= P_{\mathbf{z}}(\mathbf{z}',0) \sum_{i,j,l=1}^{d} [\sqrt{2\mathbf{D}(\mathbf{z})}]_{ij} [\sqrt{2\mathbf{D}(\mathbf{z}')}]_{il} \delta(t_1-t_2)\delta_{jl}\mathrm{d}t_1\mathrm{d}t_2 \\ &= p_{\mathrm{s}}(\mathbf{z})\delta(\mathbf{z}-\mathbf{z}') \sum_{i=1}^{d} [2\mathbf{D}(\mathbf{z})]_{ii}\delta(t_1-t_2)\mathrm{d}t_1\mathrm{d}t_2 \\ &= 2\mathrm{Tr}[\mathbf{D}(\mathbf{z})]p_{\mathrm{s}}(\mathbf{z})\delta(\mathbf{z}-\mathbf{z}')\delta(t_1-t_2)\mathrm{d}t_1\mathrm{d}t_2. \end{aligned} \tag{3.5.48}$$

Plugging this into Eq. (3.5.29), we obtain, using Eq. (3.5.46) and accounting for the $t'=0$ contribution, the result for current covariances in the form of

$$\begin{aligned} C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) =& \frac{2}{t}\int d\mathbf{z}\mathrm{Tr}[\mathbf{D}(\mathbf{z})]U^h_{\mathbf{x}}(\mathbf{z})U^h_{\mathbf{y}}(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}) \\ &+ \hat{\mathcal{I}}^{t,U}_{\mathbf{xy}}[\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}'}\cdot\hat{\mathbf{j}}_{\mathbf{z}}P_{\mathbf{z}'}(\mathbf{z},t') + \hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}\cdot\hat{\mathbf{j}}_{\mathbf{z}'}P_{\mathbf{z}}(\mathbf{z}',t') - 2\mathbf{j}_{\mathrm{s}}(\mathbf{z})\cdot\mathbf{j}_{\mathrm{s}}(\mathbf{z}')]. \end{aligned} \tag{3.5.49}$$

The second line is interpreted analogously to the current-density correlation in Eq. (3.5.45) with the only difference that the scalar product of current operators reflects scalar products of increments along individual trajectories. The first term, however, does not appear in $C^{\mathbf{xy}}_{\mathbf{J}\rho}$ and $C^{\mathbf{xy}}_{\rho\rho}$. As can be seen from the derivation in Eq. (3.5.48) this term originates from the purely diffusive (i.e., Brownian) term involving $\mathrm{d}\mathbf{x}_\tau \cdot \mathrm{d}\mathbf{x}_\tau = 2\mathrm{Tr}\mathbf{D}(\mathbf{z})\mathrm{d}\tau$ and only appears for $t_1 = t_2$, i.e., $t' = 0$. Thus, this term *cannot* be interpreted in terms of trajectories from $\mathbf{z}$ to $\mathbf{z}'$ or vice versa, but instead reflects that due to the nature of Brownian motion the square of instantaneous fluctuations $(\mathrm{d}\mathbf{x}_\tau)^2$ does not vanish but contributes on the order $\mathrm{d}\tau$. Note that since here $\mathbf{z} = \mathbf{z}'$ this term only contributes if $U^h_{\mathbf{x}}(\mathbf{z})$ and $U^h_{\mathbf{y}}(\mathbf{z}')$ have non-zero overlap.

For $\mathbf{x} = \mathbf{y}$ the covariance becomes the current variance $\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) \equiv C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ which plays a vital role in stochastic thermodynamics. As an application of the result in Eq. (3.5.49) we use the TUR-bound under concurrent variation of the coarse-graining scale h to optimize the inference of a system's dissipation via current fluctuations. Before we turn to this inference problem, we take a closer look at the limit of no coarse graining, i.e., $h \to 0$.

3.5.5 The limit of no coarse graining

In this section we consider the variance $\mathrm{var}^{\mathbf{x}}_{\rho}(t) \equiv C^{\mathbf{xx}}_{\rho\rho}(t), \mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) \equiv C^{\mathbf{xx}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ and correlations $\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t)$ in Eqs. (3.5.23), (3.5.45), and (3.5.49) with $\mathbf{x} = \mathbf{y}$ in the limit of no coarse graining, i.e., when $h \to 0$. In particular, we consider normalized window functions $\int \mathrm{d}\mathbf{z}U^h_{\mathbf{x}}(\mathbf{z}) = 1$ such that in the limit of no coarse graining $U^{h\to 0}_{\mathbf{x}}(\mathbf{z}) = \delta(\mathbf{x}-\mathbf{z})$ [see, e.g.,

Phys. Rev. Res. 4, 033243 (2022)

Eq. (3.5.3)]. Thus, the density and current observables in Eq. (3.5.1) for $h = 0$ correspond to the empirical density and current defined with a δ function,

$$\overline{\rho_{\mathbf{x}}}(t) \equiv \frac{1}{t}\int_0^t \delta(\mathbf{x}-\mathbf{x}_\tau)\mathrm{d}\tau,$$
$$\overline{\mathbf{J}_{\mathbf{x}}}(t) \equiv \frac{1}{t}\int_{\tau=0}^{\tau=t} \delta(\mathbf{x}-\mathbf{x}_\tau)\circ \mathrm{d}\mathbf{x}_\tau, \tag{3.5.50}$$

which is the definition typically adopted in the literature [34, 43–45, 107, 110, 141–144]. We show in the Appendix that in spatial dimensions $d \geq 2$, the variance and correlation functions diverge $\mathrm{var}_\rho^{\mathbf{x}}(t), \mathrm{var}_{\mathbf{J}}^{\mathbf{x}}(t), \mathbf{C}_{\mathbf{J}\rho}^{\mathbf{xx}}(t) \to \infty$ as $h \to 0$. Note that the mean values in Eqs. (3.5.15) and (3.5.17) of the observables in Eq. (3.5.50) do not diverge but instead for $U_{\mathbf{x}}^{h\to 0}(\mathbf{z}) = \delta(\mathbf{x}-\mathbf{z})$ directly simplify to $\langle\overline{\rho_{\mathbf{x}}}(t)\rangle_{\mathrm{s}} = p_{\mathrm{s}}(\mathbf{x})$ and $\langle\overline{\mathbf{J}_{\mathbf{x}}}(t)\rangle_{\mathrm{s}} = \mathbf{j}_{\mathrm{s}}(\mathbf{x})$ (see also Ref. [34]).

Before we go into the specific results for the limit $h \to 0$, let us first discuss why divergent fluctuations of the functionals in Eq. (3.5.50), although overlooked so far, are in fact *not* surprising. The simplest argument is that second moments as, e.g., $\langle\overline{\rho_{\mathbf{x}}}(t)^2\rangle_{\mathrm{s}}$ involve terms $\langle\delta(\mathbf{x}-\mathbf{x}_\tau)\delta(\mathbf{x}-\mathbf{x}_{\tau'})\rangle_{\mathrm{s}}$, which diverge for $\tau = \tau'$ since a squared δ function appears. In contrast, the mean value $\langle\overline{\rho_{\mathbf{x}}}(t)\rangle_{\mathrm{s}}$ contains $\langle\delta(\mathbf{x}-\mathbf{x}_\tau)\rangle_{\mathrm{s}} = p_{\mathrm{s}}(\mathbf{x})$, which is finite. Loosely speaking, the mean value involving $\langle\delta(\mathbf{x}-\mathbf{x}_\tau)\rangle_{\mathrm{s}}$ is given by the probability to be at the point $\mathbf{x}$, which is zero, multiplied by the height of the δ function at $\mathbf{x}$, which is infinite. Since the mean value is finite for $h \to 0$, this can be seen to yield "$0 \times \infty = p_{\mathrm{s}}(\mathbf{x})$," while the second moment contains a squared δ peak, such that the second moment loosely speaking diverges due to "$0 \times \infty^2 = p_{\mathrm{s}}(\mathbf{x}) \times \infty = \infty$". This argument illustrates that divergent fluctuations are not surprising, but this argument is oversimplified since it does not take into account the time integration. In particular, to explain why the divergence only occurs in spatial dimensions $d \geq 2$, we have to note that due to the time integration the one-dimensional case is qualitatively different. Given some point $\mathbf{z}$ in d-dimensional space, the trajectory will hit $z \equiv \mathbf{z}$ with a finite probability in $d = 1$ (i.e., with nonzero probability there is some $\tau \in [0,t]$ such that $x_\tau = z$; e.g., if $x_0 < x_t$ all points in $[x_0, x_t]$ are hit). This is qualitatively different for $d \geq 2$, since overdamped motion in $d \geq 2$ does not hit points, i.e., the probability to hit a given point $\mathbf{z}$ is zero, $\mathbb{P}(\exists\tau \in (0,\infty)\colon \mathbf{x}_\tau = \mathbf{z}) = 0$ [178]. This property is not specific to overdamped motion, but is rather due to the fact that the set of points $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$ has Lebesgue measure zero for $d \geq 2$.

To further explain the divergence and its dependence on the dimensionality in a somewhat less oversimplified way (for the detailed derivation see the Appendix), we take a second look at the term $\langle\delta(\mathbf{x}-\mathbf{x}_\tau)\delta(\mathbf{x}-\mathbf{x}_{\tau'})\rangle_{\mathrm{s}} = G(\mathbf{x}, |\tau-\tau'|\,|\mathbf{x})p_{\mathrm{s}}(\mathbf{x})$ occurring in $\langle\overline{\rho_{\mathbf{x}}}(t)^2\rangle_{\mathrm{s}}$.

Phys. Rev. Res. 4, 033243 (2022)

Here, $G(\mathbf{x},t'|\mathbf{x})$ trivially diverges if $t'=0$. However, the relevant question is whether the return integral $\int_0^t G(\mathbf{x},t'|\mathbf{x})\mathrm{d}t'$ diverges. Any divergence in the integral would come from $t'\to 0$ where $G(\mathbf{x},t'|\mathbf{x})$ diverges, i.e., from the limit of small time differences $|\tau-\tau'|$. For $t'\to 0$ the overdamped propagator $G(\mathbf{x},t'|\mathbf{x})$ becomes Gaussian with variance $\propto Dt'$ [17] [so for very small t' we have $G(\mathbf{x},t'|\mathbf{x})\propto t'^{-d/2}$ in d-dimensional space], and thus the return integral $\int_0^t G(\mathbf{x},t'|\mathbf{x})\mathrm{d}t'$ diverges if and only if $\int_0^t t'^{-d/2}\mathrm{d}t'$ diverges. Therefore, the variance $\mathrm{var}_\rho^{\mathbf{x}}(t)$ diverges in spatial dimensions $d\geq 2$.

Apart from the two arguments above providing mathematical intuition about the divergence, there is also a physical intuition that suggests divergent fluctuations. Recall that for finite $h>0$, the observables $\overline{\rho_{\mathbf{x}}^U}$ and $\overline{\mathbf{J}_{\mathbf{x}}^U}$ in Eq. (3.5.1) by definition measure the time and displacement that the trajectory $(x_\tau)_{0\leq\tau\leq t}$ accumulates in the region $U_{\mathbf{x}}^h$ of scale h around $\mathbf{x}$. Now, as $h\to 0$, only visitations of precisely the point $\mathbf{x}$ contribute. Two very similar (but not equal) trajectories may now give very different values for $\overline{\rho_{\mathbf{x}}^U}$ and $\overline{\mathbf{J}_{\mathbf{x}}^U}$, depending on whether the point $\mathbf{x}$ is hit or even slightly missed (e.g., by a distance h). Therefore, fluctuations among different trajectories of these functionals diverge as $h\to 0$. This reasoning is not restricted to overdamped stochastic motion, and indeed seems to hold for more general dynamics (see outlook in Sec. 3.5.10).

This simple illustration also explains why fluctuations do not diverge in one-dimensional space. There, points are hit, meaning that, e.g., a trajectory starting at 0 and ending at 1 always hits all points in between at some intermediate time, which is why the density and current observables have qualitatively lower fluctuations compared to higher dimensions. The reason that the divergence for $d\geq 2$ was overlooked so far is probably due to the fact that most explicit examples were analyzed in one-dimensional space only.

Explicitly, in the limit $h\to 0$ the expressions Eqs. (3.5.23), (3.5.45), and (3.5.49) with $\mathbf{x}=\mathbf{y}$ for any time t take the form

$$\begin{aligned}
\mathrm{var}_\rho^{\mathbf{x}}(t) &\overset{h\to 0}{\simeq} \frac{K}{\tilde{D}_{\mathbf{x}}t}p_{\mathrm{s}}(\mathbf{x})\begin{cases}\frac{h^{2-d}}{d-2} & \text{for } d>2\\ -\ln h & \text{for } d=2\end{cases}\\
\mathbf{C}_{\mathbf{J}\rho}^{\mathbf{xx}}(t) &\overset{h\to 0}{\simeq} \mathbf{j}_{\mathrm{s}}(\mathbf{x})\mathrm{var}_\rho^{\mathbf{x}}(t)/2p_{\mathrm{s}}(\mathbf{x})\\
\mathrm{var}_{\mathbf{J}}^{\mathbf{x}}(t) &\overset{h\to 0}{=} K'\frac{2\tilde{D}'_{\mathbf{x}}}{t}p_{\mathrm{s}}(\mathbf{x})(d-1)h^{-d}+\mathcal{O}(t^{-1})\mathcal{O}(h^{1-d}),
\end{aligned}\tag{3.5.51}$$

where $\simeq$ denotes asymptotic equality, $\tilde{D}_{\mathbf{x}}$, and $\tilde{D}'_{\mathbf{x}}$ are constants bounded by the smallest and largest eigenvalues of $\mathbf{D}(\mathbf{x})$, and K,K' are constants depending on the specific normalized window $U_{\mathbf{z}}^h$ (see the Appendix). Note that the dominant term in $\mathrm{var}_{\mathbf{J}}^{\mathbf{x}}(t)$ vanishes

Phys. Rev. Res. 4, 033243 (2022)

for $d = 1$, such that all three expressions only diverge for $d \geq 2$. Some details on the case $d = 1$ are shown in the Appendix in Sec. 3.5.12.

Thus, the empirical density and current as defined in Eq. (3.5.50) have divergent fluctuations. Note that an infinite variance contradicts Gaussian statistics on all time scales. This divergence, moreover, leads us to question whether Eq. (3.5.50) is even well-defined, i.e., whether these observables are mathematically well-defined random variables, and whether the result in the limit $h \to 0$ is unaffected by the specific choice of the $U^h_{\mathbf{x}}$ as long as $U^{h\to 0}_{\mathbf{x}}(\mathbf{z}) = \delta(\mathbf{x} - \mathbf{z})$.

3.5.6 Application to inference of dissipation

We now apply the results for the current variance $\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) \equiv C^{\mathbf{xx}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ in Eq. (3.5.49) for $\mathbf{x} = \mathbf{y}$. For an individual component, e.g., $J_y \equiv [\overline{\mathbf{J}^U_{\mathbf{x}}}]_y$, of the vector $\overline{\mathbf{J}^U_{\mathbf{x}}}$ the equivalent result reads

$$\begin{aligned}\mathrm{var}^{\mathbf{x}}_{J_y}(t) =& \frac{2}{t}\int \mathrm{d}\mathbf{z}[\mathbf{D}(\mathbf{z})]_{yy}U^h_{\mathbf{x}}(\mathbf{z})U^h_{\mathbf{x}}(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}) \\ &+ \hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[(\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}'})_y(\hat{\mathbf{j}}_{\mathbf{z}})_y P_{\mathbf{z}'}(\mathbf{z},t') + (\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}})_y(\hat{\mathbf{j}}_{\mathbf{z}'})_y P_{\mathbf{z}}(\mathbf{z}',t') - 2[\mathbf{j}_{\mathrm{s}}(\mathbf{z})]_y[\mathbf{j}_{\mathrm{s}}(\mathbf{z}')]_y]. \quad (3.5.52)\end{aligned}$$

With the dissipation rate $\dot{\Sigma}$ in Eq. (3.5.10), current observables such as $J_y \equiv [\overline{\mathbf{J}^U_{\mathbf{x}}}]_y$ satisfy the TUR [47, 126] (in the form relevant below first proven in Ref. [126]),

$$\frac{\mathrm{var}^{\mathbf{x}}_{J_y}(t)}{\langle J_y\rangle^2_{\mathrm{s}}} \geq \frac{2}{t\dot{\Sigma}}. \quad (3.5.53)$$

This bound is of particular interest since it allows to infer a lower bound on a system's dissipation from measurements of the local mean current and current fluctuations [67, 68, 153, 154, 201]. Note that Eq. (3.5.53) implicitly assumes "perfect" statistics; i.e., $\langle J_y\rangle_{\mathrm{s}}$ and $\mathrm{var}^{\mathbf{x}}_{J_y}(t)$ are the exact mean and variance for the process under consideration (not limited by sampling constraints on a finite number of realizations).

We now investigate the influence of the coarse graining on the sharpness of the bound (3.5.53). One might naively expect that coarse graining annihilates information. However, as shown in Sec. 3.4 the current fluctuations diverge in spatial dimensions $d \geq 2$ in the limit $h \to 0$ (of no coarse graining), whereas the mean converges to a constant (note that $\dot{\Sigma}$ does not at all depend on $U^h_{\mathbf{x}}$). The exact asymptotics for $h \to 0$ in Sec. 3.4 demonstrate that the bound (3.5.53) becomes entirely independent of the process (i.e., it only depends on p_{s} but contains no information about the non-equilibrium part of the dynamics). Therefore,

Phys. Rev. Res. 4, 033243 (2022)

the left-hand side of inequality (3.5.53) tends to ∞ as $h \to 0$, rendering the TUR without spatial coarse graining unable to infer dissipation beyond the statement $\dot{\Sigma} \geq 0$ for $h = 0$.

However, the naive intuition is correct in the limit of "ignorant" coarse graining $h \to \infty$, where $U^h_{\mathbf{x}}$ becomes asymptotically constant in a sufficiently large hypervolume centered at $\mathbf{x}$ [i.e., in a hypervolume A where $\int_{\mathrm{A}} p_{\mathrm{s}}(\mathbf{x})\mathrm{d}\mathbf{x} \approx 1$]. The integration over a constant $U^h_{\mathbf{x}} = c$ yields $\left\langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t) \right\rangle_{\mathrm{s}} = c \int \mathrm{d}\mathbf{z}\mathbf{j}_{\mathrm{s}}(\mathbf{z}) = \mathbf{0}$ for the mean in Eq. (3.5.17). The vanishing $\left\langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t) \right\rangle_{\mathrm{s}}$ may be seen in two ways. First, since $\nabla_{\mathbf{z}} \cdot \mathbf{j}_{\mathrm{s}}(\mathbf{z}) = 0$, curl $\mathbf{j}_{\mathrm{s}}(\mathbf{z}) = \nabla_{\mathbf{z}} \times \mathbf{f}(\mathbf{z})$ and by the Stokes theorem $\int_{\mathrm{A}} \mathrm{d}^2 z \nabla_{\mathbf{z}} \times \mathbf{f}(\mathbf{z}) = \int_{\partial \mathrm{A}} \mathbf{f} \cdot \mathrm{d}\mathbf{l}$ which vanishes since at the boundary ∂A at ∞ we have $p_{\mathrm{s}} \to 0$, thus $\mathbf{j}_{\mathrm{s}} \to 0$ and therefore the vector potential $\mathbf{f} \to \mathbf{0}$. Second, for $U^h_{\mathbf{x}} = c$ we have $\overline{\mathbf{J}^U_{\mathbf{x}}}(t) = \frac{c}{t}(\mathbf{x}_t - \mathbf{x}_0)$ [and we assume $\mathbf{x}_0$ to be sampled from $p_{\mathrm{s}}(\mathbf{x})$]. Then $\mathbf{x}_0$ and $\mathbf{x}_t$ are both distributed according to p_{s}, and thus $\langle \mathbf{x}_t \rangle_{\mathrm{s}} = \langle \mathbf{x}_0 \rangle_{\mathrm{s}}$ and $t \left\langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t) \right\rangle_{\mathrm{s}} / c = \langle \mathbf{x}_t \rangle_{\mathrm{s}} - \langle \mathbf{x}_0 \rangle_{\mathrm{s}} = 0$. Conversely, the variance remains strictly positive. Therefore, also for $h \to \infty$ the left-hand side of inequality (3.5.53) diverges, rendering the TUR with an "ignorant" coarse graining incapable of inferring dissipation (again only giving $\dot{\Sigma} \geq 0$ as for $h = 0$).

These two arguments, i.e., the necessity of coarse graining, see Sec. 3.4, and the failure of an "ignorant" coarse graining, imply that an intermediate coarse graining exists that is optimal for inferring dissipation via the TUR [Eq. (3.5.53)].

We first demonstrate this finding using a two-dimensional rotational flow (3.5.2) with Gaussian coarse graining window in Eq. (3.5.3). We evaluate the left-hand side of Eq. (3.5.53) for varying h and $\mathbf{x}$ and compare it to the constant right-hand side of Eq. (3.5.53). Particularly for $\mathbf{D}(\mathbf{z}) = D\mathbf{1}$, we have $p_{\mathrm{s}}(\mathbf{z}) = r/(2\pi D)\exp(-r\mathbf{z}^2/(2D))$ and $\mathbf{j}_{\mathrm{s}}(\mathbf{z}) = \Omega p_{\mathrm{s}}(\mathbf{z})(z_2, -z_1)^T$ and the dissipation rate in Eq. (3.5.10) is given by

$$\begin{aligned}\dot{\Sigma} &= \int \mathrm{d}\mathbf{z} \frac{\mathbf{j}_{\mathrm{s}}^T(\mathbf{z})}{p_{\mathrm{s}}(\mathbf{z})} \mathbf{D}^{-1}(\mathbf{z}) \frac{\mathbf{j}_{\mathrm{s}}(\mathbf{z})}{p_{\mathrm{s}}(\mathbf{z})} p_{\mathrm{s}}(\mathbf{z}) = \frac{\Omega^2}{D} \int \mathrm{d}\mathbf{z}\, \mathbf{z}^2 p_{\mathrm{s}}(\mathbf{z}) \\ &= \frac{\Omega^2}{D} \left\langle \mathbf{x}_0^2 \right\rangle_{\mathrm{s}} = \frac{\Omega^2}{D} \left\langle x_1^2 + x_2^2 \right\rangle_{\mathrm{s}} = \frac{\Omega^2}{D} 2 \frac{D}{r} = \frac{2\Omega^2}{r}. \end{aligned} \tag{3.5.54}$$

Thus the TUR in Eq. (3.5.53) for the rotational flow becomes

$$\frac{\mathrm{var}^{\mathbf{x}}_{J_y}(t)}{\langle J_y \rangle_{\mathrm{s}}^2} \geq \frac{r}{t\Omega^2}. \tag{3.5.55}$$

The results shown in Fig. 3.5.6a-d demonstrate, as argued above, that relative fluctuations diverge as $h \to 0, \infty$. For this example, the relative error as a function of h has a unique minimum (slightly depending on $\mathbf{x}$, and possibly on other parameters such as t). This means that (restricted to $U^h_{\mathbf{x}}$ being a Gaussian around $\mathbf{x}$) there is a coarse graining scale h that is optimal for inferring a lower bound on the dissipation, that may also provide some intu-

Phys. Rev. Res. 4, 033243 (2022)

ition about the formal optimization carried out in Ref. [201]. This result demonstrates that coarse graining trajectory data *a posteriori* can improve the inference of thermodynamical information, which is a strong motivation for considering coarse graining.

In particular, note that this method is readily applicable; i.e., one does not need to know the underlying process (as long as the dynamics is overdamped). As was done in Fig. 3.5.6e-h one simply integrates the trajectories to obtain the coarse grained current as defined in Eq. (3.5.1). Then, the mean and variance are readily obtained from the fluctuations along an ensemble of individual trajectories, and for each value of $\mathbf{x}$ and h one determines a lower bound on the dissipation via Eq. (3.5.53). Finally, one takes the best of those bounds. We here only consider Gaussian $U^h_{\mathbf{x}}$ for the coarse graining, but due to the flexibility of the theory one could even choose window functions that do not have to relate to the notion of coarse graining. Notably, a Gaussian window function is in this case better than, e.g., a rectangular indicator function (which one usually uses for binning data) due to an improved smoothing effect. Moreover, one further expects a reduced error due to discrete-time effects.

Note that compared to many of the similar existing methods [68, 132, 134], we *neither* advise to rasterize the continuous dynamics to parametrize (i.e., "count") currents *nor* advise to approximate the dynamics by a Markov-jump process. Our method is therefore not only correct (note that a Markov-jump assumption is only accurate in the presence of a timescale separation ensuring a local equilibration, e.g., as a result of high barriers separating energy minima) but also has the great advantage of not having to parametrize rates at all. Instead, one simply integrates trajectories according to Eq. (3.5.1).

A generalization to windows that are not centered at individual points, as well as the use of correlations in Eq. (3.5.45) entering the recent so-called correlation TUR (CTUR) inequality [134] will be considered in forthcoming publications.

To underscore the applicability of the above inference strategy, we apply it to a more complicated system, for which a Markov jump process description would be difficult due to the presence of low and flat barriers and extended states. The results are shown in Fig. 3.5.6e-h. The example is constructed by considering the two-dimensional potential

$$\phi(x,y) = 0.75(x^2-1)^2 + (y^2-1.5)^2((x+0.5y-0.5)^2+0.5) + c, \tag{3.5.56}$$

where c is a constant such that $p_{\mathrm{s}}(\mathbf{z}) = \exp[-\phi(\mathbf{z})]$ is normalized. We consider isotropic additive noise $\mathbf{D}(\mathbf{z}) = D\mathbf{1}$ and construct the Itô/Langevin equation for the process as

$$\mathrm{d}\mathbf{x}_\tau = -D\{\nabla\phi\}(\mathbf{x}_\tau)\mathrm{d}\tau + \mathbf{F}^{\mathrm{irrev}}(\mathbf{x}_\tau) + \sqrt{2D}\mathrm{d}\mathbf{W}_\tau, \tag{3.5.57}$$

Phys. Rev. Res. 4, 033243 (2022)

where

$$\mathbf{F}^{\text{irrev}}(\mathbf{z}) = \frac{\mathbf{j}_{\text{s}}(\mathbf{z})}{p_{\text{s}}(\mathbf{z})} \equiv -D\Omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \{\nabla\phi\}(\mathbf{z}) \tag{3.5.58}$$

is an irreversible drift that is by construction orthogonal to $\nabla\phi$ and thus does not alter the steady state [i.e., same $p_{\text{s}} = \exp[-\phi]$ for equilibrium ($\Omega = 0$) or any other Ω]. With Eq. (3.5.58) the dissipation in Eq. (3.5.10) for this process reads

$$\begin{aligned} \dot{\Sigma} =& D\Omega^2 \int \mathrm{d}^2\mathbf{x} \{\nabla\phi\}(\mathbf{x})^T \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^T \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ &\cdot \{\nabla\phi\}(\mathbf{x}) p_s(\mathbf{x}) = D\Omega^2 \int \mathrm{d}^2\mathbf{x} \{\nabla\phi\}^2(\mathbf{x}) \exp\left[-\phi(\mathbf{x})\right], \end{aligned} \tag{3.5.59}$$

which is solved numerically and gives $\dot{\Sigma} = 19.65 D\Omega^2$. We see in Fig. 3.5.6h that some intermediate coarse graining h is still optimal, but the optimal scale h now depends more intricately on $\mathbf{x}$ and the curves are not convex in h anymore.

Overall, we see that the approach is robust and easily applicable, and does not require determining and parametrize any rates. Moreover, due to the implications of the theory to the limits $h \to 0, \infty$, we can assert that some intermediate coarse graining will generally be optimal.

3.5.7 Simplifications and symmetries

In this section we list the symmetries obeyed by the results in Eqs. (3.5.15), (3.5.17), (3.5.23), (3.5.45), and (3.5.49) [with integral operator (3.5.22)]. Note that the limit $h \to 0$ was carried out in Sec. 3.5.5 and the limit $h \to \infty$ gives $U^h_{\mathbf{x}} = c$ as noted before which greatly simplifies the further analysis. The limits $t \to 0$ and $t \to \infty$ will be addressed in Sec. 3.5.9.

First consider dynamics obeying detailed balance, i.e., $\mathbf{j}_{\text{s}} = \mathbf{0}$. We then have $\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}} = -\hat{\mathbf{j}}_{\mathbf{z}} = -\hat{\mathbf{j}}^{g}_{\mathbf{z}}$ and the dual-reversal symmetry in Eq. (3.5.11), simplifies to the detailed balance statement $G(\mathbf{y}, t|\mathbf{x}) p_{\text{s}}(\mathbf{x}) = G(\mathbf{x}, t|\mathbf{y}) p_{\text{s}}(\mathbf{y})$ or $P_{\mathbf{z}}(\mathbf{z}', t) = P_{\mathbf{z}'}(\mathbf{z}, t)$. From this we obtain the following simplifications for $\mathbf{j}_{\text{s}} = \mathbf{0}$:

$$\begin{aligned} &\left\langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t) \right\rangle_{\text{s}} = \mathbf{0}, \quad \mathbf{C}^{\mathbf{xy}}_{\mathbf{J}_\rho}(t) = \mathbf{0} \\ &C^{\mathbf{xy}}_{\rho\rho}(t) = 2\hat{\mathcal{I}}^{t,U}_{\mathbf{xy}} \left[P_{\mathbf{z}}(\mathbf{z}', t') - p_{\text{s}}(\mathbf{z}) p_{\text{s}}(\mathbf{z}') \right], \\ &C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) = \frac{2}{t} \int d\mathbf{z} \mathrm{Tr}[\mathbf{D}(\mathbf{z})] U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{y}}(\mathbf{z}) p_{\text{s}}(\mathbf{z}) - 2\hat{\mathcal{I}}^{t,U}_{\mathbf{xy}} [\hat{\mathbf{j}}^g_{\mathbf{z}} \cdot \hat{\mathbf{j}}^g_{\mathbf{z}'} P_{\mathbf{z}}(\mathbf{z}', t') + \mathbf{j}_{\text{s}}(\mathbf{z}) \cdot \mathbf{j}_{\text{s}}(\mathbf{z}')]. \end{aligned} \tag{3.5.60}$$

Phys. Rev. Res. 4, 033243 (2022)

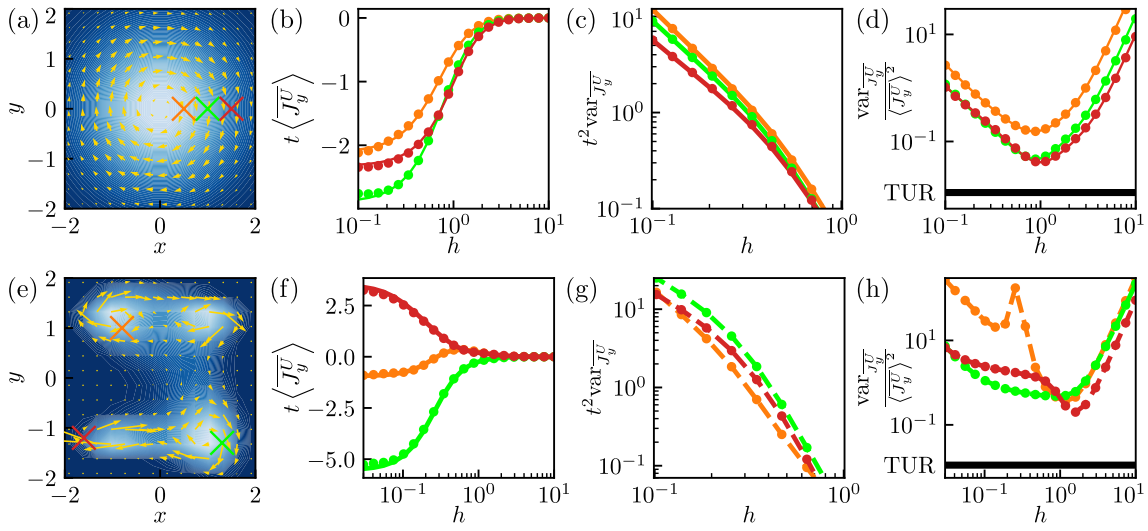

Figure 3.5.6: (a) Steady-state density (blue color gradient) and current (yellow arrows) for the rotational flow in Eq. (3.5.2) with $\Omega = 3$. Points around which the currents are evaluated in (b)-(d) are denoted by colored crosses. (b) Simulated values (circles) of the mean y component of the time-integrated current from 2000 trajectories of length $t = 10$ with time step $dt = 0.001$ (using the stochastic Euler algorithm) starting from steady-state initial conditions using a Gaussian window function [Eq. (3.5.3)] with different coarse-graining scales h. Analytical results in Eq. (3.5.17) are shown with lines. (c) As in (b) but for variances. Simulations (circles) are shown alongside analytical results [lines; the results are analytic up to one time integration; see Eq. (3.5.52)] for the variance of currents. (d) The relative error (ratio of variance and mean squared) as a function of h features a minimum at an intermediate h. At this minimum, the current fluctuations give the best lower bound on the dissipation via the TUR [Eq. (3.5.53)] at the value $2/(t\dot{\Sigma}) = 2/(10 \times 18) = 0.011$ (black line). (e) As in (a) but for the more complicated process in Eq. (3.5.57) with $D = 1$ and we choose $\Omega = 0.957$ to have the same dissipation as in (d); here the dissipation is obtained by means of a numerical integration. (f) As in (b) but for the process in (e) the "analytical" mean [Eq. (3.5.17)] had to be evaluated by means of a numerical integration. (g) As in (c) but simulated values are shown by circles and dashed lines (but without a comparison to results of numerical integration since these require the knowledge of the propagator). (h) As in (d) but for the process in (e). The relative error may display several local minima. Some intermediate h still allows for an optimal inference of the dissipation via the TUR (black line). Note that the relative error diverges (orange line) where the mean crosses zero [orange line in (b)].

For the remainder of this section we consider $\mathbf{j}_s \neq \mathbf{0}$. Note that by definition the interchange $\mathbf{x} \leftrightarrow \mathbf{y}$ leaves $C^{\mathbf{xy}}_{\rho\rho}(t)$ and $C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t)$ invariant, but *not* $C^{\mathbf{xy}}_{\mathbf{J}\rho}(t)$ since it considers currents at $\mathbf{x}$ and densities at $\mathbf{y}$.

Phys. Rev. Res. 4, 033243 (2022)

For single-point correlations and variances $\mathbf{x} = \mathbf{y}$ (more precisely $U^h_{\mathbf{x}} = U^h_{\mathbf{y}}$) the integrations over $\mathbf{z}$ and $\mathbf{z}'$ are equivalent and thus the results simplify to

$$\begin{aligned}
C^{\mathbf{xx}}_{\rho\rho}(t) &= 2\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}\left[P_{\mathbf{z}}(\mathbf{z}',t') - p_{\mathrm{s}}(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}')\right]\\
\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t) &= \hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[(\hat{\mathbf{j}}_{\mathbf{z}'} + \hat{\mathbf{j}}^{\dagger}_{\mathbf{z}})P_{\mathbf{z}}(\mathbf{z}',t') - 2\mathbf{j}_{\mathrm{s}}(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}')]\\
C^{\mathbf{xx}}_{\mathbf{J}\cdot\mathbf{J}}(t) &= \frac{2}{t}\int d\mathbf{z}\mathrm{Tr}[\mathbf{D}(\mathbf{z})][U^h_{\mathbf{x}}]^2(\mathbf{z})p_{\mathrm{s}}(\mathbf{z}) + 2\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\hat{\mathbf{j}}^{\dagger}_{\mathbf{z}}\cdot\hat{\mathbf{j}}_{\mathbf{z}'}P_{\mathbf{z}}(\mathbf{z}',t') - \mathbf{j}_{\mathrm{s}}(\mathbf{z})\cdot\mathbf{j}_{\mathrm{s}}(\mathbf{z}')].
\end{aligned} \tag{3.5.61}$$

Now we again allow $\mathbf{x} \neq \mathbf{y}$ and consider the process and the $\mathbf{j}_{\mathrm{s}} \leftrightarrow -\mathbf{j}_{\mathrm{s}}$ inverted process. Then, from Eq. (3.5.11) and $[\hat{\mathbf{j}}^{\dagger}_{\mathbf{z}'}\cdot\hat{\mathbf{j}}_{\mathbf{z}}]^{-\mathbf{j}_{\mathrm{s}}} = -\hat{\mathbf{j}}_{\mathbf{z}'}\cdot[-\hat{\mathbf{j}}^{\dagger}_{\mathbf{z}}] = \hat{\mathbf{j}}_{\mathbf{z}'}\cdot\hat{\mathbf{j}}^{\dagger}_{\mathbf{z}} = \hat{\mathbf{j}}^{\dagger}_{\mathbf{z}}\cdot\hat{\mathbf{j}}_{\mathbf{z}'}$, we get $[\hat{\mathbf{j}}^{\dagger}_{\mathbf{z}'}\cdot\hat{\mathbf{j}}_{\mathbf{z}}P_{\mathbf{z}'}(\mathbf{z},t')]^{-\mathbf{j}_{\mathrm{s}}} = \hat{\mathbf{j}}^{\dagger}_{\mathbf{z}}\cdot\hat{\mathbf{j}}_{\mathbf{z}'}P_{\mathbf{z}}(\mathbf{z}',t')$ and thus obtain

$$\begin{aligned}
\left\langle\overline{\rho^U_{\mathbf{x}}}(t)\right\rangle_{\mathrm{s}} &= \left\langle\overline{\rho^U_{\mathbf{x}}}(t)\right\rangle_{\mathrm{s}}^{-\mathbf{j}_{\mathrm{s}}}\\
\left\langle\overline{\mathbf{J}^U_{\mathbf{x}}}(t)\right\rangle_{\mathrm{s}} &= -\left\langle\overline{\mathbf{J}^U_{\mathbf{x}}}(t)\right\rangle_{\mathrm{s}}^{-\mathbf{j}_{\mathrm{s}}}\\
C^{\mathbf{xy}}_{\rho\rho}(t) &= [C^{\mathbf{xy}}_{\rho\rho}(t)]^{-\mathbf{j}_{\mathrm{s}}}\\
\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t) &= -[\mathbf{C}^{\mathbf{xy}}_{\mathbf{J}\rho}(t)]^{-\mathbf{j}_{\mathrm{s}}}\\
C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t) &= [C^{\mathbf{xy}}_{\mathbf{J}\cdot\mathbf{J}}(t)]^{-\mathbf{j}_{\mathrm{s}}}.
\end{aligned} \tag{3.5.62}$$

In addition to the symmetries of the first and second cumulants, a stronger path-wise version of the dual-reversal symmetry in Eq. (3.5.11) (or time-reversal symmetry at equilibrium) dictates symmetries of the full distributions of the functionals of steady-state trajectories under the reversal $\mathbf{j}_{\mathrm{s}} \leftrightarrow -\mathbf{j}_{\mathrm{s}}$. Notably, at equilibrium ($\mathbf{j}_{\mathrm{s}} = \mathbf{0}$) these simplify to symmetries of the process [which is a much stronger result since we do not have to compare to another (artificial) process with an inverted $\mathbf{j}_{\mathrm{s}}$].

To motivate this stronger symmetry, note that for steady-state initial conditions for any finite set of times $t_1 < t_2 < \cdots < t_n$ we have that the joint density $P_n(\cdots)$ for positions z_i at equally spaced times $t_i = i \times \Delta t$ for $i = 0, 1, \ldots, n$ is given by [since we have a Markov process by definition, i.e., Eq. (3.5.4) has no memory]

$$P_n(z_0,t_0;z_1,t_1;\ldots;z_n,t_n) = p_{\mathrm{s}}(\mathbf{x}_0)G(\mathbf{x}_1,\Delta t|\mathbf{x}_0)\cdots G(\mathbf{x}_n,\Delta t|\mathbf{x}_{n-1}). \tag{3.5.63}$$

By applying the dual-reversal symmetry in Eq. (3.5.11) $n-1$ times, we obtain

$$\begin{aligned}
P_n(z_0,t_0;z_1,t_1;\ldots;z_n,t_n) &= G^{-\mathbf{j}_{\mathrm{s}}}(\mathbf{x}_0,\Delta t|\mathbf{x}_1)\cdots G^{-\mathbf{j}_{\mathrm{s}}}(\mathbf{x}_{n-1},\Delta t|\mathbf{x}_n)p_{\mathrm{s}}(\mathbf{x}_n)\\
&= P_n^{-\mathbf{j}_{\mathrm{s}}}(z_n,0;z_{n-1},\Delta t;\ldots;z_0,n\Delta t)\\
&= P_n^{-\mathbf{j}_{\mathrm{s}}}(z_n,t_0;z_{n-1},t_1;\ldots;z_0,t_n).
\end{aligned} \tag{3.5.64}$$

Phys. Rev. Res. 4, 033243 (2022)

The $n+1$ points $(z_1,\dots,z_n)$ represent a discrete-time path for which Eq. (3.5.64) implies the path-wise discrete-time dual-reversal symmetry (denote $t=t_n=n\Delta t$)

$$P_n(z_0,t_0;z_1,t_1;\dots;z_n,t_n)=P_n^{-\mathbf{j}_\mathrm{s}}(z_n,t-t_n;z_{n-1},t-t_{n-1};\dots;z_0,t-t_0), \tag{3.5.65}$$

i.e., the probability of forward paths $(\mathbf{x}_{t_i})_{i=0,1,\dots,n}$ agrees with the probability of backwards paths of the process with inverted steady-state current $\mathbf{j}_\mathrm{s}\to-\mathbf{j}_\mathrm{s}$, i.e.,

$$\mathbb{P}\left[(\mathbf{x}_{t_i})_{i=0,1,\dots,n}\right]=\mathbb{P}^{-\mathbf{j}_\mathrm{s}}\left[(\mathbf{x}_{t-t_i})_{i=0,1,\dots,n}\right]. \tag{3.5.66}$$

Note that at equilibrium, $\mathbf{j}_\mathrm{s}=\mathbf{0}$, this is nothing but the detailed balance for discrete-time paths.

Assuming that one can take a continuum limit $\Delta t\to 0$ (and that a resulting path measure exits) one could conclude that continuous time paths fulfill the symmetry (see also Ref. [133])

$$\mathbb{P}\left[(\mathbf{x}_\tau)_{0\le\tau\le t}\right]=\mathbb{P}^{-\mathbf{j}_\mathrm{s}}\left[(\mathbf{x}_{t-\tau})_{0\le\tau\le t}\right]. \tag{3.5.67}$$

Based on this strong symmetry, and noting that densities are symmetric while currents are antisymmetric under time reversal, i.e.,

$$\begin{aligned}\overline{\rho^U_{\mathbf{x}}}\left[(x_\tau)_{0\le\tau\le t}\right]&=\overline{\rho^U_{\mathbf{x}}}\left[(x_{t-\tau})_{0\le\tau\le t}\right]\\ \overline{\mathbf{J}^U_{\mathbf{x}}}\left[(x_\tau)_{0\le\tau\le t}\right]&=-\overline{\mathbf{J}^U_{\mathbf{x}}}\left[(x_{t-\tau})_{0\le\tau\le t}\right],\end{aligned} \tag{3.5.68}$$

we obtain the following symmetries

$$\begin{aligned}\mathbb{P}\left[\overline{\rho^U_{\mathbf{x}}}(t)=u\right]&=\mathbb{P}^{-\mathbf{j}_\mathrm{s}}\left[\overline{\rho^U_{\mathbf{x}}}(t)=u\right]\\ \mathbb{P}\left[\overline{\mathbf{J}^U_{\mathbf{x}}}(t)=\mathbf{u}\right]&=\mathbb{P}^{-\mathbf{j}_\mathrm{s}}\left[\overline{\mathbf{J}^U_{\mathbf{x}}}(t)=-\mathbf{u}\right].\end{aligned} \tag{3.5.69}$$

Equation (3.5.69) implies symmetries for mean values and variances ($\mathbf{x}=\mathbf{y}$) listed in Eq. (3.5.62) since it implies that all moments of $\overline{\rho^U_{\mathbf{x}}}(t)$ agree and that the n-th moment of a current component i fulfills $\left\langle[\overline{\mathbf{J}^U_{\mathbf{x}}}(t)]_i^n\right\rangle_\mathrm{s}=\left\langle[-\overline{\mathbf{J}^U_{\mathbf{x}}}(t)]_i^n\right\rangle_\mathrm{s}^{-\mathbf{j}_\mathrm{s}}=(-1)^n\left\langle[\overline{\mathbf{J}^U_{\mathbf{x}}}(t)]_i^n\right\rangle_\mathrm{s}^{-\mathbf{j}_\mathrm{s}}$.

Note that Eq. (3.5.69) implies that the statistics of $\rho(t)$ (including all moments) in general depends on $\mathbf{j}_\mathrm{s}$ but is invariant under the inversion $\mathbf{j}_\mathrm{s}\leftrightarrow-\mathbf{j}_\mathrm{s}$. Moreover, current fluctuations

Phys. Rev. Res. 4, 033243 (2022)

at equilibrium ($\mathbf{j}_\mathrm{s} = \mathbf{0}$, hence $\mathbb{P}_\mathrm{EQ} \equiv \mathbb{P} = \mathbb{P}^{-\mathbf{j}_\mathrm{s}}$) are symmetric around the mean $\left\langle \overline{\mathbf{J}^U_\mathbf{x}} \right\rangle_\mathrm{s} = \mathbf{0}$, i.e.,

$$\mathbb{P}_\mathrm{EQ}\left[\overline{\mathbf{J}^U_\mathbf{x}}(t) = \mathbf{u}\right] = \mathbb{P}_\mathrm{EQ}\left[\overline{\mathbf{J}^U_\mathbf{x}}(t) = -\mathbf{u}\right]. \tag{3.5.70}$$

The symmetries for correlations in Eq. (3.5.62), possibly with $\mathbf{x} \neq \mathbf{y}$, may be seen as implications of the more general symmetries

$$\begin{aligned}
\mathbb{P}\left[\overline{\rho^U_\mathbf{x}}(t)\overline{\rho^U_\mathbf{y}}(t) = u\right] &= \mathbb{P}^{-\mathbf{j}_\mathrm{s}}\left[\overline{\rho^U_\mathbf{x}}(t)\overline{\rho^U_\mathbf{y}}(t) = u\right] \\
\mathbb{P}\left[\overline{\mathbf{J}^U_\mathbf{x}}(t)\overline{\rho^U_\mathbf{y}}(t) = \mathbf{u}\right] &= \mathbb{P}^{-\mathbf{j}_\mathrm{s}}\left[\overline{\mathbf{J}^U_\mathbf{x}}(t)\overline{\rho^U_\mathbf{y}}(t) = \mathbf{u}\right] \\
\mathbb{P}\left[\overline{\mathbf{J}^U_\mathbf{x}}(t)\cdot\overline{\mathbf{J}^U_\mathbf{y}}(t) = u\right] &= \mathbb{P}^{-\mathbf{j}_\mathrm{s}}\left[\overline{\mathbf{J}^U_\mathbf{x}}(t)\cdot\overline{\mathbf{J}^U_\mathbf{y}}(t) = u\right].
\end{aligned} \tag{3.5.71}$$

3.5.8 Continuity equation along individual diffusion paths

In this section, we derive a continuity equation for the time-accumulated density $t\overline{\rho^U_\mathbf{x}}(t)$ and current $t\overline{\mathbf{J}^U_\mathbf{x}}(t)$ defined with windows that satisfy $U^h_\mathbf{x}(\mathbf{x}') = U^h_\mathbf{0}(\mathbf{x}' - \mathbf{x})$. This condition in particular holds for all window functions that may be interpreted as a spatial coarse graining, as, e.g., a Gaussian around $\mathbf{x}$ or any indicator function $U^h_\mathbf{x}(\mathbf{x}') \propto \mathbb{1}_{||\mathbf{x}'-\mathbf{x}||\leq h}$ with any norm $||\cdots||$. Under this assumption, $-\nabla_\mathbf{x} U^h_\mathbf{x}(\mathbf{x}') = \nabla_{\mathbf{x}'} U^h_\mathbf{x}(\mathbf{x}') \equiv \left\{\nabla U^h_\mathbf{x}\right\}(\mathbf{x}')$ such that

$$\begin{aligned}
-\nabla_\mathbf{x}\int_{\tau=0}^{\tau=t} U^h_\mathbf{x}(\mathbf{x}_\tau)\circ\mathrm{d}\mathbf{x}_\tau &= \int_{\tau=0}^{\tau=t}\left\{\nabla U^h_\mathbf{x}\right\}(\mathbf{x}_\tau)\circ\mathrm{d}\mathbf{x}_\tau \\
&= U^h_\mathbf{x}(\mathbf{x}_t) - U^h_\mathbf{x}(\mathbf{x}_0) = \partial_t\int_0^t U^h_\mathbf{x}(\mathbf{x}_\tau)\mathrm{d}\tau,
\end{aligned} \tag{3.5.72}$$

which can be written in the form of a continuity equation,

$$\partial_t[t\overline{\rho^U_\mathbf{x}}(t)] = -\nabla_\mathbf{x}\cdot t\,\overline{\mathbf{J}^U_\mathbf{x}}(t). \tag{3.5.73}$$

This generalizes the notion of a continuity equation to individual trajectories $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$ with arbitrary initial and end points. For steady-state dynamics and normalized window functions, i.e., $\int \mathrm{d}^d z U^h_\mathbf{x}(\mathbf{z}) = 1$, taking the mean $\langle\cdot\rangle_\mathrm{s}$ of Eq. (3.5.73) leads to a continuity equation for (coarse-grained) probability densities. Conversely, for non-normalized window functions $\int \mathrm{d}^d z U^h_\mathbf{x}(\mathbf{z}) = \mathrm{Volume}(U^h_\mathbf{x})$, the mean $\langle\cdot\rangle_\mathrm{s}$ of Eq. (3.5.73) may be interpreted as a continuity equation for probabilities.

Note that the statement $\int_{\tau=0}^{\tau=t}\left\{\nabla U^h_\mathbf{x}\right\}(\mathbf{x}_\tau)\circ\mathrm{d}\mathbf{x}_\tau = U^h_\mathbf{x}(\mathbf{x}_t) - U^h_\mathbf{x}(\mathbf{x}_0)$ holds only for the Stratonovich integral but, e.g., not for an Itô integral. Therefore, the continuity equation further motivates the definition in Eq. (3.5.1) via the Stratonovich integral, which was

Phys. Rev. Res. 4, 033243 (2022)

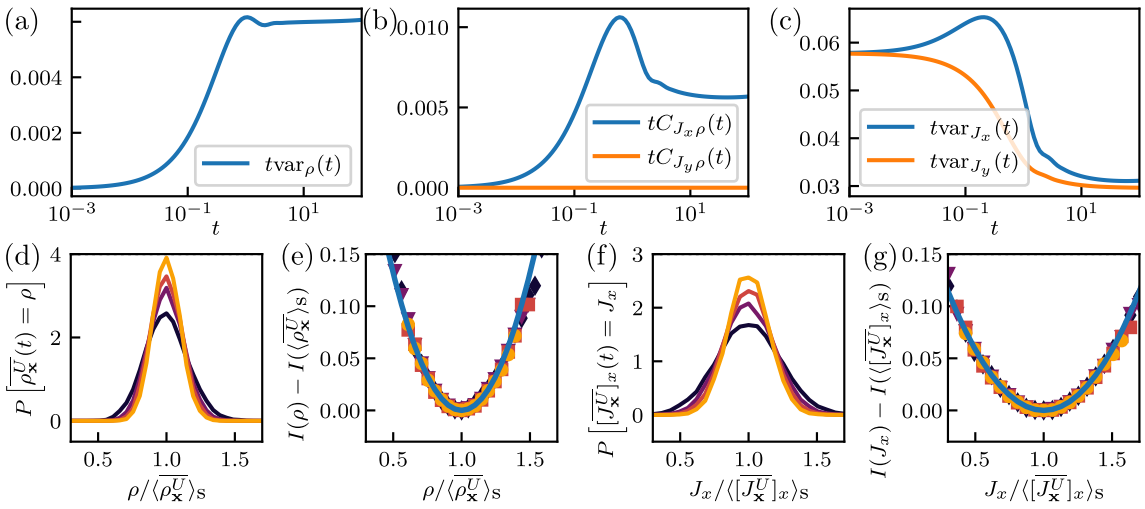

Figure 3.5.7: We consider the rotational flow in Eq. (3.5.2) with $\Omega = 3$ starting from steady-state initial conditions and use a Gaussian coarse-graining window in Eq. (3.5.3) around $\mathbf{x} = (0,1)^T$ with width $h = 0.5$. (a) Analytical result for the variance of the time-averaged density $\overline{\rho^U_{\mathbf{x}}(t)}$ multiplied by time t as a function of t. At long times the variance approaches the large deviation variance in Eq. (3.5.76). (b) As in (a) but for the components of the correlation vector $\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t)$ as in Eqs. (3.5.45) and (3.5.61). (c) As in (a) but for the variances of the current components in Eq. (3.5.52). (d) Simulation of the probability density function of the empirical density $\overline{\rho^U_{\mathbf{x}}}(t)$ assuming the parameters listed above. Colors of lines and symbols throughout denote $t = 40, 60, 80, 100$ from dark to bright. The simulated probability densities were obtained from histograms of 2×10^4 trajectories for each set of parameters. (e) Parabolic approximation for the rate function with variance from Eq. (3.5.76) (line) and simulated rate function $I(\rho) = -\frac{1}{t} \ln P[\overline{\rho^U_{\mathbf{x}}}(t) = \rho]$ (symbols). The numerical value of the rate function at the mean $\rho = \langle\overline{\rho^U_{\mathbf{x}}}(t)\rangle_s$ was subtracted. (f, g) As in (d) and (e) but for the x component of the current $[\overline{\mathbf{J}^U_{\mathbf{x}}}]_x$ instead of the density $\overline{\rho^U_{\mathbf{x}}}$.

also required for the mean empirical current [see comment below Eq. (3.5.17)] and for consistency of time reversal (e.g., to obtain the symmetry in Eqs. (3.5.45) and (3.5.49); see also Fig. 3.4.2 in Sec. 3.4.

3.5.9 Short and long trajectories and the central-limit regime

As already noted on several occasions, in the case of steady-state initial conditions the mean values of the time-averaged density and current are time independent [see Eqs. (3.5.15) and (3.5.17)]. The correlation and (co)variance results [Eqs. (3.5.23), (3.5.45), and (3.5.49) with integral operator (3.5.22)] display a non-trivial temporal behavior dictated by the time integrals $\frac{1}{t}\int_0^t \mathrm{d}t' \left(1 - \frac{t'}{t}\right)$ over two-point densities $P_{\mathbf{z}}(\mathbf{z}', t')$.

In Fig. 3.5.7a-c we depict this time-dependent behavior for the two-dimensional rotational flow in Eq. (3.5.2) for $\mathbf{x} = \mathbf{y}$. The short-time behavior can be obtained by analogy to

Phys. Rev. Res. 4, 033243 (2022)

the short-time expansion in the Appendix. Note that the short-time limit of fluctuations of time-integrated currents recently attracted much attention in the context of inference of dissipation, since in this limit the thermodynamic uncertainty relation becomes sharp [153, 154]. The long-time behavior shows that $\mathbf{C}(t), \mathrm{var}(t) \propto t^{-1}$, as expected from the central limit theorem (and large deviation theory) due to sufficiently many sufficiently uncorrelated visits of the window region. Accordingly, a serious problem is encountered in dimensions ≥ 2 in the limit $h \to 0$ because diffusive trajectories do not hit points (for a detailed discussion see Sec. 3.4).

The limit of $t\mathbf{C}(t)$, $t\mathrm{var}(t)$ for large t can be obtained as follows. We have $\int_{t'}^{\infty} \mathrm{d}t''[P_{\mathbf{y}}(\mathbf{x}, t'') - p_{\mathrm{s}}(\mathbf{x})] \to 0$ for $t' \to \infty$ since $P_{\mathbf{y}}(\mathbf{x}, t') \overset{t'\to\infty}{\longrightarrow} p_{\mathrm{s}}(\mathbf{x})$ and $\hat{\mathbf{j}}_{\mathbf{x}} P_{\mathbf{y}}(\mathbf{x}, t') \overset{t'\to\infty}{\longrightarrow} \mathbf{j}_{\mathrm{s}}(\mathbf{x})$ with exponentially decaying deviations. This implies that for large t, we can replace $\frac{1}{t}\int_0^t \mathrm{d}t' \left(1 - \frac{t'}{t}\right)$ by $\frac{1}{t}\int_0^{\infty} \mathrm{d}t'$ in the integral operator (3.5.22). This replacement of integrals and the scaling are also confirmed by a spectral expansion (see, e.g., Ref. [41] for spectral-theoretic results for the empirical density).

We now discuss the central-limit regime, which is contained in large deviation theory as small deviations from the mean. According to the central limit theorem [for not almost surely constant $U_{\mathbf{x}}^h$, and for finite variances (i.e., strictly positive h, see Sec. 3.5.5)], the probability distributions $p(A_t = a)$ for $A_t = \overline{\rho_{\mathbf{x}}^U}(t)$ and $A_t = \overline{\mathbf{J}_{\mathbf{x}}^U}(t)$ become Gaussian for large t. This is contained in large deviation theory in terms of a parabola that locally (for $a \approx \mu$) approximates the rate function

$$I(a) = -\lim_{t\to\infty} \frac{1}{t} \ln p(A_t = a) \approx \frac{(a-\mu)^2}{2\sigma_A^2}, \tag{3.5.74}$$

where the mean μ is given by $\langle \rho_{\mathbf{x}}^U(t) \rangle_{\mathrm{s}} = \int \mathrm{d}\mathbf{z} U_{\mathbf{x}}^h(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z})$ and $\langle \mathbf{J}_{\mathbf{x}}^U(t) \rangle_{\mathrm{s}} = \int \mathrm{d}\mathbf{z} U_{\mathbf{x}}^h(\mathbf{z}) \mathbf{j}_{\mathrm{s}}(\mathbf{z})$ [see Eqs. (3.5.15) and (3.5.17)] and the large deviation variance σ_A^2 follows by the above arguments from Eqs. (3.5.23) and (3.5.49) for $\mathbf{x} = \mathbf{y}$ as in Eq. (3.5.61) as

$$\begin{aligned} \sigma^2_{\overline{\rho_{\mathbf{x}}^U}} &\equiv \lim_{t\to\infty} t\, \mathrm{var}_{\rho}^{\mathbf{x}}(t) \\ &= 2\int_0^{\infty} \mathrm{d}t' \int \mathrm{d}\mathbf{z} \int \mathrm{d}\mathbf{z}' U_{\mathbf{x}}^h(\mathbf{z}) U_{\mathbf{x}}^h(\mathbf{z}') \left[P_{\mathbf{z}'}(\mathbf{z}, t) - p_{\mathrm{s}}(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}')\right], \end{aligned} \tag{3.5.75}$$

as well as

$$\begin{aligned} \sigma^2_{\overline{\mathbf{J}_{\mathbf{x}}^U}} &\equiv \lim_{t\to\infty} t\, \mathrm{var}_{\mathbf{J}}^{\mathbf{x}}(t) \\ &= 2\int \mathrm{d}\mathbf{z}\, \mathrm{Tr}[\mathbf{D}(\mathbf{z})][U_{\mathbf{x}}^h]^2(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}) \end{aligned}$$

Phys. Rev. Res. 4, 033243 (2022)

$$+2\int_0^\infty \mathrm{d}t' \int \mathrm{d}\mathbf{z} \int \mathrm{d}\mathbf{z}' U_{\mathbf{x}}^h(\mathbf{z}) U_{\mathbf{x}}^h(\mathbf{z}') \left[\hat{\mathbf{j}}_{\mathbf{z}} \cdot \hat{\mathbf{j}}_{\mathbf{z}'}^{\ddagger} P_{\mathbf{z}'}(\mathbf{z}, t) - \mathbf{j}_{\mathrm{s}}(\mathbf{z}) \mathbf{j}_{\mathrm{s}}(\mathbf{z}')\right]. \tag{3.5.76}$$

For any Lebesgue integrable window function $U_{\mathbf{x}}^h$ (i.e., if the window size h fulfills $h > 0$), and in $d = 1$ even for the δ function, this variance is finite, and the central limit theorem applies as described above. The parabolic approximation for the rate function for a two-dimensional system with finite window size $h > 0$ is shown for the density $\overline{\rho_{\mathbf{x}}^U}(t)$ and current $\overline{\mathbf{J}_{\mathbf{x}}^U}(t)$ in Fig. 3.5.7e,g. The agreement of the simulation and the variance given by Eqs. (3.5.75) and (3.5.76) is readily confirmed.

If we instead take the limit of no coarse graining $h \to 0$ in multidimensional space $d \geq 2$, the variances diverge [see Eq. (3.5.51)]. Figure 3.5.8 depicts the distribution of the empirical density $\overline{\rho_{\mathbf{x}}^U}(t)$ in a fixed point $\mathbf{x}$ for different times t and window sizes h. We see that the distribution becomes non-Gaussian for small h, in particular the most probable value departs from the mean and approaches zero. Even though a Gaussian distribution is restored for longer times (see Fig. 3.5.8b), for even smaller window sizes the distribution again becomes non-Gaussian (see Fig. 3.5.8c). This behavior is not surprising, since Gaussian distributions are only expected for sufficiently many (sufficiently uncorrelated) visits of the coarse graining window. For $h \to 0$ the recurrence time to return to the window diverges and thus for any finite t one cannot expect a Gaussian distribution. Note that it is not clear whether a limit in distribution for $h \to 0$ of $\overline{\rho_{\mathbf{x}}^U}$ and $\overline{\mathbf{J}_{\mathbf{x}}^U}$ even exists. We hypothesize that, if the limit $h \to 0$ of the distribution indeed exists, then it does so only as a scaling limit with $h \to 0$ and $t \to \infty$ simultaneously.

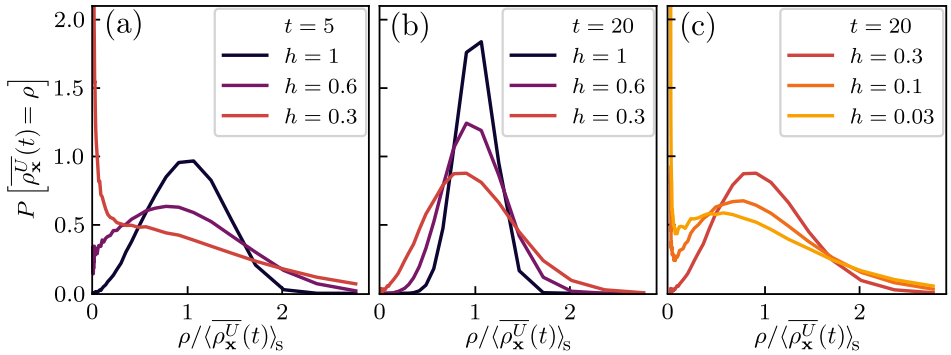

Figure 3.5.8: Simulation of the probability density function of the empirical density $\overline{\rho_{\mathbf{x}}^U}(t)$ for $\mathbf{x} = (0, 1)^T$ and Gaussian window function $U_{\mathbf{x}}^h$ in Eq. (3.5.3) with width h for the two-dimensional driven Ornstein-Uhlenbeck process in Eq. (3.5.2) with $\Omega = 3$ with $\mathbf{x}_0$ starting from the steady state. The simulated probability densities were obtained from histograms of 2×10^5 trajectories for each set of parameters.

Phys. Rev. Res. 4, 033243 (2022)

3.5.10 Outlook beyond overdamped dynamics

In this section we give a brief outlook on the relevance of our findings in the limit $h \to 0$ for processes that are *not* described by purely overdamped dynamics. In particular, we highlight that although in physical systems the assumption of overdamped dynamics breaks down at very small time or length scales (which often may not be observable), the predicted divergence of fluctuations in the limit $h \to 0$ does not break down, or at least it remains true for sufficiently small finite h that empirical densities and currents attain numerically very large values, i.e., effectively diverge. We emphasize that this section only establishes an outlook that underscores the experimental relevance of our approach, but does not contain quantitative theoretical results. Note that beyond the examples given here, the results in the limit $h \to 0$ also apply to Markov jump processes, as illustrated in the Supplemental Material of in Subsec. 3.4.7.

To go beyond the assumption of Markovian overdamped motion assumed in Eq. (3.5.4), Fig. 3.5.9 depicts the fluctuations of the empirical density and current for two very different types of stochastic dynamics. In Fig. 3.5.9a,b we evaluate the functionals in Eq. (3.5.1)

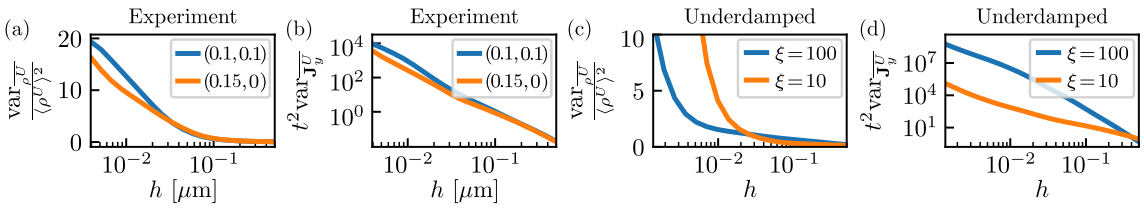

Figure 3.5.9: (a) Variance divided by squared mean of the empirical density and (b) variance of the empirical current with Gaussian coarse graining window [see Eqs. (3.5.1) and (3.5.3)] around $\mathbf{x} = (0.1, 0.1)^T \mu$m and $\mathbf{x} = (0.15, 0)^T \mu$m for 200 experimental trajectories $(\mathbf{x}_\tau - \mathbf{x}_0)_{\tau \leq 50\ \mathrm{s}}$ measured in particle tracking in living cells [91] with time-step $\mathrm{d}t = 0.1$ s and spatial resolution 10^{-3} μm. (c, d) Analogous results for underdamped dynamics simulated according to Eq. (3.5.77) for 1000 trajectories with an Euler integration scheme with time step $\mathrm{d}t = 0.02$, total time 10, initial positions $\mathbf{x}_0 = (0, 0)^T$, and Maxwellian initial velocities $\mathbf{v}_0$ (i.e., zero-mean Gaussian with variance $k_\mathrm{B}T$), evaluated with Gaussian coarse graining window around $\mathbf{x} = (0.1, 0)^T$. The observables in (a) and (c) and in (b) and (d) relate to the predicted analytical curves in Figs. 3.5.3 and 3.5.6c, respectively, but now for completely different underlying dynamics. We note that processes not following an overdamped motion, in fact even non-Markovian processes, also seem to feature the divergence of fluctuations in the limit $h \to 0$, at least down to relevant small values of $h > 0$ (larger than the resolution yet small enough to display impractically large fluctuations to be experimentally useful).

with a Gaussian window function from Eq. (3.5.3) for particle-tracking data in living cells that was found to be well described by a two-state fractional Brownian motion [90, 91, 202]. The latter in particular is non-Markovian with subdiffusive antipersistence on a given

Phys. Rev. Res. 4, 033243 (2022)

timescale. We observe that, even though the assumption of Markovian overdamped motion in Eq. (3.5.4) is obviously violated (on some time and spatial scales), and thus the results of our work do not necessarily apply, we still find divergent fluctuations in the limit of small coarse-graining scales h. Note that the resolution of the measurement is $h = 10^{-3}\,\mu$m [91]. In Fig. 3.5.9a,b we observe that even for h above this resolution limit, the fluctuations approach impracticably large values. Therefore, we propose that in general scenarios (e.g., in this experimental setup that extends way beyond the discussed overdamped dynamics) coarse graining empirical densities and currents may even in the case of very good statistics be necessary to obtain experimentally meaningful values with limited fluctuations.

In Fig. 3.5.9c,d we similarly evaluate the coarse-grained empirical density and current for two-dimensional *underdamped* harmonically confined Langevin dynamics with friction constant ξ (setting for convenience the mass $m = 1$ and temperature $k_{\mathrm{B}}T = 1$) simulated by integrating the equations of motion

$$\begin{aligned} \mathrm{d}\mathbf{x}_t &= \mathbf{v}_t \mathrm{d}t, \\ \mathrm{d}\mathbf{v}_t &= -\xi \mathbf{v}_t \mathrm{d}t - \mathbf{x}_t \mathrm{d}t + \sqrt{2\xi}\mathrm{d}\mathbf{W}_t. \end{aligned} \tag{3.5.77}$$

This dynamics exhibits persistence on timescales around or below m/ξ (i.e., the ballistic regime). Again, we find in Fig. 3.5.9c,d that the divergence predicted in the limit $h \to 0$ for overdamped dynamics is qualitatively preserved. The quantitative order of divergence will depend on the details of the process. We hypothesize that on time scales h^2/D (with diffusion constant $D = k_{\mathrm{B}}T/\xi$) that are smaller that m/ξ the ballistic regime will cause deviations from the predicted divergence results in Sec. 3.4. Following the arguments in Sec. 3.5.5, the expressions will still diverge since the probability to hit points in $(d \geq 2)$-dimensional space becomes zero.

The influence of the details of the process, such as memory effects and ballistic transport, constitutes an interesting direction for future research that, however, goes beyond the scope of the present work. From the qualitative behavior found in Fig. 3.5.9 we may already conclude that the relevance of coarse-graining empirical densities and currents to ensure finite and manageable fluctuations appears to be a quite general result, exceeding beyond the overdamped dynamics discussed in this work.

3.5.11 Conclusion

In this extended exposé accompanying the Letter reproduced in Sec. 3.4 we presented the conceptual and technical background that is required to describe and understand the statistics of the empirical density and current of steady-state diffusions, which are central

Phys. Rev. Res. 4, 033243 (2022)

to statistical mechanics and thermodynamics on the level of individual trajectories. In order to gain deeper insight into the meaning of fluctuations of the empirical density and current, we made use of a generalized time-reversal symmetry. We carried out a systematic analysis of the effect of a spatial coarse graining. A systematic variation of the coarse-graining scale in an *a posteriori* smoothing of trajectory data was proposed as an efficient method to infer bounds on a system's dissipation. Moreover, we discussed symmetries in the statistics of the empirical current and density that arise as a result of the (generalized) time-reversal symmetry. Throughout the work, we advocated the application of stochastic calculus, which is very powerful in the analysis of related problems and represents a more direct alternative to Feynman-Kac theory and path-integral methods. The technical background and concepts presented here may serve as a basis for forthcoming publications, including the generalization of the presented inference strategy to windows that are not centered at an individual point, as well as the use of the correlations result entering the CTUR inequality [134].

Acknowledgments—We thank Diego Krapf and Matthias Weiss for kindly providing traces from their particle-tracking experiments. Financial support from Studienstiftung des Deutschen Volkes (to C. D.) and the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2 (to A. G.) is gratefully acknowledged.

Phys. Rev. Res. 4, 033243 (2022)

3.5.12 Appendix: Derivations in the limit of no coarse graining

We now take the limit to very small window sizes, i.e., the limit to no coarse graining, which will turn out to depend only on the properties of the two-point functions $P_{\mathbf{x}}(\mathbf{y}, t')$ for small time differences $t' = t_2 - t_1$. This allows us to derive the bounds in Eq. (3.5.51). We consider normalized window functions such that for a window size $h \to 0$ the window function becomes a δ distribution $U^h_{\mathbf{x}}(\mathbf{z}) \to \delta(\mathbf{x} - \mathbf{z})$.

Density variance

For the variance of the density $\mathrm{var}^{\mathbf{x}}_{\rho}(t) \equiv \langle \overline{\rho^U_{\mathbf{x}}}(t)^2 \rangle_{\mathrm{s}} - \langle \overline{\rho^U_{\mathbf{x}}}(t) \rangle^2_{\mathrm{s}}$, we have [see Eq. (3.5.23)]

$$\mathrm{var}^{\mathbf{x}}_{\rho}(t) = 2\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z}, t') - p_{\mathrm{s}}(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}')]. \tag{3.5.78}$$

For window size $h \to 0$ the mean remains finite such that $\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[p_{\mathrm{s}}(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}')] \overset{h\to 0}{\longrightarrow} -2p_{\mathrm{s}}(\mathbf{x})^2 = O(h^0)$. Now consider (the notation $\mathrm{d}^d z$ denotes, exactly as $\mathrm{d}\mathbf{z}$, the volume integral $\mathrm{d}z_1 \mathrm{d}z_2 \cdots \mathrm{d}z_d$)

$$\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z}, t')] = \frac{1}{t} \int_0^t \mathrm{d}t' \left(1 - \frac{t'}{t}\right) \int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') P_{\mathbf{z}'}(\mathbf{z}, t'). \tag{3.5.79}$$

For $t' > \varepsilon > 0$, $P_{\mathbf{z}'}(\mathbf{z}, t')$ is bounded and thus $\int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') P_{\mathbf{z}'}(\mathbf{z}, t')$ is bounded using $||P_{\mathbf{z}'}(\mathbf{z}, \varepsilon)||_\infty = O(h^0)$. Contributions diverging for $h \to 0$ can thus only come from the $t' \to 0$ part of the integral, i.e., from **small time differences** $t' = t_2 - t_1$ (but **not** small absolute time t) in the Dyson series. To get the dominant divergent contribution, we can thus set $1 - t'/t \to 1$ and replace the two-point function $P_{\mathbf{z}'}(\mathbf{z}, t')$ by the short time propagator $P_{\mathbf{y}}(\mathbf{z}, t') \to p_{\mathrm{s}}(\mathbf{z}') G_{\mathrm{short}}(\mathbf{z}, t'|\mathbf{z}')$ which reads (for simplicity take $\mathbf{D}(\mathbf{z}) = D\mathbf{1}$, which we generalize later) [17]

$$\begin{aligned} G_{\mathrm{short}}(\mathbf{z}, t'|\mathbf{z}') &= (4\pi D t')^{-d/2} \exp\left[-\frac{[\mathbf{z} - \mathbf{z}' - F(\mathbf{z}')t']^2}{4Dt'}\right] \\ &= (4\pi D t')^{-d/2} \exp\left[-\frac{[\mathbf{z} - \mathbf{z}']^2}{4Dt'} + \frac{2(\mathbf{z} - \mathbf{z}') \cdot \mathbf{F}(\mathbf{z}')t'}{4Dt'} + O(t')\right] \\ &\overset{\mathbf{z} \approx \mathbf{z}'}{\approx} (4\pi D t')^{-d/2} \left[1 + \frac{1}{4D}(\mathbf{z} - \mathbf{z}') \cdot \mathbf{F}(\mathbf{z}')\right] \exp\left[-\frac{[\mathbf{z} - \mathbf{z}']^2}{4Dt'}\right]. \end{aligned} \tag{3.5.80}$$

We write for $t' \to 0, \mathbf{z} - \mathbf{z}' \to \mathbf{0}$,

$$\begin{aligned} G_{\mathrm{short},2} &\equiv (4\pi D t')^{-d/2} \left[1 + \frac{1}{2D}(\mathbf{z} - \mathbf{z}') \cdot \mathbf{F}(\mathbf{z}')\right] \exp\left[-\frac{[\mathbf{z} - \mathbf{z}']^2}{4Dt'}\right], \\ G_{\mathrm{short},3} &\equiv (4\pi D t')^{-d/2} \exp\left[-\frac{[\mathbf{z} - \mathbf{z}']^2}{4Dt'}\right], \end{aligned} \tag{3.5.81}$$

where $G_{\mathrm{short},2}$ can be replaced by $G_{\mathrm{short},3}$ (since $\mathbf{z} - \mathbf{z}'$ is small) if $G_{\mathrm{short},3}$ does not give zero in the integrals.

For Gaussian window functions

$$U^h_{\mathbf{x}}(\mathbf{z}) = (2\pi h^2)^{-d/2} \exp\left[-\frac{(\mathbf{z} - \mathbf{x})^2}{2h^2}\right], \tag{3.5.82}$$

we obtain for the spatial integrals

$$\begin{aligned} &\int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') G_{\mathrm{short},3}(\mathbf{z}, t'|\mathbf{z}') p_{\mathrm{s}}(\mathbf{z}') \\ &\simeq p_{\mathrm{s}}(\mathbf{x}) \int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') G_{\mathrm{short},3}(\mathbf{z}, t'|\mathbf{z}') \\ &= p_{\mathrm{s}}(\mathbf{x})(2\pi h^2)^{-d}(4\pi D t')^{-d/2} \int \mathrm{d}^d z \int \mathrm{d}^d z' \exp\left[-\frac{(\mathbf{z} - \mathbf{z})^2}{2h^2} - \frac{(\mathbf{z}' - \mathbf{z})^2}{2h^2} - \frac{(\mathbf{z} - \mathbf{z}')^2}{4Dt'}\right] \\ &= p_{\mathrm{s}}(\mathbf{x})(2\pi h^2)^{-d}(4\pi D t')^{-d/2} \left(\int \mathrm{d}x_1 \int \mathrm{d}y_1 \exp\left[-\frac{x_1^2}{2h^2} - \frac{y_1^2}{2h^2} - \frac{(x_1 - y_1)^2}{4Dt'}\right]\right)^d \end{aligned}$$

Phys. Rev. Res. 4, 033243 (2022)

$$= p_{\mathrm{s}}(\mathbf{x}) \left[\frac{\sqrt{2Dh^2t' + h^4}}{2\sqrt{\pi}h^2\sqrt{2Dt' + h^2}\sqrt{\frac{Dt'}{h^2} + 1}} \right]^d$$
$$= p_{\mathrm{s}}(\mathbf{x})(4\pi)^{-d/2}(Dt' + h^2)^{-d/2}. \tag{3.5.83}$$

This implies, throughout denoting by $\simeq$ asymptotic equality in the limit $h \to 0$ (i.e., equality of the largest order),

$$\begin{aligned}
\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z}, t')] &\simeq (4\pi)^{-d/2} \frac{p_{\mathrm{s}}(\mathbf{x})}{t} \int_0^t \mathrm{d}t' (Dt' + h^2)^{-d/2} \\
&= (4\pi)^{-d/2} \frac{p_{\mathrm{s}}(\mathbf{x})}{t} \times \begin{cases} -\frac{h^{2-d}}{D\left(1-\frac{d}{2}\right)} + \frac{\left(D+h^2\right)^{1-\frac{d}{2}}}{D\left(1-\frac{d}{2}\right)} & \text{for } d \neq 2 \\ -\frac{\log\left(h^2\right)}{D} + \frac{\log\left(D+h^2\right)}{D} & \text{otherwise} \end{cases} \\
&\simeq (4\pi)^{-d/2} \frac{2p_{\mathrm{s}}(\mathbf{x})}{Dt} \times \begin{cases} \frac{h^{2-d}}{d-2} & \text{for } d > 2 \\ -\log(h) & \text{for } d = 2. \end{cases}
\end{aligned} \tag{3.5.84}$$

This gives for Gaussian U with width h the result

$$\mathrm{var}^{\mathbf{x}}_{\rho}(t) \overset{h\to 0}{\simeq} (4\pi)^{-d/2} \frac{4p_{\mathrm{s}}(\mathbf{x})}{Dt} \times \begin{cases} \frac{h^{2-d}}{d-2} & \text{for } d > 2 \\ -\log(h) & \text{for } d = 2, \end{cases} \tag{3.5.85}$$

where only the numerical prefactor changes if we choose other indicator functions, since the relevant part (close to $\mathbf{x}$) of any finite size window function can be bounded from above and below by Gaussian window functions.

To extend to general diffusion matrices $\mathbf{D}(\mathbf{z})$, we first note that for $h \to 0$ only the *local* diffusion matrix $\mathbf{D}(\mathbf{x})$ at position $\mathbf{x}$ will enter the result, and, if the local $\mathbf{D}(\mathbf{x})$ is not isotropic we transform to coordinates where the diffusion matrix is diagonal, $\mathbf{D}(\mathbf{x}) = \mathrm{diag}(D_1(\mathbf{x}), \ldots, D_d(\mathbf{x}))$. One can check this by Taylor expanding around $\mathbf{x}$ in h in the local coordinate frame, isolating the leading-order term, keeping in mind that $\mathbf{D}(\mathbf{z})$ was assumed to be smooth. In the local coordinates we then need to evaluate the integral

$$\int_0^t \mathrm{d}t' \prod_{i=1}^{d} (D_i(\mathbf{x})t' + h^2)^{-1/2}, \tag{3.5.86}$$

whose integrand can be bounded by

$$(\max_j(D_j(\mathbf{x}))t' + h^2)^{-1/2} \leq (D_i(\mathbf{x})t' + h^2)^{-1/2} \leq (\min_j(D_j(\mathbf{x}))t' + h^2)^{-1/2}, \tag{3.5.87}$$

Phys. Rev. Res. 4, 033243 (2022)

implying that in the final result D in Eq. (3.5.85) can be replaced by $\tilde{D}(\mathbf{x}) \in [\min(D_i(\mathbf{x})), \max(D_i(\mathbf{x}))]$,

$$\mathrm{var}_\rho^{\mathbf{x}}(t) \overset{h\to 0}{\simeq} (4\pi)^{-d/2} \frac{4p_{\mathrm{s}}(\mathbf{x})}{\tilde{D}(\mathbf{x})t} \times \begin{cases} \frac{h^{2-d}}{d-2} & \text{for } d > 2 \\ -\log(h) & \text{for } d = 2. \end{cases} \tag{3.5.88}$$

The entries $D_i(\mathbf{x})$ of the diagonalized $\mathbf{D}(\mathbf{x})$ are the eigenvalues; hence in general $\tilde{D}(\mathbf{x}) \in [\lambda(\mathbf{x})_{\min}, \lambda(\mathbf{x})_{\max}]$ is bounded by the lowest and highest eigenvalues $\lambda(\mathbf{x})_{\min}$ and $\lambda(\mathbf{x})_{\max}$ of the matrix $\mathbf{D}(\mathbf{x})$. This proves the density variance result in Eq. (3.5.51).

Correlation of current and density

Now consider the small-window limit for correlations with $\mathbf{x} = \mathbf{y}$ defined as $\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t) \equiv \langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t)\overline{\rho^U_{\mathbf{x}}}(t)\rangle_{\mathrm{s}} - \langle \overline{\mathbf{J}^U_{\mathbf{x}}}(t)\rangle_{\mathrm{s}}\langle \overline{\rho^U_{\mathbf{x}}}(t)\rangle_{\mathrm{s}}$, given according to Eq. (3.5.45) by

$$\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}\rho}(t) = \hat{\mathcal{I}}^{t,U}_{\mathbf{xx}} \left[\hat{\mathbf{j}}_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z}, t') + \hat{\mathbf{j}}^\dagger_{\mathbf{z}} P_{\mathbf{z}}(\mathbf{z}', t') - 2\mathbf{j}_{\mathrm{s}}(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}') \right]. \tag{3.5.89}$$

Recall that $\hat{\mathbf{j}}_{\mathbf{z}} = \mathbf{F}(\mathbf{z}) - \mathbf{D}(\mathbf{z})\nabla_{\mathbf{z}}$. As always the term involving the mean values is finite for $h \to 0$. We first look at $\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\hat{\mathbf{j}}_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z}, t')]$, again, first for constant isotropic diffusion $\mathbf{D}(\mathbf{z}) = D\mathbf{1}$.

Here we have to use $G_{\mathrm{short},2}$ [see Eq. (3.5.81)], since the integrals over $G_{\mathrm{short},3}$ vanish. Hence, consider

$$\begin{aligned} \hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\hat{\mathbf{j}}_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z}, t')] &= \frac{1}{t} \int_0^t \mathrm{d}t' \left(1 - \frac{t'}{t}\right) \int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') \hat{\mathbf{j}}_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z}, t') \\ &\simeq \frac{p_{\mathrm{s}}(\mathbf{x})}{t} \int_0^t \mathrm{d}t' \int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') [\mathbf{F}(\mathbf{z}) - D\nabla_{\mathbf{z}}] G_{\mathrm{short},2}(\mathbf{z}, t'|\mathbf{z}'), \end{aligned} \tag{3.5.90}$$

where we can use $\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\mathbf{F}(\mathbf{z}) P_{\mathbf{z}'}(\mathbf{z}, t')] \simeq \mathbf{F}(\mathbf{x}) \hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z}, t')] \simeq \mathbf{F}(\mathbf{x}) \times$ Eq. (3.5.84) and we compute

$$\begin{aligned} &\nabla_{\mathbf{z}} G_{\mathrm{short},2}(\mathbf{z}, t'|\mathbf{z}') \\ =&(4\pi D t')^{-d/2} \left[1 + \frac{1}{2D}(\mathbf{z} - \mathbf{z}') \cdot \mathbf{F}(\mathbf{z}')\right] \nabla_{\mathbf{z}} \exp\left[-\frac{[\mathbf{z} - \mathbf{z}']^2}{4Dt'}\right] \\ &+ (4\pi D t')^{-d/2} \frac{\mathbf{F}(\mathbf{z}')}{2D} \exp\left[-\frac{[\mathbf{z} - \mathbf{z}']^2}{4Dt'}\right] \\ =&-(4\pi D t')^{-d/2} \left[1 + \frac{1}{2D}(\mathbf{z} - \mathbf{z}') \cdot \mathbf{F}(\mathbf{z}')\right] \frac{\mathbf{z} - \mathbf{z}'}{2Dt'} \exp\left[-\frac{[\mathbf{z} - \mathbf{z}']^2}{4Dt'}\right] \end{aligned}$$

Phys. Rev. Res. 4, 033243 (2022)

$$+ (4\pi Dt')^{-d/2} \frac{\mathbf{F}(\mathbf{z}')}{2D} \exp\left[-\frac{[\mathbf{z}-\mathbf{z}']^2}{4Dt'}\right]. \tag{3.5.91}$$

By symmetry, the spatial integrals over $(\mathbf{z}-\mathbf{z}')\exp\left[-\frac{[\mathbf{z}-\mathbf{z}']^2}{4Dt'}\right]$ vanish, and we are left to compute

$$\begin{aligned}
&- D\int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') \nabla_{\mathbf{z}} G_{\mathrm{short},2}(\mathbf{z}, t'|\mathbf{z}') \\
&\simeq - D(4\pi Dt')^{-d/2} \int \mathrm{d}^d x \int \mathrm{d}^d y U(\mathbf{x}) U(\mathbf{y}) \left(-\frac{[(\mathbf{x}-\mathbf{y})\cdot\mathbf{F}(\mathbf{y})](\mathbf{x}-\mathbf{y})}{4D^2 t'} + \frac{\mathbf{F}(\mathbf{z}')}{2D}\right) \\
&\times \exp\left[-\frac{[\mathbf{z}-\mathbf{z}']^2}{4Dt'}\right],
\end{aligned} \tag{3.5.92}$$

where the second term gives $-\frac{1}{2}\mathbf{F}(\mathbf{x})\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z},t')]$. The remaining term, noting that $\mathbf{F}(\mathbf{z}') \simeq \mathbf{F}(\mathbf{x})$ and integrating out all directions except k for the $F_k(\mathbf{x})$ component [by symmetry $(z_i - z_i')(z_j - z_j')$ integrates to zero if $i \neq j$], becomes

$$\begin{aligned}
&\frac{(4\pi Dt')^{-d/2}}{4Dt'} \int \mathrm{d}^d z \int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}') [(\mathbf{z}-\mathbf{z}')\cdot\mathbf{F}(\mathbf{x})](\mathbf{z}-\mathbf{z}') \exp\left[-\frac{[\mathbf{z}-\mathbf{z}']^2}{4Dt'}\right] \\
&= \frac{\mathbf{F}(\mathbf{x})}{4Dt'} \int \mathrm{d}z_1 \int \mathrm{d}z_1' U^1(z_1) U^1(z_1') (z_1 - z_1')^2 G_{\mathrm{short},3,\mathrm{one-dim}}(z_1, t'|z_1') \\
&= \frac{\mathbf{F}(\mathbf{x})}{4Dt'} \frac{Dh^2 t'}{\sqrt{\pi}\,(Dt'+h^2)^{\frac{3}{2}}} \\
&= \frac{\mathbf{F}(\mathbf{x}) h^2}{4\sqrt{\pi}\,(Dt'+h^2)^{\frac{3}{2}}}.
\end{aligned} \tag{3.5.93}$$

This term is subdominant as we see from the time integral

$$\begin{aligned}
\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[-D\nabla_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z},t')] &\simeq -\frac{\mathbf{F}(\mathbf{x})}{2}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z},t')] + \frac{p_{\mathrm{s}}(\mathbf{x})\mathbf{F}(\mathbf{x})h^2}{4\sqrt{\pi t}} \underbrace{\int_0^t \mathrm{d}t' \left(Dt'+h^2\right)^{-\frac{3}{2}}}_{h^{-1}} \\
&\simeq -\frac{\mathbf{F}(\mathbf{x})}{2}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z},t')].
\end{aligned} \tag{3.5.94}$$

Hence, overall we get

$$\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\hat{\mathbf{j}}_{\mathbf{z}} P_{\mathbf{z}'}(\mathbf{z},t')] = \frac{\mathbf{F}(\mathbf{x})}{2}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z},t')]. \tag{3.5.95}$$

The generalization to non-constant or non-isotropic $\mathbf{D}(\mathbf{z})$ only changes $\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}'}(\mathbf{z},t')]$ but Eq. (3.5.95) is retained.

Phys. Rev. Res. 4, 033243 (2022)

Now consider $\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}P_{\mathbf{z}}(\mathbf{z}',t')]$. Since this involves derivatives of both G and $p_{\rm s}$ (at the initial point) we instead take the form $\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}=\mathbf{j}_{\rm s}(\mathbf{z})/p_{\rm s}(\mathbf{z})+Dp_{\rm s}(\mathbf{z})\nabla_{\mathbf{z}}p_{\rm s}(\mathbf{z})^{-1}$ such that $\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}P_{\mathbf{z}}(\mathbf{z}',t')=[\mathbf{j}_{\rm s}(\mathbf{z})+Dp_{\rm s}(\mathbf{z})\nabla_{\mathbf{z}}]G(\mathbf{z}',t'|\mathbf{z})$ giving

$$\begin{aligned}
&\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}P_{\mathbf{z}}(\mathbf{z}',t')]\\
&=\frac{1}{t}\int_0^t \mathrm{d}t'\left(1-\frac{t'}{t}\right)\int \mathrm{d}^d z\int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z})U^h_{\mathbf{x}}(\mathbf{z}')[\mathbf{j}_{\rm s}(\mathbf{z})+Dp_{\rm s}(\mathbf{z})\nabla_{\mathbf{z}}]G(\mathbf{z}',t'|\mathbf{z})\\
&\simeq \frac{\mathbf{j}_{\rm s}(\mathbf{x})}{p_{\rm s}(\mathbf{x})}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}}(\mathbf{z}',t')]+Dp_{\rm s}(\mathbf{x})\int_0^t \mathrm{d}t'\int \mathrm{d}^d z\int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z})U^h_{\mathbf{x}}(\mathbf{z}')\nabla_{\mathbf{z}}G_{\rm short,2}(\mathbf{z}',t'|\mathbf{z}),
\end{aligned}\tag{3.5.96}$$

where [note that $G_{\rm short,3}(\mathbf{z},t'|\mathbf{z}')=G_{\rm short,3}(\mathbf{z}',t'|\mathbf{z})$]

$$\begin{aligned}
&\nabla_{\mathbf{z}}G_{\rm short,2}(\mathbf{z}',t'|\mathbf{z})\\
=&(4\pi Dt')^{-d/2}\left[1+\frac{1}{2D}(\mathbf{z}'-\mathbf{z})\cdot\mathbf{F}(\mathbf{z}')\right]\nabla_{\mathbf{z}}\exp\left[-\frac{[\mathbf{z}-\mathbf{z}']^2}{4Dt'}\right]\\
&+(4\pi Dt')^{-d/2}\frac{-\mathbf{F}(\mathbf{z})}{2D}\exp\left[-\frac{[\mathbf{z}-\mathbf{z}']^2}{4Dt'}\right]\\
=&-(4\pi Dt')^{-d/2}\left[1+\frac{1}{2D}(\mathbf{z}-\mathbf{z}')\cdot\mathbf{F}(\mathbf{z}')\right]\frac{\mathbf{z}-\mathbf{z}'}{2Dt'}\exp\left[-\frac{[\mathbf{z}-\mathbf{z}']^2}{4Dt'}\right]\\
&-\frac{\mathbf{F}(\mathbf{z}')}{2D}G_{\rm short,3}(\mathbf{z}',t'|\mathbf{z}).
\end{aligned}\tag{3.5.97}$$

As before, asymptotically only the last term contributes, giving

$$\begin{aligned}
&\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}}P_{\mathbf{z}}(\mathbf{z}',t')]\\
&\simeq \frac{\mathbf{j}_{\rm s}(\mathbf{x})}{p_{\rm s}(\mathbf{x})}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}}(\mathbf{z}',t')]+Dp_{\rm s}(\mathbf{x})\int_0^t \mathrm{d}t'\int \mathrm{d}^d z\int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z})U^h_{\mathbf{x}}(\mathbf{z}')\nabla_{\mathbf{z}}G_{\rm short,2}(\mathbf{z}',t'|\mathbf{z})\\
&\simeq \frac{\mathbf{j}_{\rm s}(\mathbf{x})}{p_{\rm s}(\mathbf{x})}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}}(\mathbf{z}',t')]\\
&\quad+Dp_{\rm s}(\mathbf{x})\int_0^t \mathrm{d}t'\int \mathrm{d}^d z\int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z})U^h_{\mathbf{x}}(\mathbf{z}')\frac{-\mathbf{F}(\mathbf{z}')}{2D}G_{\rm short,3}(\mathbf{z}',t'|\mathbf{z})\\
&\simeq\left[\frac{\mathbf{j}_{\rm s}(\mathbf{x})}{p_{\rm s}(\mathbf{x})}-\frac{\mathbf{F}(\mathbf{x})}{2}\right]\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}}(\mathbf{z}',t')].
\end{aligned}\tag{3.5.98}$$

Overall, this gives for the correlations (having the same form for anisotropic diffusion)

$$\mathbf{C}^{\mathbf{xx}}_{\mathbf{J}_\rho}(t)\overset{h\to 0}{\simeq}\frac{\mathbf{j}_{\rm s}(\mathbf{x})}{p_{\rm s}(\mathbf{x})}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}[P_{\mathbf{z}}(\mathbf{z}',t')]\simeq\frac{\mathbf{j}_{\rm s}(\mathbf{x})}{2p_{\rm s}(\mathbf{x})}\mathrm{var}^{\mathbf{x}}_{\rho}(t).\tag{3.5.99}$$

This proves the correlation result in Eq. (3.5.51).

Phys. Rev. Res. 4, 033243 (2022)

Current variance

We now turn to the current variance [see Eq. (3.5.49) for $\mathbf{x} = \mathbf{y}$],

$$\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) = \frac{2}{t}\int \mathrm{d}^d z \mathrm{Tr}\mathbf{D}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}) + 2\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}\left[\hat{\mathbf{j}}_{\mathbf{z}}\cdot\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}'} P_{\mathbf{z}'}(\mathbf{z},t') - \mathbf{j}_{\mathrm{s}}(\mathbf{z})\cdot\mathbf{j}_{\mathrm{s}}(\mathbf{z}')\right]. \tag{3.5.100}$$

The first term for $h \to 0$ gives

$$\frac{2\mathrm{Tr}\mathbf{D}(\mathbf{z})}{t}\int \mathrm{d}^d x U^h_{\mathbf{x}}(\mathbf{z}) U^h_{\mathbf{x}}(\mathbf{z}) p_{\mathrm{s}}(\mathbf{z}) \simeq \frac{2\mathrm{Tr}\mathbf{D}(\mathbf{x})}{t} p_{\mathrm{s}}(\mathbf{x}) U^h_{\mathbf{x}}(\mathbf{x}), \tag{3.5.101}$$

where $U^h_{\mathbf{x}}(\mathbf{x}) \propto h^{-d}$ is the height of the δ function approximation, e.g., $U^h_{\mathbf{x}}(\mathbf{x}) = (2\pi)^{-d/2}h^{-d}$ for Gaussian $U^h_{\mathbf{x}}$. In the derivation (see Sec. 3.5.4) this term occurred from cross correlations $\mathrm{d}W_{t_1}\mathrm{d}W_{t_1} = \mathrm{d}t' \neq 0$ in the noise part, hence it can be seen to come from zero time differences $t' = t_2 - t_1 = 0$. Such a term does not appear in the density variance or density-current correlation, since there $\mathrm{d}t_1\mathrm{d}t_2 = 0$ and $\mathrm{d}t_1\mathrm{d}W_{t_2} = 0$ would occur instead of $\mathrm{d}W_{t_1}\mathrm{d}W_{t_2}$.

Due to the fast h^{-d} divergence, the dominant limit does not depend on terms with no or only one derivative, since they were shown to scale at most as h^{2-d}. The only new term is the second derivative, for which we see that

$$\begin{aligned}\int_0^t \mathrm{d}t' D\nabla_{\mathbf{z}}\cdot(-\nabla_{\mathbf{z}'})G_{\text{short},3}(\mathbf{z},t'|\mathbf{z}') &= \int_0^t D\nabla^2_{\mathbf{z}} G_{\text{short},3}(\mathbf{z},t'|\mathbf{z}') \\ &= \int_0^t \mathrm{d}t'\partial_{t'} G_{\text{short},3}(\mathbf{z},t'|\mathbf{z}') \\ &= \left[G_{\text{short},3}(\mathbf{z},t'|\mathbf{z}')\right]_0^t \\ &= G_{\text{short},3}(\mathbf{z},t|\mathbf{z}') - \delta(\mathbf{z}-\mathbf{z}'),\end{aligned} \tag{3.5.102}$$

such that

$$\begin{aligned}\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}}\left[\hat{\mathbf{j}}_{\mathbf{z}}\cdot\hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}'}P_{\mathbf{z}'}(\mathbf{z},t')\right] &\simeq -D^2\frac{p_{\mathrm{s}}(\mathbf{x})}{t}\int_0^t \mathrm{d}t'\int \mathrm{d}^d z\int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z})U^h(\mathbf{z})\nabla_{\mathbf{z}}\cdot\nabla_{\mathbf{z}'}G_{\text{short},3}(\mathbf{z},t'|\mathbf{z}') \\ &\simeq -D\frac{p_{\mathrm{s}}(\mathbf{x})}{t}\int \mathrm{d}^d z\int \mathrm{d}^d z' U^h_{\mathbf{x}}(\mathbf{z})U^h(\mathbf{z}')\delta(\mathbf{z}-\mathbf{z}') \\ &\simeq -D\frac{p_{\mathrm{s}}(\mathbf{x})}{t}U^h(\mathbf{x}).\end{aligned} \tag{3.5.103}$$

Phys. Rev. Res. 4, 033243 (2022)

For non-isotropic and possibly non-constant $\mathbf{D}(\mathbf{z}) \neq D\mathbf{1}$, we again note that for $h \to 0$ only $\mathbf{D}(\mathbf{z})$ at $\mathbf{z} = \mathbf{x}$ matters, and move to the basis where $\mathbf{D} = \mathbf{D}(\mathbf{x})$ is diagonal, where we have

$$D^2 \nabla_{\mathbf{z}} \cdot \nabla_{\mathbf{z}'} \longrightarrow \sum_{i=1}^{d} D_i^2 \partial_{z_i} \partial_{z_{i'}}. \tag{3.5.104}$$

The operator we need is $\nabla_{\mathbf{z}} \mathbf{D} \nabla_{\mathbf{z}'} = \sum_i D_i \partial_{z_i} \partial_{z_{i'}}$ so we bound one of the D_i in D_i^2 by $D' \in [\min(D_i), \max(D_i)]$ such that we get

$$\hat{\mathcal{I}}^{t,U}_{\mathbf{xx}} \left[\hat{\mathbf{j}}_{\mathbf{z}} \cdot \hat{\mathbf{j}}^{\ddagger}_{\mathbf{z}'} P_{\mathbf{z}'}(\mathbf{z}, t') \right] \simeq -D' \frac{p_{\mathrm{s}}(\mathbf{x})}{t} U^h(\mathbf{x}). \tag{3.5.105}$$

Since $\mathrm{Tr}\mathbf{D} = \sum_i D_i$ we have $\tilde{D}' \equiv \frac{\mathrm{Tr}\mathbf{D} - D'}{d-1} \in [\min(D_i), \max(D_i)]$ and we can write

$$\mathrm{var}^{\mathbf{x}}_{\mathbf{J}}(t) \simeq \frac{2\mathrm{Tr}\mathbf{D}}{t} p_{\mathrm{s}}(\mathbf{x}) U^h(\mathbf{x}) - 2D' \frac{p_{\mathrm{s}}(\mathbf{x})}{t} U^h(\mathbf{x}) = \frac{2\tilde{D}'}{t} p_{\mathrm{s}}(\mathbf{x})(d-1) U(\mathbf{x}), \tag{3.5.106}$$

where $U(\mathbf{x}) \propto h^{-d}$. This proves the current variance result in Eq. (3.5.51). Thus, we see that the current fluctuations diverge for $h \to 0$, except in one-dimensional space where $d - 1 = 0$.

Limit of no coarse graining in the one-dimensional case

In the one-dimensional case, the variance of empirical density and current remain finite for $h \to 0$ which allows to take the limit to $U_x^{h=0}(x') = \delta(x - x')$. In terms of the stochastic integrals, the one-dimensional case is much simpler, since any one-dimensional function $U_x^h(x')$ possesses an antiderivative—a primitive function $\mathcal{U}_x^h(x') = \int^{x'} U_x^h(x'')\mathrm{d}x''$ such that $U_x^h(x') = \partial_{x'} \mathcal{U}_x(x')$. This implies for the Stratonovich integral that

$$t\overline{J_x^U}(t) = \int_0^t U_x^h(x_\tau) \circ \mathrm{d}x_\tau = \mathcal{U}_x(x_t) - \mathcal{U}_x(x_0). \tag{3.5.107}$$

Thus, the stochastic current is no longer a functional but only a function of the initial and end points of the trajectory. Its moments are directly accessible, e.g.,

$$\langle [\overline{J_x^U}(t)]^2 \rangle_{\mathrm{s}} = \frac{1}{t^2} \langle [\mathcal{U}_x(x_t) - \mathcal{U}_x(x_0)]^2 \rangle_{\mathrm{s}} = \frac{1}{t^2} \int \mathrm{d}z \int \mathrm{d}z' \left[\mathcal{U}_x(z) - \mathcal{U}_x(z') \right]^2 P_{z'}(z, t). \tag{3.5.108}$$

If U is Gaussian, then $\mathcal{U}_x$ is the error function such that $[\mathcal{U}_x(x) - \mathcal{U}_x(y)]^2 \leq 1$ and thus $\langle [\overline{J_x^U}(t)]^2 \rangle_{\mathrm{s}} \leq 1/t^2$. This also holds in the limit of a δ function where the primitive function becomes a Heaviside step function, and we get that the current can only be 0 or $\pm t^{-2}$,

Phys. Rev. Res. 4, 033243 (2022)

see Fig. 3.5.10. The current defined with a δ function at x simply counts the net number of crossing through x such that all crossings except maybe one cancel out. Note that the reasoning above only holds for the current defined with a Stratonovich integral—the same definition with an Itô or anti-Itô integral would give a divergent current for the δ function.

To obtain a $1/t$ term as in large deviations one would need to have a steady-state current which could, e.g., be achieved by generalizing to periodic boundary conditions. Then the current would depend on the initial and final points and, in addition, also on the net number of crossings of the full interval between the boundaries of the system.

Figure 3.5.10 shows the time-integrated density and current, i.e., the empirical density and current in Eq. (3.5.1) multiplied by the total time t. Fluctuations remain in the same order of magnitude for $h \to 0$ (see Fig. 3.5.10c,e). We see that the time-integrated current is bounded by 1 which is due to the fact that it simply counts the net number of crossings. According to Eq. (3.5.107) it only depends on the initial point x_0 and end point x_t, in this case x_{10}.

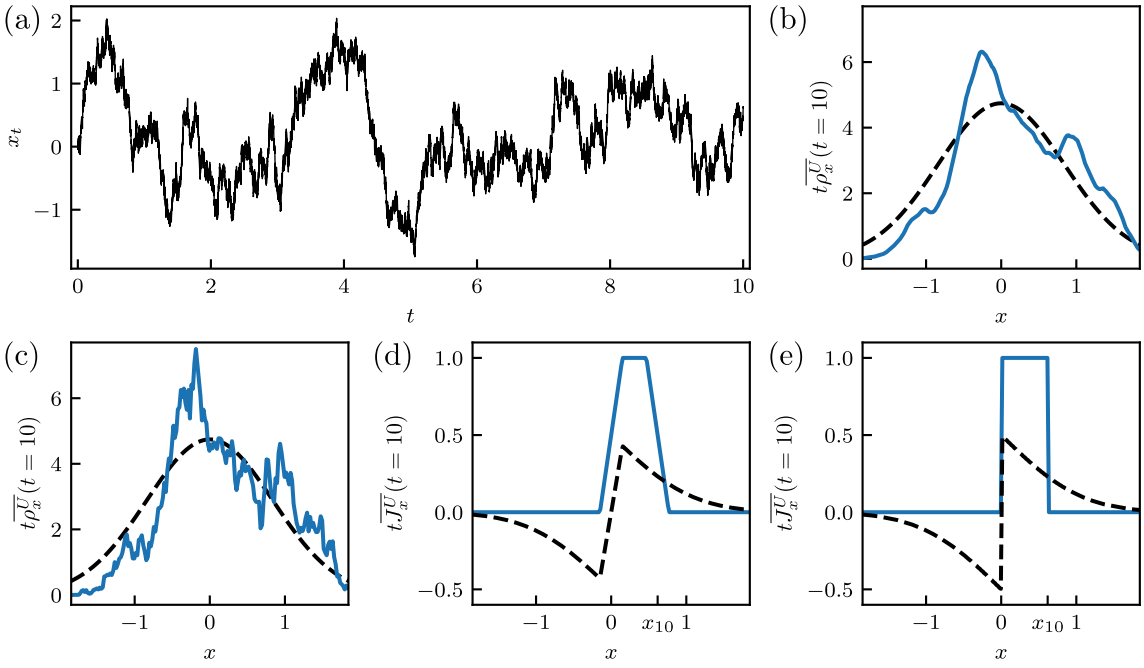

Figure 3.5.10: (a) One-dimensional Brownian motion in a harmonic potential (see Eq. (3.5.2) in 1d with parameters $r = \sqrt{2}$ and $D = 1$) starting at $x_0 = 0$ and ending at $x_{10} = 0.62$. (b) Time-integrated density of the trajectory in (a) as a function of x for normalized window function $U_x^h(x') = h^{-1}\mathbb{1}_{|x-x'|\leq h/2}$ with width $h = 0.3$. The dashed line shows the expectation value of the time-integrated density conditioned on $\mathbf{x}_0 = 0$. (c) As in (b) with width $h = 0.001$. (d) Time-integrated current for window as in (b) width $h = 0.3$. The dashed line shows the expectation value of the time-integrated current conditioned on $\mathbf{x}_0 = 0$. (e) As in (d) for width $h = 0.001$.

Phys. Rev. Res. 4, 033243 (2022)

3.6 On correlations and fluctuations of time-averaged densities and currents with general time dependence (J. Phys. A 2022)

This section is a slightly adapted version of the publication C. Dieball and A. Godec, "On correlations and fluctuations of time-averaged densities and currents with general time dependence", *J. Phys. A: Math. Theor.* **55**, 475001 (2022).

My personal contribution as first author of the publication was in performing calculations and simulations, analyzing the results, and co-writing the manuscript.

On correlations and fluctuations of time-averaged densities and currents with general time dependence

Cai Dieball and Aljaž Godec

Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

Abstract

We present technical results required for the description and understanding of correlations and fluctuations of the empirical density and current, as well as diverse time-integrated and time-averaged thermodynamic currents of diffusion processes with a general time dependence on all timescales. In particular, we generalize the results from Refs. [83, 84, 86] (Secs. 3.4, 3.5, and 3.7) to additive functionals with explicit time dependence and transient or non-ergodic overdamped diffusion. As an illustration, we apply the results to two-dimensional harmonically confined overdamped diffusion in a rotational flow evolving from a non-stationary initial distribution.

J. Phys. A: Math. Theor. **55** 475001 (2022)

3.6.1 Introduction

"Time-average statistical mechanics" focuses on the study of additive functionals of stochastic paths and is important in the analysis of single-particle tracking [37, 130, 203], large deviation theory [43, 107, 110, 143, 204], and stochastic thermodynamics [28, 49, 54, 104, 133, 134, 152, 162], to name but a few. The most important functionals from a physical point of view include the "empirical density" (also known as local or occupation time) [33, 36, 40–42, 135–139], time-integrated and time-averaged currents [34, 43–45, 83, 84, 107, 110, 133, 134, 141–144, 169], and the time-averaged mean squared displacement (see, e.g., [37, 130, 205–209]).

Fluctuations of time-averaged observables have a noise floor—they are bounded from below by the dissipation in a system, which is embodied within the "thermodynamic uncertainty relation" (TUR) [47–49, 67, 71, 126, 132, 210–212]. One may fruitfully exploit this universal lower bound on current fluctuations, e.g., to gauge the thermodynamic cost of precision [47, 152, 213], infer dissipation from fluctuations [67, 83, 84, 132], or to derive thermodynamic limits on the temporal extent of anomalous diffusion [214].

Recent works addressed fluctuations of additive functionals in transient non-equilibrium systems [71, 126, 210], as well as in periodically [49, 211, 212] and generally driven systems [49].

Our aim here is to generalize the direct, stochastic-calculus approach we developed for steady-state systems in [83,84] to transients and systems as well as functionals with explicit time dependence. Note that this includes non-ergodic systems (see, e.g., [40, 139]).

The paper is structured as follows. We first set up the formal background and define the additive functionals in Sec. 3.6.2. In Section 3.6.3 we evaluate the first moments. In Sec. 3.6.4 we present our main result—a Lemma that allows a direct evaluation of fluctuations and correlations of general additive functionals in systems with explicit time dependence—and derive general results for current fluctuations and current-density correlations. In Sec. 3.6.5 we illustrate how to apply the newly developed results by evaluating current-density correlations in overdamped diffusion in a rotational flow evolving from a non-stationary initial distribution. We conclude with a brief outlook.

J. Phys. A: Math. Theor. **55** 475001 (2022)

3.6.2 Set-up

Consider overdamped Langevin dynamics with possibly multiplicative noise and explicit time dependence, described by the anti-Itô (or Hänggi-Klimontovich [197, 198]) stochastic differential equation[9]

$$\mathrm{d}\mathbf{x}_\tau = \mathbf{F}(\mathbf{x}_\tau,\tau)\mathrm{d}\tau + \boldsymbol{\sigma}(\mathbf{x}_\tau,\tau) \circledast \mathrm{d}\mathbf{W}_\tau, \tag{3.6.1}$$

with positive definite diffusion matrix $\mathbf{D}(\mathbf{x}_\tau,\tau) = \boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\boldsymbol{\sigma}^T(\mathbf{x}_\tau,\tau)/2$. Assume that the drift $\mathbf{F}(\mathbf{x}_\tau,\tau)$ and noise amplitude $\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)$ are sufficiently well-behaved for Eq. (3.6.1) to be well-defined with a unique strong solution (e.g., assume that a weak solution exists and $\mathbf{F}$ and $\boldsymbol{\sigma}$ are locally Lipschitz continuous [14]). The anti-Itô convention $\circledast\mathrm{d}\mathbf{W}_\tau = \mathbf{W}_\tau - \mathbf{W}_{\tau-\mathrm{d}\tau}$ is the thermodynamically consistent choice [84, 104, 169], in particular it ensures Boltzmann statistics if the drift $\mathbf{F}(\mathbf{x}_\tau,\tau)$ is such that the system settles into thermodynamic equilibrium [84]. The time-evolution of the probability density $P(\mathbf{x},\tau)$ for any initial density $P(\mathbf{x},\tau=0)$ obeys a Fokker-Planck equation [6, 17]

$$\begin{aligned}\partial_\tau P(\mathbf{x},\tau) &= [-\nabla_\mathbf{x}\cdot\mathbf{F}(\mathbf{x},\tau) + \nabla_\mathbf{x}^T\mathbf{D}(\mathbf{x},\tau)\nabla_\mathbf{x}]P(\mathbf{x},\tau)\\ &\equiv L(\mathbf{x},\tau)P(\mathbf{x},\tau),\end{aligned} \tag{3.6.2}$$

which is equivalent to a continuity equation $[\partial_\tau + \nabla_\mathbf{x}\cdot\hat{\mathbf{j}}(\mathbf{x},\tau)]P(\mathbf{x},\tau) = 0$, with the current operator

$$\hat{\mathbf{j}}(\mathbf{x},\tau) \equiv \mathbf{F}(\mathbf{x},\tau) - \mathbf{D}(\mathbf{x},\tau)\nabla_\mathbf{x} \tag{3.6.3}$$

[9] For more precise statements in the case of multiplicative noise in multidimensional space see Sec. 2.8, in particular the additional condition in Eq. (2.8.6).

that gives the instantaneous current as $\mathbf{j}(\mathbf{x},\tau) = \hat{\mathbf{j}}(\mathbf{x},\tau)P(\mathbf{x},\tau)$.

As a special case of Eq. (3.6.1) we will also study time-homogeneous non-equilibrium steady-state systems, where the stochastic equation of motion reads (curly brackets throughout denote that derivatives only act inside brackets)

$$\mathrm{d}\mathbf{x}_\tau = \left[\mathbf{D}(\mathbf{x}_\tau)\{\nabla \log p_\mathrm{s}\}(\mathbf{x}_\tau) + \frac{\mathbf{j}_\mathrm{s}(\mathbf{x}_\tau)}{p_\mathrm{s}(\mathbf{x}_\tau)}\right]\mathrm{d}\tau + \boldsymbol{\sigma}(\mathbf{x}_\tau) \circledast \mathrm{d}\mathbf{W}_\tau, \tag{3.6.4}$$

where p_s and $\mathbf{j}_\mathrm{s}$ denote the steady-state density and current [84]. Note that (as opposed to [83, 84]) we do *not* assume that the initial distribution is sampled from p_s.

Based on the dynamics defined in Eqs. (3.6.1) or (3.6.4), we consider time-averaged density and current functionals of the trajectories $[\mathbf{x}_\tau]_{0\leq\tau\leq t}$ defined as

$$\begin{aligned}\rho_t^V &= \frac{1}{t}\int_0^t V(\mathbf{x}_\tau,\tau)\mathrm{d}\tau \\ \mathbf{J}_t^U &= \frac{1}{t}\int_{\tau=0}^{\tau=t} U(\mathbf{x}_\tau,\tau)\circ\mathrm{d}\mathbf{x}_\tau,\end{aligned} \tag{3.6.5}$$

with U, V differentiable and square integrable functions and $\circ$ denotes the Stratonovich convention of the stochastic integral. The density functional ρ_t^V measures the time spent in the region $V(\mathbf{x}) \neq 0$, weighted by $V(\mathbf{x})$, while the current $\mathbf{J}_t^U$ functional measures weighted displacements accumulated in U. In particular, for positive V, U that are centered around some point $\mathbf{x}$ and decay on a finite length scale, one can interpret ρ_t^V and $\mathbf{J}_t^U$ as the coarse-grained empirical density and current at $\mathbf{x}$ [83, 84].

In the following, we will derive expressions for the mean values, correlations and fluctuations of these stochastic quantities and illustrate them with an example, thereby generalizing the results in [83, 84] to non-steady-state initial conditions and even systems with explicit time dependence, and thus in particular also without the existence of a steady state.

3.6.3 First moments

Consider overdamped Langevin dynamics as defined in Eq. (3.6.1) starting from an arbitrary initial density $P(\mathbf{x},\tau=0)$. Let $P(\mathbf{x},\tau)$ be the probability density to find the particle at position $\mathbf{x}$ after time τ, i.e., the solution of the Fokker-Planck equation in Eq. (3.6.2). Then the mean value of the density functional in Eq. (3.6.5) is given by

$$\langle\rho_t^V\rangle = \frac{1}{t}\int_0^t \langle V(\mathbf{x}_\tau,\tau)\rangle\mathrm{d}\tau$$

J. Phys. A: Math. Theor. **55** 475001 (2022)

$$= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}^d x V(\mathbf{x},\tau)P(\mathbf{x},\tau). \tag{3.6.6}$$

The mean value of the current is in turn given accordingly by (following closely the approach [84] using that the Itô-$\mathrm{d}\mathbf{W}_\tau$-term vanishes on average, integrating by parts, and using $\mathbf{D} = \mathbf{D}^T$)

$$\begin{aligned}
\langle \mathbf{J}_t^U \rangle &= \frac{1}{t}\int_0^t \langle U(\mathbf{x}_\tau,\tau) \circ \mathrm{d}\mathbf{x}_\tau \rangle \\
&= \frac{1}{t}\int_{\tau=0}^{\tau=t} \langle U(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{x}_\tau \rangle + \frac{1}{t}\int_{\tau=0}^{\tau=t} \frac{1}{2}\langle \mathrm{d}U(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{x}_\tau \rangle_{\mathrm{s}} \\
&= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} P(\mathbf{x},\tau)\times \\
&\qquad \left[U(\mathbf{x},\tau)\mathbf{F}(\mathbf{x},\tau) + \left\{\nabla_{\mathbf{x}}^T \mathbf{D}(\mathbf{x},\tau)\right\} U(\mathbf{x},\tau) + \mathbf{D}(\mathbf{x},\tau)\left\{\nabla_{\mathbf{x}} U(\mathbf{x},\tau)\right\}\right] \\
&= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} P(\mathbf{x},\tau)\left[U(\mathbf{x},\tau)\mathbf{F}(\mathbf{x},\tau) + \nabla_{\mathbf{x}}^T \mathbf{D}(\mathbf{x},\tau) U(\mathbf{x},\tau)\right] \\
&= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} U(\mathbf{x},\tau)\left[\mathbf{F}(\mathbf{x},\tau) - \mathbf{D}(\mathbf{x},\tau)\nabla_{\mathbf{x}}\right] P(\mathbf{x},\tau) \\
&= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} U(\mathbf{x},\tau)\hat{\mathbf{j}}(\mathbf{x},\tau) P(\mathbf{x},\tau) \\
&= \frac{1}{t}\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} U(\mathbf{x},\tau)\mathbf{j}(\mathbf{x},\tau).
\end{aligned} \tag{3.6.7}$$

The expressions Eq. (3.6.6) and (3.6.7) average the probability density and current over the function $U(\mathbf{x},\tau)$ and over time $\tau \in [0,t]$, i.e., one can interpret ρ_t^V and $\mathbf{J}_t^U$ as estimators of space and time averages of $P(\mathbf{x},\tau)$ and $\mathbf{j}(\mathbf{x},\tau)$. Note that for time-homogeneous steady-state dynamics (see Eq. (3.6.4)) these results are unchanged. They only further simplify for dynamics in Eq. (3.6.4) if also the initial condition is sampled from the steady state $P(\mathbf{x},\tau=0) = p_{\mathrm{s}}(\mathbf{x})$, in which case $P(\mathbf{x},\tau) = p_{\mathrm{s}}(\mathbf{x})$ and $\mathbf{j}(\mathbf{x},\tau) = \mathbf{j}_{\mathrm{s}}(\mathbf{x})$ implies that $\langle \rho_t^V \rangle$ and $\langle \mathbf{J}_t^U \rangle$ become independent of t.

3.6.4 Correlations and fluctuations

We now derive second moments and linear correlations of the time-averaged density and current in Eq. (3.6.5). The derivations for higher moments of currents are more involved than the first moments, but as in [84] we solve the complications in the derivation by means of a single Lemma derived in the following subsection. Note that one could alternatively derive the following results using a Feynman-Kac approach (and optionally functional calculus) by appropriately generalizing the approach in Sec. 3.7.

J. Phys. A: Math. Theor. 55 475001 (2022)

Lemma

In the derivation of expressions for fluctuations and correlations of the time-averaged quantities, we must evaluate correlations of noise increments $\mathrm{d}\mathbf{W}_\tau$ and functions of $\mathbf{x}_{\tau'}$. Correlations for $\tau' \leq \tau$ vanish by the properties of the Wiener process. Conversely, correlations for $\tau' > \tau$ are non-trivial. This problem was solved for steady-state dynamics in [84] and via Doob conditioning [104, 107, 200] for general time-homogeneous Langevin systems in the Supplemental Material of [133]. We now generalize the direct approach from [84] to overdamped Langevin systems with explicit time dependence.

Consider the k-th component $[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau]_k$ of a noise increment in an expectation value $\langle f(\mathbf{x}_\tau,\mathbf{x}_{\tau'},\tau,\tau')\,[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau]_k\rangle$ with some (differentiable, square integrable) function f. For $\tau' \leq \tau$ this term vanishes due to vanishing correlations and zero mean of $\mathrm{d}\mathbf{W}_\tau$. Now consider $\tau' > \tau$.

Given a point $\mathbf{x}_\tau = \mathbf{x}$ and writing $\boldsymbol{\varepsilon} \equiv \boldsymbol{\sigma}(\mathbf{x},\tau)\mathrm{d}\mathbf{W}_\tau$, the equation of motion (3.6.1) rewritten in Itô form (writing out the anti-Itô correction term) implies a displacement $\mathrm{d}\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon}) = [\mathbf{F}(\mathbf{x},\tau) + \nabla_\mathbf{x}^T\mathbf{D}(\mathbf{x},\tau)]\mathrm{d}\tau + \boldsymbol{\varepsilon}$. With this we can write the expectation $\langle f(\cdots)\,[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau]_k\rangle$ as ε_k integrated over the probability to be at points $\mathbf{x}, \mathbf{x} + \mathrm{d}\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon}), \mathbf{y}$ at times $\tau < \tau + \mathrm{d}\tau < \tau'$, i.e., (with joint density $P(\mathbf{y},\tau';\mathbf{x},\tau)$ and conditional density $P(\mathbf{y},\tau'|\mathbf{x},\tau) \equiv P(\mathbf{y},\tau';\mathbf{x},\tau)/P(\mathbf{x},\tau)$; we write $\mathbb{1}_{\tau<\tau'}$ for 1 if $\tau < \tau'$ and 0 else)

$$\begin{aligned}&\langle f(\mathbf{x}_\tau,\mathbf{x}_{\tau'},\tau,\tau')\,[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau]_k\rangle \\ &= \mathbb{1}_{\tau<\tau'}\int \mathrm{d}\mathbf{x}\int \mathrm{d}\mathbf{y} f(\mathbf{x},\mathbf{y},\tau,\tau')\int \mathrm{d}\boldsymbol{\varepsilon}\,\mathbb{P}(\boldsymbol{\varepsilon})\varepsilon_k P(\mathbf{y},\tau'|\mathbf{x}+\mathrm{d}\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon}),\tau+\mathrm{d}\tau)P(\mathbf{x},\tau)\,,\end{aligned} \tag{3.6.8}$$

where the probability $\mathbb{P}(\boldsymbol{\varepsilon})$ is given by a Gaussian distribution with zero mean and covariance matrix $2\mathbf{D}(\mathbf{x},\tau)\mathrm{d}\tau$. Since this distribution is symmetric around $\mathbf{0}$, only terms with even powers of the components of $\boldsymbol{\varepsilon}$ survive the $\mathrm{d}\boldsymbol{\varepsilon}\mathbb{P}(\boldsymbol{\varepsilon})$-integration. Note that

$$P(\mathbf{y},\tau'|\mathbf{x}+\mathrm{d}\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon}),\tau+\mathrm{d}\tau) \overset{\mathrm{d}\tau\to 0}{\longrightarrow} [1+\mathrm{d}\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon})\cdot\nabla_\mathbf{x}]P(\mathbf{y},\tau'|\mathbf{x},\tau)+\mathcal{O}(\mathrm{d}\tau), \tag{3.6.9}$$

and we can neglect the higher orders $\mathcal{O}(\mathrm{d}\tau)$ since $\varepsilon_k\mathcal{O}(\mathrm{d}\tau) = \mathcal{O}(\mathrm{d}\tau^{3/2})$, which (unlike $\varepsilon_k\mathcal{O}(\mathrm{d}\tau^{1/2})$) will still give zero after integration in τ. From the zeroth and first order contribution, we see that the only even power of the components of $\boldsymbol{\varepsilon}$ in the above integration gives

$$\langle f(\mathbf{x}_\tau,\mathbf{x}_{\tau'},\tau,\tau')\,[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau]_k\rangle$$

J. Phys. A: Math. Theor. 55 475001 (2022)

$$= \mathbb{1}_{\tau<\tau'} \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} f(\mathbf{x},\mathbf{y},\tau,\tau') P(\mathbf{x},\tau) \int \mathrm{d}\boldsymbol{\varepsilon}\, \mathbb{P}(\boldsymbol{\varepsilon}) \varepsilon_k \boldsymbol{\varepsilon} \cdot \nabla_{\mathbf{x}} P(\mathbf{y},\tau'|\mathbf{x},\tau), \tag{3.6.10}$$

which, using $\int \mathrm{d}\varepsilon \mathbb{P}(\boldsymbol{\varepsilon})\varepsilon_k \varepsilon_j = 2D_{kj}(\mathbf{x},\tau)\mathrm{d}\tau$, yields the result for $\tau < \tau'$

$$\begin{aligned} &\langle f(\mathbf{x}_\tau, \mathbf{x}_{\tau'}, \tau, \tau') \left[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau\right]_k \rangle \\ &= \mathbb{1}_{\tau<\tau'} \mathrm{d}\tau \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} P(\mathbf{x},\tau) f(\mathbf{x},\mathbf{y},\tau,\tau') \left[2\mathbf{D}(\mathbf{x},\tau)\nabla_{\mathbf{x}} P(\mathbf{y},\tau'|\mathbf{x},\tau)\right]_k . \end{aligned} \tag{3.6.11}$$

For scalar products with vector valued functions $\mathbf{f}$ the result (3.6.11) can be summed over components f_k to obtain

$$\begin{aligned} &\langle \mathbf{f}(\mathbf{x}_\tau, \mathbf{x}_{\tau'}, \tau, \tau') \cdot \boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau \rangle \\ &= \mathbb{1}_{\tau<\tau'} \mathrm{d}\tau \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} P(\mathbf{x},\tau) \mathbf{f}(\mathbf{x},\mathbf{y},\tau,\tau') \cdot 2\mathbf{D}(\mathbf{x},\tau)\nabla_{\mathbf{x}} P(\mathbf{y},\tau'|\mathbf{x},\tau). \end{aligned} \tag{3.6.12}$$

Eq. (3.6.12) is the central result of this work that allows us to directly deduce expressions for fluctuations and correlations of densities and currents. Upon integrating by parts and using symmetry $\mathbf{D}^T(\mathbf{x},\tau) = \mathbf{D}(\mathbf{x},\tau)$ Eq. (3.6.12) could also be rewritten as

$$\begin{aligned} &\langle \mathbf{f}(\mathbf{x}_\tau, \mathbf{x}_{\tau'}, \tau, \tau') \cdot \boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau \rangle \\ &= -\mathbb{1}_{\tau<\tau'} \mathrm{d}\tau \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} P(\mathbf{y},\tau'|\mathbf{x},\tau) \nabla_{\mathbf{x}} \cdot [P(\mathbf{x},\tau) 2\mathbf{D}(\mathbf{x},\tau)\mathbf{f}(\mathbf{x},\mathbf{y},\tau,\tau')]. \end{aligned} \tag{3.6.13}$$

Fluctuations and correlations of densities and currents

Following the developed approach and generalizing the results obtained in [84] we now derive expressions for fluctuations and correlations of densities and currents for arbitrary initial conditions.

For two time-averaged densities ρ_t^U, ρ_t^V, the covariance (variance for $U = V$) is given by

$$\begin{aligned} &\langle \rho_t^U \rho_t^V \rangle - \langle \rho_t^U \rangle \langle \rho_t^V \rangle \\ &= t^{-2} \int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' \left[\langle U(\mathbf{x}_\tau,\tau) V(\mathbf{x}_{\tau'},\tau') - \langle U(\mathbf{x}_\tau,\tau)\rangle \langle V(\mathbf{x}_{\tau'},\tau')\rangle\right] \\ &= t^{-2} \int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} U(\mathbf{x},\tau) V(\mathbf{y},\tau') \left[P(\mathbf{x},\tau;\mathbf{y},\tau') - P(\mathbf{x},\tau)P(\mathbf{y},\tau')\right] . \end{aligned} \tag{3.6.14}$$

Note that this result can be interpreted as correlations caused by differences of $P(\mathbf{x},\tau;\mathbf{y},\tau')$ and $P(\mathbf{x},\tau)P(\mathbf{y},\tau')$, averaged over time and over functions U, V. More precisely, the two-point function $P(\mathbf{x},\tau;\mathbf{y},\tau')$ can be understood to be characterized by all paths with $\mathbf{x}_\tau = \mathbf{x}$

J. Phys. A: Math. Theor. 55 475001 (2022)

and $\mathbf{x}_{\tau'} = \mathbf{y}$. For further interpretation, in particular for the case of steady-state dynamics, see [83, 84].

For the correlation of $\mathbf{J}_t^U$ and ρ_t^V we first consider the expectation of the product and carry out the same steps as in Eq. (3.6.7),

$$\begin{aligned} t^2\langle \mathbf{J}_t^U \rho_t^V\rangle &= \int_0^t \mathrm{d}\tau' \int_{\tau=0}^{\tau=t} \langle U(\mathbf{x}_\tau,\tau)\circ \mathrm{d}\mathbf{x}_\tau V(\mathbf{x}_{\tau'},\tau')\rangle \\ &= \int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} U(\mathbf{x},\tau)V(\mathbf{y},\tau')\hat{\mathbf{j}}(\mathbf{x},\tau)P(\mathbf{y},\tau';\mathbf{x},\tau) \\ &\quad + \int_0^t \mathrm{d}\tau' \int_{\tau=0}^{\tau=t} \langle U(\mathbf{x}_\tau,\tau)\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau V(\mathbf{x}_{\tau'},\tau')\rangle. \end{aligned} \tag{3.6.15}$$

Comparing with the calculation in Eq. (3.6.7), the noise term no longer vanishes since terms with $\tau < \tau'$ give non-trivial correlations according to Eq. (3.6.11), which in turn gives

$$\begin{aligned} &\int_0^t \mathrm{d}\tau' \int_{\tau=0}^{\tau=t} \langle U(\mathbf{x}_\tau,\tau)\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau V(\mathbf{x}_{\tau'},\tau')\rangle = \int_0^t \mathrm{d}\tau' \int_0^t \mathrm{d}\tau \mathbb{1}_{\tau<\tau'}\times \\ &\int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} U(\mathbf{x},\tau)V(\mathbf{y},\tau') \left[2P(\mathbf{x},\tau)\mathbf{D}(\mathbf{x},\tau)\nabla_{\mathbf{x}}P(\mathbf{x},\tau)^{-1}\right] P(\mathbf{y},\tau';\mathbf{x},\tau), \end{aligned} \tag{3.6.16}$$

where we rewrote $P(\mathbf{y},\tau'|\mathbf{x},\tau) = P(\mathbf{x},\tau)^{-1}P(\mathbf{y},\tau';\mathbf{x},\tau)$. Introducing the adapted current operator

$$\hat{\mathbf{j}}^\ddagger(\mathbf{x},\tau) \equiv \hat{\mathbf{j}}(\mathbf{x},\tau) + 2P(\mathbf{x},\tau)\mathbf{D}(\mathbf{x},\tau)\nabla_{\mathbf{x}}P(\mathbf{x},\tau)^{-1}, \tag{3.6.17}$$

we thus obtain from Eq. (3.6.15) an expression for the current-density correlation that reads

$$\begin{aligned} \langle \mathbf{J}_t^U \rho_t^V\rangle - \langle \mathbf{J}_t^U\rangle\langle \rho_t^V\rangle = t^{-2}\int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} U(\mathbf{x},\tau)V(\mathbf{y},\tau')\times \\ \left[\mathbb{1}_{\tau>\tau'}\hat{\mathbf{j}}(\mathbf{x},\tau) + \mathbb{1}_{\tau<\tau'}\hat{\mathbf{j}}^\ddagger(\mathbf{x},\tau)\right]\left[P(\mathbf{y},\tau';\mathbf{x},\tau) - P(\mathbf{x},\tau)P(\mathbf{y},\tau')\right]. \end{aligned} \tag{3.6.18}$$

Note that to write the expression more compactly, we used that $\hat{\mathbf{j}}(\mathbf{x},\tau)P(\mathbf{x},\tau) = \hat{\mathbf{j}}^\ddagger(\mathbf{x},\tau)P(\mathbf{x},\tau) = \mathbf{j}(\mathbf{x},\tau)$. To write the result in a more symmetric way, and since the difference vanishes (as a set of measure 0 in two-dimensional integration), we wrote $\mathbb{1}_{\tau>\tau'}$ instead of $\mathbb{1}_{\tau\geq\tau'}$.

The expression (3.6.18) is a natural generalization of Eq. (3.6.14) with the current operators $\hat{\mathbf{j}}, \hat{\mathbf{j}}^\ddagger$ appearing. Recall that $\hat{\mathbf{j}}$ is the current operator entering the Fokker-Planck equation, see Eqs. (3.6.2)-(3.6.3). The adapted operator $\hat{\mathbf{j}}^\ddagger$ defined in Eq. (3.6.17) ac-

J. Phys. A: Math. Theor. 55 475001 (2022)

counts for the fact that trajectories contributing to $P(\mathbf{y},\tau';\mathbf{x},\tau)$ that first visit $\mathbf{x}$ and later $\mathbf{y}$ (i.e., $\tau<\tau'$) have, compared to the Fokker-Planck evolution, altered statistics, since displacements at $\mathbf{x}$ correlate with probabilities of reaching $\mathbf{y}$ later. For the particular case of steady-state systems, the special case of the correlation result (3.6.18) and the adapted current operator were discussed in detail, and explained using a generalized time-reversal symmetry, in [83, 84]. Note that for the case of time-homogeneous dynamics (in particular steady-state dynamics defined in Eq. (3.6.4)), the Fokker-Planck current operator Eq. (3.6.3) does not have an explicit time dependence such that $\hat{\mathbf{j}}(\mathbf{x},\tau)$ in Eq. (3.6.18) simplifies to $\hat{\mathbf{j}}(\mathbf{x})$. However, the adapted current operator $\hat{\mathbf{j}}^{\ddagger}(\mathbf{x},\tau)$ defined in Eq. (3.6.17) retains explicit time dependence even for time-homogeneous dynamics. Only in the case of steady-state systems with steady-state initial conditions (where $P(\mathbf{x},\tau)=p_{\mathrm{s}}(\mathbf{x})$ for all τ) $\hat{\mathbf{j}}^{\ddagger}$ has no explicit time dependence, and simplifies to the negative $\hat{\mathbf{j}}$ with inverted steady-state current $\mathbf{j}_{\mathrm{s}}\to-\mathbf{j}_{\mathrm{s}}$ [83, 84].

Covariances of components m,n of time-integrated currents $J^U_{t,m}$ and $J^V_{t,n}$ can be obtained analogously by considering

$$t^2\langle J^U_{t,m}J^V_{t,n}\rangle=\int_{\tau'=0}^{\tau'=t}\int_{\tau=0}^{\tau=t}\langle U(\mathbf{x}_\tau,\tau)\circ\mathrm{d}x^m_\tau V(\mathbf{x}_{\tau'},\tau')\circ\mathrm{d}x^n_{\tau'}\rangle, \tag{3.6.19}$$

where both $\circ\mathrm{d}\mathbf{x}_t$ increments split into $\mathrm{d}t$ and $\mathrm{d}\mathbf{W}_t$ terms. The $\mathrm{d}\tau\mathrm{d}\tau'$ terms give rise to the current operator $\hat{\mathbf{j}}$ as in Eqs. (3.6.7), (3.6.15), but now its components $\hat{\mathrm{j}}_m(\mathbf{x},\tau)$ and $\hat{\mathrm{j}}_n(\mathbf{y},\tau')$ appear. The $\mathrm{d}W_\tau\mathrm{d}W_{\tau'}$ term yields (by Itô's isometry, i.e., "delta-correlated white noise")

$$\begin{aligned}&\int_{\tau'=0}^{\tau'=t}\int_{\tau=0}^{\tau=t}\langle U(\mathbf{x}_\tau,\tau)\left[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau\right]_m V(\mathbf{x}_{\tau'},\tau')\left[\boldsymbol{\sigma}(\mathbf{x}_{\tau'},\tau')\mathrm{d}\mathbf{W}_{\tau'}\right]_n\rangle\\&=\int_0^t\langle U(\mathbf{x}_\tau,\tau)V(\mathbf{x}_\tau,\tau)2D_{mn}(\mathbf{x}_\tau)\rangle\mathrm{d}\tau\\&=2\int_0^t\mathrm{d}\tau\int\mathrm{d}\mathbf{x}U(\mathbf{x},\tau)V(\mathbf{x},\tau)D_{mn}(\mathbf{x})P(\mathbf{x},\tau).\end{aligned} \tag{3.6.20}$$

The mixed term $\mathrm{d}\tau'\mathrm{d}\mathbf{W}_\tau$ (and equivalently $\mathrm{d}\tau\mathrm{d}\mathbf{W}_{\tau'}$) in Eq. (3.6.19) according to calculations as in Eq. (3.6.7) and using Eq. (3.6.11) gives

$$\begin{aligned}&\int_0^t\mathrm{d}\tau'\int_{\tau=0}^{\tau=t}\langle U(\mathbf{x}_\tau,\tau)\left[\boldsymbol{\sigma}(\mathbf{x}_\tau,\tau)\mathrm{d}\mathbf{W}_\tau\right]_m\left[V(\mathbf{x}_{\tau'},\tau')\mathbf{F}(\mathbf{x}_{\tau'},\tau')+\{\nabla\mathbf{D}V\}(\mathbf{x}_{\tau'},\tau')\right]_n\rangle\\&=\mathbb{1}_{\tau<\tau'}\int_0^t\mathrm{d}\tau\int_0^t\mathrm{d}\tau'\int\mathrm{d}\mathbf{x}\int\mathrm{d}\mathbf{y}U(\mathbf{x},\tau)V(\mathbf{y},\tau')\hat{\mathrm{j}}_n(\mathbf{y},\tau')\times\\&\qquad\left[2P(\mathbf{x},\tau)\nabla_{\mathbf{x}}P(\mathbf{x},\tau)^{-1}\right]_m P(\mathbf{y},\tau';\mathbf{x},\tau).\end{aligned} \tag{3.6.21}$$

J. Phys. A: Math. Theor. 55 475001 (2022)

Collecting all terms and using the and notation $\hat{\mathrm{j}}^{\ddagger}_m$ for the components of $\hat{\mathbf{j}}^{\ddagger}$ in Eq. (3.6.17), we obtain for Eq. (3.6.19)

$$t^2\langle J^U_{t,m}J^V_{t,n}\rangle = 2\int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} U(\mathbf{x},\tau)V(\mathbf{x},\tau)D_{mn}(\mathbf{x},\tau)P(\mathbf{x},\tau) + \int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} \times$$
$$U(\mathbf{x},\tau)V(\mathbf{y},\tau')\left[\mathbb{1}_{\tau<\tau'}\hat{\mathrm{j}}^{\ddagger}_m(\mathbf{x},\tau)\hat{\mathrm{j}}_n(\mathbf{y},\tau') + \mathbb{1}_{\tau>\tau'}\hat{\mathrm{j}}_m(\mathbf{x},\tau)\hat{\mathrm{j}}^{\ddagger}_n(\mathbf{y},\tau')\right]P(\mathbf{y},\tau';\mathbf{x},\tau). \quad (3.6.22)$$

From the derivation one sees that the first term is the $\tau = \tau'$-contribution (see also [84]). This is the natural generalization of the results in Eqs. (3.6.14) and (3.6.18), with the interpretation of non-trivial displacements (and thus $\hat{\mathbf{j}}^{\ddagger}$ instead of $\hat{\mathbf{j}}$) for currents evaluated at earlier times (see above and [83,84]). As before, for time-homogeneous dynamics $\hat{\mathbf{j}}(\mathbf{x},\tau)$ simplifies to $\hat{\mathbf{j}}(\mathbf{x})$ and in the special case of steady-state dynamics (see Eq. (3.6.4)) with steady-state initial conditions, $\hat{\mathbf{j}}^{\ddagger}(\mathbf{x},\tau)$ simplifies to $\hat{\mathbf{j}}^{\ddagger}(\mathbf{x})$. This special case was discussed and explained using generalized time-reversal symmetry in [83, 84].

3.6.5 Example

To present a concrete minimal example, we consider a two-dimensional harmonically confined overdamped diffusion in a rotational flow (i.e., an irreversible Ornstein-Uhlenbeck process)

$$\mathrm{d}\mathbf{x}_t = -\begin{bmatrix} 1 & -\Omega \\ \Omega & 1 \end{bmatrix}\mathbf{x}\mathrm{d}t + \sqrt{2}\mathrm{d}\mathbf{W}_t. \quad (3.6.23)$$

Assuming that the initial density $P(\mathbf{x},\tau=0)$ is Gaussian, the solution $P(\mathbf{x},\tau)$ of the Fokker-Planck equation corresponding to Eq. (3.6.23) is well known to be a Gaussian density for all $\tau \geq 0$ (see, e.g., [6]). We choose U to be a two-dimensional Gaussian centered at $\mathbf{z}$ with width h, i.e.,

$$U_{\mathbf{z}}(\mathbf{x}) = \frac{1}{2\pi h^2}\exp\left[-\frac{(\mathbf{x}-\mathbf{z})^2}{2h^2}\right]. \quad (3.6.24)$$

Due to the Gaussianity of $P(\mathbf{x},\tau)$ and $U_{\mathbf{z}}(\mathbf{x})$, all spatial integrals entering the results Eqs. (3.6.14), (3.6.18) and (3.6.22) can be performed analytically, e.g., using the computer algebra system SymPy [185] (as outlined in the Supplemental Material of [83]). The two remaining time-integrals are computed numerically. For simplicity, we only consider the (non-steady-state) initial condition in a point, i.e., $P(\mathbf{x},\tau=0) = \delta(\mathbf{x}-\mathbf{x}_0)$. For this

J. Phys. A: Math. Theor. **55** 475001 (2022)

initial condition, via a left-right decomposition for the process Eq. (3.6.23) (see, e.g., [17]) or by solving the Lyapunov equation, we have the time-dependent density

$$P(\mathbf{x},\tau) = \frac{1}{2\pi(1-\mathrm{e}^{-2\tau})} \exp\left[\frac{-\left(\mathbf{x} - \mathrm{e}^{-\tau}\begin{bmatrix}\cos(\Omega\tau) & \sin(\Omega\tau)\\ -\sin(\Omega\tau) & \cos(\Omega\tau)\end{bmatrix}\mathbf{x}_0\right)^2}{2(1-\mathrm{e}^{-2\tau})}\right], \tag{3.6.25}$$

i.e., the mean value $\langle \mathbf{x}_\tau\rangle = \mathrm{e}^{-\tau}\begin{bmatrix}\cos(\Omega\tau) & \sin(\Omega\tau)\\ -\sin(\Omega\tau) & \cos(\Omega\tau)\end{bmatrix}\mathbf{x}_0$ moves on a spiral shape towards the center. The case $\Omega = 0$ corresponds to the equilibrium process, i.e., harmonically confined overdamped diffusion without rotational flow.

For this example, we compute the density-current correlation vector as in Eq. (3.6.18),

$$\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0) \equiv \langle \mathbf{J}_t^{U_\mathbf{z}}\rho_t^{U_\mathbf{z}}\rangle_{\mathbf{x}_0} - \langle \mathbf{J}_t^{U_\mathbf{z}}\rangle_{\mathbf{x}_0}\langle\rho_t^{U_\mathbf{z}}\rangle_{\mathbf{x}_0} = t^{-2}\int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{y} \times$$
$$U_\mathbf{z}(\mathbf{x})U_\mathbf{z}(\mathbf{y})\left[\mathbb{1}_{\tau>\tau'}\hat{\mathbf{j}}(\mathbf{x}) + \mathbb{1}_{\tau<\tau'}\hat{\mathbf{j}}^\ddagger(\mathbf{x},\tau)\right]\left[P(\mathbf{y},\tau';\mathbf{x},\tau) - P(\mathbf{x},\tau)P(\mathbf{y},\tau')\right]. \tag{3.6.26}$$

with Gaussian $U_\mathbf{z}$ as in Eq. (3.6.24). In Fig. 3.6.1 we show the time evolution and spatial dependence of this correlation vector. For long times without driving $\Omega = 0$, we see that $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0) \to 0$. This corresponds to the limit when the initial condition is forgotten, i.e., for long times $\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0)$ approaches the result of $\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t)$ for steady-state initial conditions where in equilibrium ($\Omega = 0$) we have $P(\mathbf{y},\tau';\mathbf{x},\tau) = P(\mathbf{x},\tau';\mathbf{y},\tau)$ (time-reversal symmetry) and $\hat{\mathbf{j}}^\ddagger(\mathbf{x}) = -\hat{\mathbf{j}}(\mathbf{x})$ [84], implying $\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t) = \mathbf{0}$ at all $\mathbf{z}$. In the case $\Omega \neq 0$, the correlation $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0)$ becomes constant for long times, where $\mathbf{C}_{\mathbf{j}\rho} \propto t^{-1}$ represents the large-deviation limit of the correlation result, which agrees with the large-deviation limit for the process starting in steady-state initial conditions [84]. This has a spatial dependence similar to the steady-state current $\mathbf{j}_\mathrm{s}(\mathbf{z})$ but averaged over the Gaussian $U_\mathbf{z}$. By comparison with the color gradient we see in all panels in Fig. 3.6.1, as expected, that large values of the correlation can only occur at points that are visited for a significant amount of time, i.e., with not too small $\langle\rho_t^{U_\mathbf{z}}\rangle$.

In addition to the qualitative behavior shown in Fig. 3.6.1, we present a quantitative evaluation of the correlation result multiplied by time, $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0)$, for a single $\mathbf{z}$ in Fig. 3.6.2. Simulations shown in Fig. 3.6.2 confirm the theoretical result in Eq. (3.6.18) (re-stated in Eq. (3.6.26)). As mentioned above, this result approaches the large-deviation limit for long times. Moreover, for long times the initial condition will become irrelevant, i.e., $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0)$ approaches the result for $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t)$ for steady-state initial conditions [84].

J. Phys. A: Math. Theor. 55 475001 (2022)

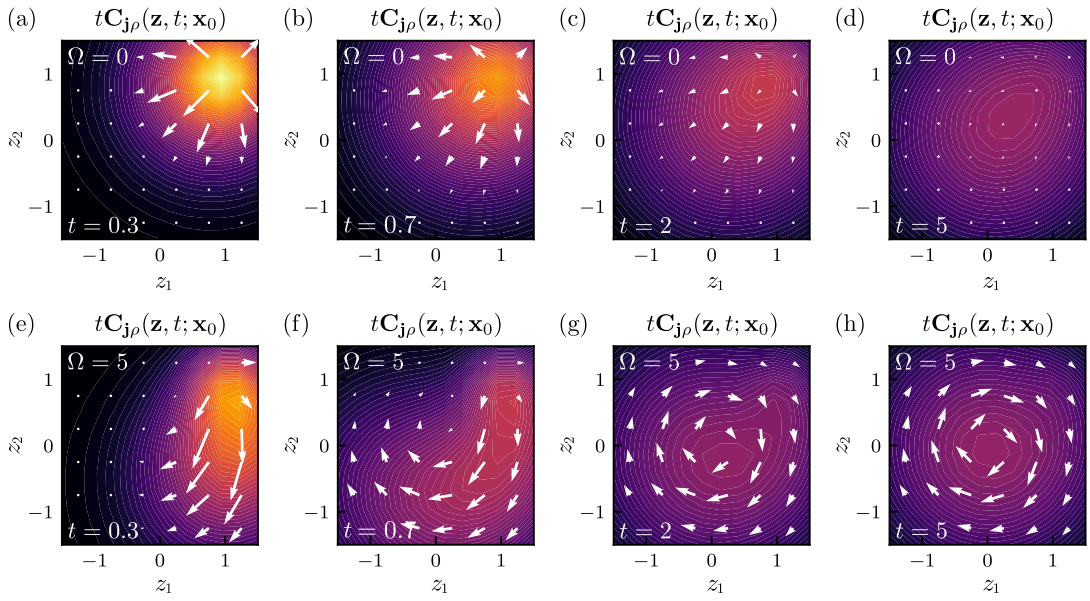

Figure 3.6.1: White arrows depict the correlation result multiplied by time, $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0)$ as in Eq. (3.6.18) with $\mathbf{x}_0 = (1,1)^T$ for the process in Eq. (3.6.23) with $\Omega = 0$ in (a-d), $\Omega = 5$ in (e-h), and $U_\mathbf{z}$ as in Eq. (3.6.24). The position $\mathbf{z} = (z_1, z_2)^T$ around which the correlation is evaluated varies along the respective axes. The color gradient depicts the mean time-averaged density $\langle \rho_t^{U_\mathbf{z}} \rangle$, i.e., the time spent around $\mathbf{z}$ weighted by $U_\mathbf{z}$. Time increases from left (a,e) to right (d,h), $t = 0.3,\ 0.7,\ 2,\ 5$

First note that, due to the time-integration, deviations for short times are only slowly 'forgotten' with order t^{-1} (instead of exponentially fast with some Poincaré timescale). Interestingly, we see in Fig. 3.6.2a that for substantial coarse-graining (i.e., rather large $h = 0.5$ in $U_\mathbf{z}$ in Eq. (3.6.24)), the result for $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z},t;\mathbf{x}_0)$ starting in a point only approaches the corresponding value for steady-state initial condition (green curve) in the large deviation regime (black line), but *not* before. Going to smaller coarse graining $h = 0.15$ in Fig. 3.6.2b, we see that the process starting in the center $\mathbf{x}_0 = (0,0)^T$ (blue line) features the arguably more intuitive behavior, by first approaching the steady-state (green line) and later the large deviation result (black line). However, for a different initial condition (see orange line) the steady state curve is again only approached in the large deviation limit. This highlights the long-lasting and non-trivial effects of the time-integration and underscores why interpreting time-average observables, in particular those involving currents, remains challenging.

3.6.6 Conclusion

To summarize, we presented a new Lemma (3.6.12) that enabled us to derive results Eqs. (3.6.14), (3.6.18) and (3.6.22) for correlations and fluctuations of the time-averaged

J. Phys. A: Math. Theor. **55** 475001 (2022)

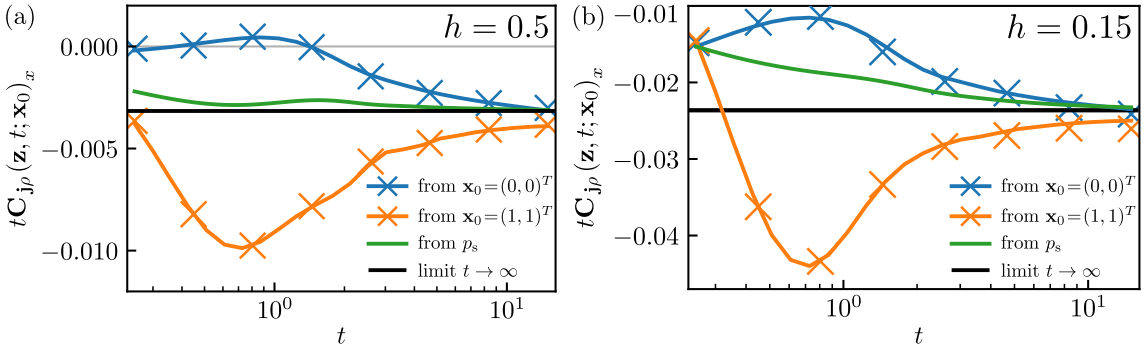

Figure 3.6.2: Quantitative depiction of the time dependence of the x-component of the current-density correlation $t\mathbf{C}_{\mathbf{j}\rho}(\mathbf{z}, t; \mathbf{x}_0)_x$ with $\mathbf{z} = (0, -0.2)^T$ for the process in Eq. (3.6.23) with $\Omega = 3$ and $U_{\mathbf{z}}$ as in Eq. (3.6.24) with (a) $h = 0.5$ and (b) $h = 0.15$ for different initial conditions (colors). The new analytical result (blue and orange lines; Eq. (3.6.18)) is confirmed by simulations (crosses; for each t, (a) 10^5 and (b) 10^6 trajectories with 10^3 time-steps each were simulated according to the stochastic Euler algorithm). For $t \to \infty$, irrespective of the initial condition, all result approach the same large-deviation limit.

density and current Eq. (3.6.5) for general Langevin dynamics defined in Eq. (3.6.1) with general initial conditions. This generalization of the recent results derived for non-equilibrium steady states [83, 84] may improve the understanding of inference of densities and currents with the estimators ρ_t^V and $\mathbf{J}_t^U$ (in particular in connection with the notion of coarse graining [83, 84]) in cases where the dynamics does *not* evolve from the steady-state, or is not time-homogeneous. Importantly, the strategy of inferring dissipation from the current variance (see Eq. (3.6.22)) via the thermodynamic uncertainty relation (TUR) [47, 48, 67, 132] remains valid. Generalized versions of the TUR, e.g., for general initial conditions [126] or time-dependent dynamics [49], already exist. A recently improved version of the TUR that includes current-density correlations (see Eq. (3.6.18)) is, however, so far only available for steady-state systems with steady-state initial conditions [134]. Notably, as we will show in a forthcoming publication, Lemma (3.6.12) allows the correlation-TUR to also be proved for transient dynamics.

Acknowledgments.— Financial support from Studienstiftung des Deutschen Volkes (to C. D.) and the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2 (to A. G.) is gratefully acknowledged.

J. Phys. A: Math. Theor. **55** 475001 (2022)

3.7 Feynman-Kac theory of time-integrated functionals: Itô versus functional calculus **(J. Phys. A 2023)**

This section is a slightly adapted version of the publication C. Dieball and A. Godec, "Feynman-Kac theory of time-integrated functionals: Itô versus functional calculus", *J. Phys. A: Math. Theor.* **56**, 155002 (2023).

My personal contribution as first author of the publication was in performing calculations, analyzing the results, and co-writing the manuscript.

Feynman-Kac theory of time-integrated functionals: Itô versus functional calculus

Cai Dieball and Aljaž Godec

Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

Abstract

The fluctuations of dynamical functionals such as the empirical density and current as well as heat, work and generalized currents in stochastic thermodynamics are usually studied within the Feynman-Kac tilting formalism, which in the Physics literature is typically derived by some form of Kramers-Moyal expansion, or in the Mathematical literature via the Cameron-Martin-Girsanov approach. Here we derive the Feynman-Kac theory for general additive dynamical functionals directly via Itô calculus and via functional calculus, where the latter results in fact appears to be new. Using Dyson series, we then independently recapitulate recent results on steady-state (co)variances of general additive dynamical functionals derived recently in Dieball and Godec (2022 *Phys. Rev. Lett.* **129** 140601) and Dieball and Godec (2022 *Phys. Rev. Res.* **4** 033243) (see Secs. 3.4 and 3.5, respectively). We hope for our work to put the different approaches to the statistics of dynamical functionals employed in the field on a common footing, and to illustrate more easily accessible ways to the tilting formalism.

J. Phys. A: Math. Theor. **56** 155002 (2023)

3.7.1 Introduction

Dynamical functionals and diverse path-based observables [60,157,215–220], such as local and occupation times (also known as the "empirical density") [33,36,40,42,135–139] as well as diverse time-integrated and time-averaged currents [34,43–45,83,84,107,110,133,134,141–144] are central to "time-average statistical mechanics" [37,41,130], large deviation theory (see, e.g., [43,107,110,143,204]), macroscopic fluctuation theory [109,159,173], and path-wise, stochastic thermodynamics [28,49,54,104,133,134,152,162].

Several techniques are available for the study of dynamical functionals, presumably best known is the Lie-Trotter-Kato formalism [135,221] that was employed by Kac in his seminal work [42]. The techniques typically employed in physics rely on an analogy to quantum mechanical problems (see, e.g., [33]) or assume some form of the Kramers-Moyal

expansion [137, 138, 222, 223] (see also interesting generalizations to anomalous dynamics [40, 139]).

Deriving Feynman-Kac theory [42] of such additive functionals amounts to obtaining a "tilted" generator which generates the time-evolution of the observables under consideration. The tilted evolution operator can be obtained using the Cameron-Martin-Girsanov theorem [224,225] — a well-known technical theorem often employed in the Mathematical Physics literature [44, 107, 142].

In this paper we develop the Itô [14, 41] and functional calculus [226, 227] approaches to Feynman-Kac theory, whereby we focus on the methodology and accessibility for readers that are unfamiliar with the Cameron-Martin-Girsanov approach to "tilting". We thereby hope to provide two accessible alternative (but equivalent) ways to obtaining the tilted generator. While the Itô approach already exists (see, e.g., [41] for the empirical density), our functional calculus approach is a generalization of the pedagogical work of Fox [226, 227] and is aimed towards readers who prefer to avoid Itô calculus. Since both methods are equivalent, they yield the same tilted generator. This generator is subsequently used to rederive recent results on the statistics of time-integrated densities and currents obtained in References [83, 84] using a different, more direct, stochastic calculus approach that avoids tilting. In particular, these results illustrate the use of the tilted generator to derive the statistics of time-integrated observables for finite times, i.e., extending beyond large deviation theory.

The outline of the paper is as follows. In Sec. 3.7.2 we provide the mathematical setup of the problem. In Sec. 3.7.2 we derive the Feynman-Kac equation for a general dynamical functional of diffusion processes using Itô calculus. By generalizing the approach by Fox [226, 227] we derive in Sec. 3.7.2 the Feynman-Kac equation using functional calculus. In Sec. 3.7.3 we apply the formalism to compute steady-state (co)variances of general dynamical functionals using a Dyson-series approach. We conclude with a brief perspective.

3.7.2 Tilted Generator

In this section, we first introduce the considered stochastic dynamics and define what we call "dynamical functionals". Subsequently, we derive the tilted generator (i.e., the operator generating the time-evolution of time-integrated functionals) based on Itô calculus, and finally equivalently also via functional calculus.

J. Phys. A: Math. Theor. **56** 155002 (2023)

J. Phys. A: Math. Theor. 56 155002 (2023)

Set-Up

We consider overdamped stochastic motion in d-dimensional space described by the stochastic differential equation

$$\mathrm{d}\mathbf{x}_t = \mathbf{F}(\mathbf{x}_t)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t, \tag{3.7.1}$$

where $\mathrm{d}\mathbf{W}_t$ is denotes increment of the Wiener process [14]. The corresponding diffusion constant is $\mathbf{D} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T/2$. For simplicity, we stick to additive noise whereas all present results generalize to multiplicative noise $\mathbf{D}(\mathbf{x})$ as described in [84]. In the physics literature, Eq. (3.7.1) is typically written in the form of a Langevin equation

$$\dot{\mathbf{x}}_t = \mathbf{F}(\mathbf{x}_t) + \mathbf{f}(t), \tag{3.7.2}$$

with white noise amplitude $\langle \mathbf{f}(t)\mathbf{f}(t')^T\rangle = 2\mathbf{D}\delta(t-t')$. Comparing the two equations, $\mathbf{f}(t)$ corresponds to the derivative of $\mathbf{W}_t$, which however (with probability one) is not differentiable; more precisely, upon taking $\mathrm{d}t \to 0$ one has $||\mathrm{d}\mathbf{W}_t/\mathrm{d}t|| = \infty$ with probability one, which is why the mathematics literature prefers Eq. (3.7.1).

If one describes the system on the level of probability densities instead of trajectories, the above equations translate to the Fokker-Planck equation $\partial_t G(\mathbf{x},t|\mathbf{x}_0) = \hat{L}(\mathbf{x})G(\mathbf{x},t|\mathbf{x}_0)$ with conditional density $G(\mathbf{x},t|\mathbf{x}_0)$ to be at $\mathbf{x}$ at time t after starting in $\mathbf{x}_0$ and the Fokker-Planck operator [6, 17]

$$\hat{L}(\mathbf{x}) = -\nabla_\mathbf{x}\cdot\mathbf{F}(\mathbf{x}) + \nabla_\mathbf{x}\cdot\mathbf{D}\nabla_\mathbf{x} = -\nabla_\mathbf{x}\cdot\hat{\mathbf{j}}_\mathbf{x}, \tag{3.7.3}$$

where we have defined the current operator $\hat{\mathbf{j}}_\mathbf{x} \equiv \mathbf{F}(\mathbf{x}) - \mathbf{D}\nabla_\mathbf{x}$. Note that all differential operators act on all functions to the right, e.g., $\nabla_\mathbf{x}\cdot\mathbf{F}(\mathbf{x})g(\mathbf{x}) = g(\mathbf{x})\nabla_\mathbf{x}\cdot\mathbf{F}(\mathbf{x}) + \mathbf{F}(\mathbf{x})\cdot\nabla_\mathbf{x}g(\mathbf{x})$. Although the approach presented here is more general, we restrict our attention to (possibly non-equilibrium) steady states where the drift $\mathbf{F}(\mathbf{x})$ is sufficiently smooth and confining to assure the existence of a steady-state (invariant) density $p_\mathrm{s}(\mathbf{x}) = \lim_{t\to\infty} G(\mathbf{x},t|\mathbf{x}_0)$ and steady-state current $\mathbf{j}_\mathrm{s}(\mathbf{x}) = \hat{\mathbf{j}}_\mathbf{x}p_\mathrm{s}(\mathbf{x})$. The special case $\mathbf{j}_\mathrm{s}(\mathbf{x}) = \mathbf{0}$ corresponds to equilibrium steady states. For systems that eventually evolve into a steady state we can rewrite the current operator as [84] (again the differential operator in $\nabla_\mathbf{x}p_\mathrm{s}^{-1}(\mathbf{x})$ also acts on functions to the right if $\hat{\mathbf{j}}_\mathbf{x}$ is applied to a function)

$$\hat{\mathbf{j}}_\mathbf{x} = \mathbf{j}_\mathrm{s}(\mathbf{x})p_\mathrm{s}^{-1}(\mathbf{x}) - \mathbf{D}p_\mathrm{s}(\mathbf{x})\nabla_\mathbf{x}p_\mathrm{s}^{-1}(\mathbf{x}). \tag{3.7.4}$$

We will later also restrict the treatment to systems evolving from steady-state initial conditions, i.e., the initial condition $\mathbf{x}_{t=0}$ is drawn according to the density p_s.

We define the two fundamental additive dynamical functionals— time-integrated current and density— as

$$\begin{aligned}\mathbf{J}_t &= \int_{\tau=0}^{\tau=t} U(\mathbf{x}_\tau) \circ \mathrm{d}\mathbf{x}_\tau \\ \rho_t &= \int_0^t V(\mathbf{x}_\tau)\mathrm{d}\tau,\end{aligned} \tag{3.7.5}$$

with differentiable and square-integrable (real-valued) functions $U, V\colon \mathbb{R}^d \to \mathbb{R}$ and $\circ$ denoting the Stratonovich integral [7, 14, 228]. These objects depend on the whole trajectory $[\mathbf{x}_\tau]_{0\leq\tau\leq t}$ and are thus random functionals with non-trivial statistics. In the following we will derive an equation for the characteristic function of the joint distribution of $\mathbf{x}_t, \rho_t, \mathbf{J}_t$ via a Feynman-Kac approach, which will then yield the moments (including variances and correlations) via a Dyson series. The formalism was already applied to the time-averaged density ρ_t/t (under the term of local/occupation time fraction) [41, 42, 174]. To do so, we need to derive a tilted Fokker-Planck equation, which we first do via Itô calculus and then, equivalently, via a functional calculus. Note that the tilted generator can also be found in the literature on large deviation theory [44, 107] (in this case obtained via the Feynman-Kac-Girsanov approach).

Tilting via Itô's Lemma

We first derive a tilted the Fokker-Planck equation using Itô calculus. From the Itô-Stratonovich correction term $\mathrm{d}U(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau/2$ and $\mathrm{d}\mathbf{x}_\tau\mathrm{d}\mathbf{x}_\tau^T = 2\mathbf{D}\mathrm{d}\tau$ (where $\mathbf{D} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T/2$) we obtain from Eqs. (3.7.1) and (3.7.5) the increments (curly brackets $\{\nabla \dots\}$ throughout denote that derivatives only act inside brackets)

$$\begin{aligned}\mathrm{d}\mathbf{J}_\tau &= U(\mathbf{x}_\tau) \circ \mathrm{d}\mathbf{x}_\tau = U(\mathbf{x}_\tau)\mathrm{d}\mathbf{x}_\tau + \mathbf{D}\,\{\nabla_{\mathbf{x}} U\}(\mathbf{x}_\tau)\mathrm{d}\tau \\ \mathrm{d}\rho_\tau &= V(\mathbf{x}_\tau)\mathrm{d}\tau.\end{aligned} \tag{3.7.6}$$

We use Itô's Lemma [14] in d dimensions for a twice differentiable test function $f = f(\mathbf{x}_t, \rho_t, \mathbf{J}_t)$ and Eqs. (3.7.1) and (3.7.6), to obtain

$$\begin{aligned}\mathrm{d}f &= \sum_{i=1}^d \frac{\partial f}{\partial x_i}\mathrm{d}x_t^i + \frac{\partial f}{\partial \rho}\mathrm{d}\rho_t + \sum_{i=1}^d \frac{\partial f}{\partial J_i}\mathrm{d}J_t^i \\ &\quad + \frac{1}{2}\sum_{i,j=1}^d \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\mathrm{d}x_t^i\mathrm{d}x_t^j + \frac{\partial^2 f}{\partial J_i \partial J_j}\mathrm{d}J_t^i\mathrm{d}J_t^j + 2\frac{\partial^2 f}{\partial x_i \partial J_j}\mathrm{d}x_t^i\mathrm{d}J_t^j\right) \\ &= [(\nabla_{\mathbf{x}} f) + (\nabla_{\mathbf{J}} f)U(\mathbf{x}_t)][\mathbf{F}(\mathbf{x}_t)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t] + (\nabla_{\mathbf{J}} f)\mathbf{D}\{\nabla_{\mathbf{x}} U\}(\mathbf{x}_t)\mathrm{d}t + V(\mathbf{x}_t)\partial_\rho f \mathrm{d}t\end{aligned}$$

J. Phys. A: Math. Theor. 56 155002 (2023)

$$+\left(\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{x}}+U(\mathbf{x}_t)^2\nabla_{\mathbf{J}}^T\mathbf{D}\nabla_{\mathbf{J}}+2U(\mathbf{x}_t)\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{J}}\right)f\mathrm{d}t. \tag{3.7.7}$$

For the time derivative of f this gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}f(\mathbf{x}_t,\rho_t,\mathbf{J}_t)=&\Bigg[\left(\mathbf{F}+\boldsymbol{\sigma}\frac{\mathrm{d}\mathbf{W}_t}{\mathrm{d}t}\right)(\nabla_{\mathbf{x}}+U\nabla_{\mathbf{J}})+\{\nabla_{\mathbf{x}}U\}\mathbf{D}\nabla_{\mathbf{J}}\\&+V\partial_\rho+\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{x}}+U^2\nabla_{\mathbf{J}}^T\mathbf{D}\nabla_{\mathbf{J}}+2U\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{J}}\Bigg]f(\mathbf{x}_t,\rho_t,\mathbf{J}_t).\end{aligned} \tag{3.7.8}$$

Following this formalism, we move towards a tilted Fokker-Planck equation [41,42]. Using the conditional probability density $Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0)$, we may write (omitting $\mathbf{x}$ dependence in $\mathbf{F},U,V$ for brevity) the evolution equation for $\langle f(\mathbf{x}_t,\rho_t,\mathbf{J}_t)\rangle_{\mathbf{x}_0}$, i.e., the expected value of $f(\mathbf{x}_t,\rho_t,\mathbf{J}_t)$ over the ensemble of paths propagating between $\mathbf{x}_0$ and $\mathbf{x}$ in time t. Using Eq. (3.7.8) and integration by parts, we obtain (note that non-negative functions $V\geq 0$ imply $\rho\geq 0$, such that one would restrict the ρ-integration to $\int_0^\infty \mathrm{d}\rho$ as in Ref. [41])

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle f(\mathbf{x}_t,\rho_t,\mathbf{J}_t)\rangle_{\mathbf{x}_0}&=\int\mathrm{d}^dx\int_{-\infty}^{\infty}\mathrm{d}\rho\int\mathrm{d}^dJ f(\mathbf{x},\rho,\mathbf{J})\partial_t Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0)\\
&=\int\mathrm{d}^dx\int_{-\infty}^{\infty}\mathrm{d}\rho\int\mathrm{d}^dJ\,Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0)\big[\mathbf{F}(\nabla_{\mathbf{x}}+U\nabla_{\mathbf{J}})+\{\nabla_{\mathbf{x}}U\}\mathbf{D}\nabla_{\mathbf{J}}\\
&\quad+V\partial_\rho+\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{x}}+U^2\nabla_{\mathbf{J}}^T\mathbf{D}\nabla_{\mathbf{J}}+2U\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{J}}\big]f(\mathbf{x},\rho,\mathbf{J})\\
&=\int\mathrm{d}^dx\int_{-\infty}^{\infty}\mathrm{d}\rho\int\mathrm{d}^dJ\,f(\mathbf{x},\rho,\mathbf{J})\Big[-\nabla_{\mathbf{x}}\mathbf{F}-U\mathbf{F}\nabla_{\mathbf{J}}-\{\nabla_{\mathbf{x}}U\}\mathbf{D}\nabla_{\mathbf{J}}-V\partial_\rho\\
&\quad+\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{x}}+U^2\nabla_{\mathbf{J}}^T\mathbf{D}\nabla_{\mathbf{J}}+2U\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{J}}\Big]Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0).\end{aligned} \tag{3.7.9}$$

Since the test function f is an arbitrary twice differentiable function, the resulting tilted Fokker-Planck equation reads

$$\partial_t Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0)=\hat{\mathcal{L}}_{\mathbf{x},\rho,\mathbf{J}}Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0), \tag{3.7.10}$$

with the tilted Fokker-Planck operator [10]

$$\begin{aligned}\hat{\mathcal{L}}_{\mathbf{x},\rho,\mathbf{J}}=&-\nabla_{\mathbf{x}}\cdot\mathbf{F}(\mathbf{x})+\nabla_{\mathbf{x}}^T\mathbf{D}\nabla_{\mathbf{x}}-V(\mathbf{x})\partial_\rho-U(\mathbf{x})\mathbf{F}(\mathbf{x})\cdot\nabla_{\mathbf{J}}\\
&-\{\nabla_{\mathbf{x}}U(\mathbf{x})\}^T\mathbf{D}\nabla_{\mathbf{J}}+U(\mathbf{x})^2\nabla_{\mathbf{J}}^T\mathbf{D}\nabla_{\mathbf{J}}+2\nabla_{\mathbf{J}}^T\mathbf{D}\nabla_{\mathbf{x}}U(\mathbf{x})\\
=&-[\nabla_{\mathbf{x}}+U(\mathbf{x})\nabla_{\mathbf{J}}]\mathbf{F}(\mathbf{x})-V(\mathbf{x})\partial_\rho+[\nabla_{\mathbf{x}}+U(\mathbf{x})\nabla_{\mathbf{J}}]^T\,\mathbf{D}\,[\nabla_{\mathbf{x}}+U(\mathbf{x})\nabla_{\mathbf{J}}]\,.\end{aligned} \tag{3.7.11}$$

[10] For non-negative functions $V\geq 0$ an additional boundary term appears at $\rho=0$ upon partial integration in Eq. (3.7.9), leading to an extra term $-V(\mathbf{x})\delta(\rho)$ in Eq. (3.7.10) that ensures conservation of probability (see Ref. [41]).

J. Phys. A: Math. Theor. 56 155002 (2023)

We see that the ρ dependence enters in standard Feynman-Kac form [41, 42], whereas the $\mathbf{J}$ dependence enters less trivially and shifts the gradient operator $\nabla_{\mathbf{x}} \to \nabla_{\mathbf{x}} + U(\mathbf{x})\nabla_{\mathbf{J}}$.

Tilting via functional calculus

We now rederive the tilted Fokker-Planck operator in Eq. (3.7.11) using a functional calculus approach [226, 227] instead of the Itô calculus in the previous section. This shows that both alternative approaches are equivalent, as expected. We closely follow the derivation of the Fokker-Planck equation in reference [226] but for d-dimensional space, and we generalize the approach to include the functionals defined in Eq. (3.7.5). The following approach is equivalent to a Stratonovich interpretation of stochastic calculus, which is manifested in the convention $\int_0^t \delta(t')\mathrm{d}t' = \int_0^t \delta(t-t')\mathrm{d}t' = 1/2$ [226]. The white noise term $\mathbf{f}(\tau)$ with $\left\langle \mathbf{f}(\tau)\mathbf{f}(\tau')^T \right\rangle_{\mathrm{s}} = 2\mathbf{D}\delta(\tau-\tau')$ in the Langevin equation (3.7.2) can be considered to be described by a path-probability measure [226]

$$P[\mathbf{f}] = N \exp\left[-\frac{1}{2}\int_0^t \mathbf{f}(\tau)^T \mathbf{D}^{-1}\mathbf{f}(\tau)\mathrm{d}\tau\right], \tag{3.7.12}$$

with normalization constant N, which may formally be problematic but always cancels out.

We now derive a tilted Fokker-Planck equation for the joint conditional density Q of $\mathbf{x}_t$ and the functionals $\mathbf{J}_t, \rho_t$, as defined in Eq. (3.7.5), given a deterministic initial condition $\mathbf{x}_0$ at time $t = 0$,

$$Q_t(\mathbf{x}, \rho, \mathbf{J}|\mathbf{x}_0) \equiv \int \mathcal{D}\mathbf{f} P[\mathbf{f}]\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t). \tag{3.7.13}$$

Note for the time derivatives that $\dot{\mathbf{J}}_t = U(\mathbf{x}_t)\dot{\mathbf{x}}_t$ and $\dot{\rho}_t = V(\mathbf{x}_t)$ to obtain (as a generalization of the calculation in reference [226] to dynamical functionals)

$$\begin{aligned}
\partial_t Q(\mathbf{x}, \rho, \mathbf{J}, t|\mathbf{x}_0) &= \partial_t \int \mathcal{D}\mathbf{f} P[\mathbf{f}]\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t) \\
&= \int \mathcal{D}\mathbf{f} P[\mathbf{f}] \left[-\nabla_{\mathbf{x}}\cdot\dot{\mathbf{x}}_t - \partial_\rho\dot{\rho}_t - \nabla_{\mathbf{J}}\cdot\dot{\mathbf{J}}_t\right]\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t) \\
&= \int \mathcal{D}\mathbf{f} P[\mathbf{f}] \left[-\nabla_{\mathbf{x}}\cdot[\mathbf{F}(\mathbf{x}_t)+\mathbf{f}(t)] - V(\mathbf{x}_t)\partial_\rho - U(\mathbf{x}_t)\left[\mathbf{F}(\mathbf{x}_t)+\mathbf{f}(t)\right]\nabla_{\mathbf{J}}\right] \times \\
&\quad \delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t) \\
&= \left[-\nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x}) - V(\mathbf{x})\partial_\rho - U(\mathbf{x})\mathbf{F}(\mathbf{x})\nabla_{\mathbf{J}}\right] Q_t(\mathbf{x}, \rho, \mathbf{J}|\mathbf{x}_0) \\
&\quad - \left[\nabla_{\mathbf{x}} + U(\mathbf{x})\nabla_{\mathbf{J}}\right]\cdot\int \mathcal{D}\mathbf{f} P[\mathbf{f}]\mathbf{f}(t)\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t).
\end{aligned} \tag{3.7.14}$$

J. Phys. A: Math. Theor. 56 155002 (2023)

The functional derivative of Eq. (3.7.12) reads [226]

$$\frac{\delta P[\mathbf{f}]}{\delta \mathbf{f}(t)} = -\frac{1}{2}\mathbf{D}^{-1}\mathbf{f}(t)P[\mathbf{f}], \tag{3.7.15}$$

which we use to obtain, via an integration by parts in $\delta\mathbf{f}(t)$,

$$\begin{aligned}
&-\int \mathcal{D}\mathbf{f} P[\mathbf{f}]\mathbf{f}(t)\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t)\\
&= 2\mathbf{D}\int \mathcal{D}\mathbf{f}\frac{\delta P[\mathbf{f}]}{\delta \mathbf{f}(t)}\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t)\\
&= -2\mathbf{D}\int \mathcal{D}\mathbf{f} P[\mathbf{f}]\frac{\delta}{\delta \mathbf{f}(t)}\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t).
\end{aligned} \tag{3.7.16}$$

As before, differentials are understood to act on all functions to the right, i.e., $\frac{\delta}{\delta\mathbf{f}(t)}$ here acts on the full product of delta functions. We obtain

$$\begin{aligned}
&\frac{\delta}{\delta \mathbf{f}(t)}\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t)\\
&= \left[-\nabla_{\mathbf{x}}\frac{\delta \mathbf{x}_t}{\delta \mathbf{f}(t)} - \partial_\rho \frac{\delta \rho_t}{\delta \mathbf{f}(t)} - \nabla_{\mathbf{J}}\frac{\delta \mathbf{x}_t}{\delta \mathbf{f}(t)}\right]\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t),
\end{aligned} \tag{3.7.17}$$

and we use that $\delta\rho_t/\delta\mathbf{f}(t) = \mathbf{0}$, and $\delta\mathbf{x}_t/\delta\mathbf{f}(t) = \mathbf{1}/2$ [226] which implies $\delta\mathbf{J}_t/\delta\mathbf{f}(t) = U(\mathbf{x}_t)\mathbf{1}/2$, to get

$$\frac{\delta}{\delta \mathbf{f}(t)}\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t) = \frac{1}{2}\left[-\nabla_{\mathbf{x}} - U(\mathbf{x}_t)\nabla_{\mathbf{J}}\right]\delta(\mathbf{x}-\mathbf{x}_t)\delta(\rho-\rho_t)\delta(\mathbf{J}-\mathbf{J}_t). \tag{3.7.18}$$

Plugging Eq. (3.7.18) first into Eq. (3.7.16) and then into Eq. (3.7.14) yields the tilted Fokker-Planck equation for the joint conditional density

$$\begin{aligned}
\partial_t Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0) = \Big[&-\nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x}) - V(\mathbf{x})\partial_\rho - U(\mathbf{x})\mathbf{F}(\mathbf{x})\nabla_{\mathbf{J}}\\
&+ \left[\nabla_{\mathbf{x}} + U(\mathbf{x})\nabla_{\mathbf{J}}\right]^T \mathbf{D}\left[\nabla_{\mathbf{x}} + U(\mathbf{x})\nabla_{\mathbf{J}}\right]\Big] Q_t(\mathbf{x},\rho,\mathbf{J}|\mathbf{x}_0).
\end{aligned} \tag{3.7.19}$$

Note that Eq. (3.7.19) fully agrees with Eq. (3.7.11) derived via Itô calculus thus establishing the announced equivalence of the two approaches.

J. Phys. A: Math. Theor. **56** 155002 (2023)

3.7.3 Steady-state covariance via Dyson expansion of the tilted propagator

In this section, we employ the tilted Fokker-Planck equation (3.7.19) to derive results for the mean value and (co)variances of time-integrated densities and currents. These follow as derivatives of the characteristic function evaluated at zero, and it thus suffices to treat the tilt as a perturbation of the "bare" generator (see [41]). The derivation is based on a Dyson expansion of the exponential of a Fourier-transformed tilted generator (i.e., tilted Fokker-Planck operator). Therefore, consider a one-dimensional Fourier variable ν and a d-dimensional Fourier variable $\boldsymbol{\omega} = (\omega_1, \dots, \omega_d)$ and define the Fourier transform of $Q_t(\mathbf{x}, \rho, \mathbf{J}|\mathbf{x}_0)$ as

$$\tilde{Q}_t(\mathbf{x}, \nu, \boldsymbol{\omega}|\mathbf{x}_0) \equiv \int_{-\infty}^{\infty} \mathrm{d}\rho \int \mathrm{d}^d J \, Q_t(\mathbf{x}, \rho, \mathbf{J}|\mathbf{x}_0) \exp\left(-\mathrm{i}\nu\rho - \mathrm{i}\boldsymbol{\omega} \cdot \mathbf{J}\right). \tag{3.7.20}$$

In the case $V \geq 0$ where $\rho \geq 0$ one would instead take the Laplace transform in the ρ-coordinate, see Ref. [41]. Recall the (untilted) Fokker-Planck operator $\hat{L}(\mathbf{x}) = -\nabla_{\mathbf{x}} \cdot \hat{\mathbf{j}}_{\mathbf{x}}$ with the current operator $\hat{\mathbf{j}}_{\mathbf{x}} = \mathbf{F}(\mathbf{x}) - \mathbf{D}\nabla_{\mathbf{x}}$ from Eq. (3.7.3). The Fourier transform of the tilted Fokker-Planck operator in Eqs. (3.7.11) and (3.7.19) reads

$$\begin{aligned} \hat{\mathcal{L}}(\mathbf{x}, \nu, \boldsymbol{\omega}) &= \hat{L}(\mathbf{x}) - \mathrm{i}\nu V(\mathbf{x}) - \mathrm{i}\boldsymbol{\omega}^T \cdot \hat{\mathbf{L}}^U(\mathbf{x}) - U(\mathbf{x})^2 \boldsymbol{\omega}^T \mathbf{D} \boldsymbol{\omega}, \\ \hat{\mathbf{L}}^U(\mathbf{x}) &\equiv U(\mathbf{x}) \hat{\mathbf{j}}_{\mathbf{x}} - \mathbf{D}\nabla_{\mathbf{x}} U(\mathbf{x}). \end{aligned} \tag{3.7.21}$$

As always, the differential operators act on all functions to the right unless written inside curly brackets, i.e., $\nabla_{\mathbf{x}} U(\mathbf{x}) = \{\nabla_{\mathbf{x}} U(\mathbf{x})\} + U(\mathbf{x})\nabla_{\mathbf{x}}$. Note that whereas we obtained the tilted generator directly and only subsequently Fourier transformed it, there are also approaches that directly target the Fourier image of the tilted generator (see, e.g., [229]). Compared to the tilt of the density (i.e., the ν-term; see also [41]), the tilt corresponding to the current observable ($\boldsymbol{\omega}$-terms) involves more terms and even a term that is second order in $\boldsymbol{\omega}$. The second order term occurs since $(\mathrm{d}\mathbf{W}_\tau)^2 \sim \mathrm{d}\tau$ and therefore (in contrast to $\mathrm{d}\tau \mathrm{d}\mathbf{W}_\tau$ and $\mathrm{d}\tau^2$) contributes in the tilting of the generator.

We now restrict our attention to dynamics starting in the steady state p_{s}, and denote the average over an ensemble over paths propagating from the steady state by $\langle \cdot \rangle_{\mathrm{s}}$. Extensions of the formalism to any initial distribution are straightforward and introduce additional

J. Phys. A: Math. Theor. **56** 155002 (2023)

transient terms. For the derivation of the moments of ρ_t and $\mathbf{J}_t$, we introduce and expand the characteristic function (also known as moment-generating function)

$$\tilde{\mathcal{P}}_t^{\rho\mathbf{J}}(\nu,\boldsymbol{\omega}|p_\mathrm{s}) \equiv \left\langle \mathrm{e}^{-\mathrm{i}\nu\rho_t - \mathrm{i}\boldsymbol{\omega}\cdot\mathbf{J}_t}\right\rangle_\mathrm{s} = 1 - \mathrm{i}\nu\left\langle\rho_t\right\rangle_\mathrm{s} - \mathrm{i}\boldsymbol{\omega}\cdot\left\langle\mathbf{J}_t\right\rangle_\mathrm{s} - \nu\boldsymbol{\omega}\cdot\left\langle\rho_t\mathbf{J}_t\right\rangle_\mathrm{s} + O(\boldsymbol{\omega}^2,\nu^2). \tag{3.7.22}$$

This expansion in $\nu,\boldsymbol{\omega}$ will now be compared to the Dyson expansion of the exponential of Eq. (3.7.21) which yields expressions for $\left\langle\rho_t\right\rangle_\mathrm{s}, \left\langle\mathbf{J}_t\right\rangle_\mathrm{s}, \left\langle\rho_t\mathbf{J}_t\right\rangle_\mathrm{s}$ by comparing individual orders.

The Dyson expansion allows to expand for small $|\nu|, |\boldsymbol{\omega}|$ (see also [41])

$$\begin{aligned}\mathrm{e}^{\hat{\mathcal{L}}(\mathbf{x}_1,v,\boldsymbol{\omega})t} =& 1 - \mathrm{i}\int_0^t \mathrm{d}t_1 \mathrm{e}^{\hat{L}(\mathbf{x}_1)(t-t_1)}\left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_1)\right]\mathrm{e}^{\hat{L}(\mathbf{x}_1)t_1} \\ &- \int_0^t \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1 \mathrm{e}^{\hat{L}(\mathbf{x}_1)(t-t_2)}\left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_1)\right]\mathrm{e}^{\hat{L}(\mathbf{x}_1)(t_2-t_1)} \\ &\left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_1)\right]\mathrm{e}^{\hat{L}(\mathbf{x}_1)t_1} + O(\boldsymbol{\omega}^2,\nu^2).\end{aligned} \tag{3.7.23}$$

Using that the first propagation only differs from 1 by total derivatives (recall $\hat{L}(\mathbf{x}) = -\nabla_\mathbf{x}\cdot\hat{\mathbf{j}}_\mathbf{x}$), and using for the last propagation term $\mathrm{e}^{\hat{L}(\mathbf{x}_1)t_1}p_\mathrm{s}(\mathbf{x}_1) = p_\mathrm{s}(\mathbf{x}_1)$, we obtain

$$\begin{aligned}\tilde{\mathcal{P}}_t^{\rho\mathbf{J}}(v,\boldsymbol{\omega}|p_\mathrm{s}) =& \int \mathrm{d}^d x_1\, \mathrm{e}^{\hat{\mathcal{L}}(\mathbf{x}_1,v,\boldsymbol{\omega})t}p_\mathrm{s}(\mathbf{x}_1) \\ =& 1 - \mathrm{i}\int \mathrm{d}^d x_1 \int_0^t \mathrm{d}t_1 \left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_1)\right]p_\mathrm{s}(\mathbf{x}_1) \\ &- \sum_{l,m=1}^{d}\int \mathrm{d}^d x_1 \int_0^t \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1 \left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_1)\right]\mathrm{e}^{\hat{L}(\mathbf{x}_1)(t_2-t_1)} \\ &\left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_1)\right]p_\mathrm{s}(\mathbf{x}_1) + O(\boldsymbol{\omega}^2,v^2).\end{aligned} \tag{3.7.24}$$

We substitute the one-step propagation by the conditional density $G(\mathbf{x}_2,t|\mathbf{x}_1) = \mathrm{e}^{\hat{L}(\mathbf{x}_1)t}\delta(\mathbf{x}_2-\mathbf{x}_1)$ [7, 16],

$$\int \mathrm{d}^d x_1 f(\mathbf{x}_1)\mathrm{e}^{\hat{L}(\mathbf{x}_1)(t_2-t_1)}g(\mathbf{x}_1) = \int \mathrm{d}^d x_1 \int \mathrm{d}^d x_2 f(\mathbf{x}_2)G(\mathbf{x}_2,t_2-t_1|\mathbf{x}_1)g(\mathbf{x}_1), \tag{3.7.25}$$

which yields

$$\begin{aligned}\tilde{\mathcal{P}}_t^{\rho\mathbf{J}}(v,\boldsymbol{\omega}|p_\mathrm{s}) =& 1 - \mathrm{i}\int \mathrm{d}^d x_1 \int_0^t \mathrm{d}t_1 \left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_1)\right]p_\mathrm{s}(\mathbf{x}_1) \\ &- \int \mathrm{d}^d x_1 \int \mathrm{d}^d x_2 \int_0^t \mathrm{d}t_2 \int_0^{t_2}\mathrm{d}t_1 \left[\nu V(\mathbf{x}_2) + \boldsymbol{\omega}^T\cdot\hat{\mathbf{L}}^U(\mathbf{x}_2)\right]\end{aligned}$$

J. Phys. A: Math. Theor. 56 155002 (2023)

$$G(\mathbf{x}_2, t_2 - t_1|\mathbf{x}_1)\left[\nu V(\mathbf{x}_1) + \boldsymbol{\omega}^T \cdot \hat{\mathbf{L}}^U(\mathbf{x}_1)\right] p_{\rm s}(\mathbf{x}_1) + O(\boldsymbol{\omega}^2, \nu^2). \quad (3.7.26)$$

This concludes the expansion of the exponential of the Fourier transformed tilted generator. Now, by comparing the definition and expansion of the characteristic function Eq. (3.7.22) with the result Eq. (3.7.26) from the Dyson expansion, we obtain the moments and correlations of the functionals $\mathbf{J}_t = \int_{\tau=0}^{\tau=t} U(\mathbf{x}_\tau) \circ \mathrm{d}\mathbf{x}_\tau$ and $\rho_t = \int_0^t V(\mathbf{x}_\tau)\mathrm{d}\tau$.

Note that the first moments (i.e., the mean values for steady-state initial conditions) can also be obtained directly [34,84], but we obtain them here by comparing the terms of order ν and $\boldsymbol{\omega}$ in Eqs. (3.7.22) and (3.7.26),

$$\begin{aligned}\langle \rho_t \rangle_{\rm s} &= \int_0^t \mathrm{d}t_1 \int \mathrm{d}^d x_1 V(\mathbf{x}_1) p_{\rm s}(\mathbf{x}_1) = t \int \mathrm{d}^d x_1 V(\mathbf{x}_1) p_{\rm s}(\mathbf{x}_1) \\ \langle \mathbf{J}_t \rangle_{\rm s} &= t \int \mathrm{d}^d x_1 [U(\mathbf{x}_1)\hat{\mathbf{j}}_{\mathbf{x}_1} - \mathbf{D}\nabla_{\mathbf{x}_1} U(\mathbf{x}_1)] p_{\rm s}(\mathbf{x}_1) = t \int \mathrm{d}^d x_1 U(\mathbf{x}_1)\mathbf{j}_{\rm s}(\mathbf{x}_1), \end{aligned} \quad (3.7.27)$$

where $\nabla_{\mathbf{x}_1} U(\mathbf{x}_1) p_{\rm s}(\mathbf{x}_1)$ vanishes after integration by parts and $\mathbf{j}_{\rm s}(\mathbf{x}_1) \equiv \hat{\mathbf{j}}_{\mathbf{x}_1} p_{\rm s}(\mathbf{x}_1)$ is the steady-state current.

By comparing the terms of order $\nu\boldsymbol{\omega}$ in Eqs. (3.7.22) and (3.7.26) we have for the steady-state expectation $\langle \mathbf{J}_t \rho_t \rangle_{\rm s}$ that

$$\begin{aligned}\langle \mathbf{J}_t \rho_t \rangle_{\rm s} &= \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \int \mathrm{d}^d x_1 \int \mathrm{d}^d x_2 \\ &\quad \left[\hat{\mathbf{L}}^U(\mathbf{x}_2) G(\mathbf{x}_2, t_2 - t_1|x_1) V(\mathbf{x}_1) + V(\mathbf{x}_2) G(\mathbf{x}_2, t_2 - t_1|x_1) \mathbf{L}_U(\mathbf{x}_1)\right] p_{\rm s}(\mathbf{x}_1) \\ &= \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \int \mathrm{d}^d x_1 \int \mathrm{d}^d x_2 \Big[U(\mathbf{x}_2)\hat{\mathbf{j}}_{\mathbf{x}_2} G(\mathbf{x}_2, t_2 - t_1|\mathbf{x}_1) V(\mathbf{x}_1) \\ &\quad + V(\mathbf{x}_2) G(\mathbf{x}_2, t_2 - t_1|\mathbf{x}_1)[U(\mathbf{x}_1)\hat{\mathbf{j}}_{\mathbf{x}_1} - \mathbf{D}\nabla_{\mathbf{x}_1} U(\mathbf{x}_1)]\Big] p_{\rm s}(\mathbf{x}_1). \end{aligned} \quad (3.7.28)$$

We note that for any function f the following identity holds

$$\int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 f(t_2 - t_1) = \int_0^t \mathrm{d}t'(t - t') f(t'), \quad (3.7.29)$$

and further introduce the shorthand notation

$$\hat{\mathcal{I}}^t_{\mathbf{xy}}[\cdots] = \int_0^t \mathrm{d}t'(t - t') \int \mathrm{d}^d x_1 \int \mathrm{d}^d x_2 U(\mathbf{x}_1) V(\mathbf{x}_2)[\cdots]. \quad (3.7.30)$$

Moreover, we define the joint density $P_{\mathbf{y}}(\mathbf{x}, t) \equiv G(\mathbf{x}, t|\mathbf{y}) p_{\rm s}(\mathbf{y})$ and following Ref. [84] introduce the dual-reversed current operator $\hat{\mathbf{j}}^\ddagger_{\mathbf{x}} \equiv \mathbf{j}_{\rm s}(\mathbf{x})/p_{\rm s}(\mathbf{x}) + \mathbf{D}p_{\rm s}(\mathbf{x})\nabla_{\mathbf{x}} p_{\rm s}^{-1}(\mathbf{x}) = -\hat{\mathbf{j}}_{\mathbf{x}}(\mathbf{j}_{\rm s} \to -\mathbf{j}_{\rm s})$. With these notations, using integration by parts, and by relabeling $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$ in one

J. Phys. A: Math. Theor. 56 155002 (2023)

term, we rewrite Eq. (3.7.28) to obtain for the correlation, reproducing the main result of Refs. [83, 84],

$$\begin{aligned}\langle \mathbf{J}_t \rho_t \rangle_\mathrm{s} - \langle \mathbf{J}_t \rangle \langle \rho_t \rangle_\mathrm{s} =& \hat{\mathcal{I}}^t_{\mathbf{xy}} \Big[\hat{\mathbf{j}}_{\mathbf{x}_1} P_{\mathbf{x}_2}(\mathbf{x}_1, t') + \mathbf{j}_\mathrm{s}(\mathbf{x}_1) p_\mathrm{s}^{-1}(\mathbf{x}_1) P_{\mathbf{x}_1}(\mathbf{x}_2, t') \\ &+ \mathbf{D} p_\mathrm{s}(\mathbf{x}_1) \nabla_{\mathbf{x}_1} p_\mathrm{s}(\mathbf{x}_1)^{-1} P_{\mathbf{x}_1}(\mathbf{x}_2, t') \Big] - \langle \mathbf{J}_t \rangle \langle \rho_t \rangle_\mathrm{s} \\ =& \hat{\mathcal{I}}^t_{\mathbf{xy}} \left[\hat{\mathbf{j}}_{\mathbf{x}_1} P_{\mathbf{x}_2}(\mathbf{x}_1, t') + \hat{\mathbf{j}}^\ddagger_{\mathbf{x}_1} P_{\mathbf{x}_1}(\mathbf{x}_2, t') - 2\mathbf{j}_\mathrm{s}(\mathbf{x}_1) p_\mathrm{s}(\mathbf{x}_2) \right], \end{aligned} \tag{3.7.31}$$

We will discuss this result below, but first derive analogous results for (co)variances of densities and currents, respectively.

Instead of obtaining $\langle \rho_t^2 \rangle_\mathrm{s}$ from the ν^2 order in Eq. (3.7.26) we here consider a generalization to two densities, $\rho_t = \int_0^t V(\mathbf{x}_\tau) \mathrm{d}\tau$ and $\rho'_t = \int_0^t U(\mathbf{x}_\tau) \mathrm{d}\tau$. The Fourier-transformed tilted generator in Eq. (3.7.21) with Fourier variables ν, ν' corresponding to ρ_t, ρ'_t is obtained equivalently and gives $\hat{\mathcal{L}}(\mathbf{x}, v, v') = \hat{L}(\mathbf{x}) - \mathrm{i}\nu V(\mathbf{x}) - \mathrm{i}\nu' U(\mathbf{x})$. The related term in the Dyson series (by an adaption of Eq. (3.7.26) including $\nu' U$) becomes $[\nu V(\mathbf{x}_2) + \nu' U(\mathbf{x}_2)] G(\mathbf{x}_2, t_2 - t_1 | \mathbf{x}_1) [\nu V(\mathbf{x}_1) + \nu' U(\mathbf{x}_1)] p_\mathrm{s}(\mathbf{x}_1)$ (see also [41]). By comparison with the characteristic function in Eq. (3.7.22) including ρ'_t, one obtains the known result [41, 42],

$$\langle \rho_t \rho'_t \rangle_\mathrm{s} - \langle \rho_t \rangle_\mathrm{s} \langle \rho'_t \rangle_\mathrm{s} = \hat{\mathcal{I}}^t_{\mathbf{xy}} [P_{\mathbf{x}_2}(\mathbf{x}_1, t') + P_{\mathbf{x}_1}(\mathbf{x}_2, t') - 2 p_\mathrm{s}(\mathbf{x}_1) p_\mathrm{s}(\mathbf{x}_2)]. \tag{3.7.32}$$

For $U = V$ this becomes the variance of ρ_t which can also be obtained from the order ν^2 in Eqs. (3.7.22) and (3.7.26).

To obtain the current covariance, we accordingly require a tilted generator with two Fourier variables $\boldsymbol{\omega}, \boldsymbol{\omega}'$ corresponding to $\mathbf{J}_t = \int_{\tau=0}^{\tau=t} U(\mathbf{x}_\tau) \circ \mathrm{d}\mathbf{x}_\tau$ and $\mathbf{J}'_t = \int_{\tau=0}^{\tau=t} V(\mathbf{x}_\tau) \circ \mathrm{d}\mathbf{x}_\tau$, which can, by the same formalism, be derived as

$$\begin{aligned}\hat{\mathcal{L}}(\mathbf{x}, \boldsymbol{\omega}, \boldsymbol{\omega}') =& \hat{L}(\mathbf{x}) - \mathrm{i}\boldsymbol{\omega}^T \cdot \hat{\mathbf{L}}^U(\mathbf{x}) - \mathrm{i}\boldsymbol{\omega}'^T \cdot \hat{\mathbf{L}}^V(\mathbf{x}) - U(\mathbf{x})^2 \boldsymbol{\omega}^T \mathbf{D} \boldsymbol{\omega} - V(\mathbf{x})^2 \boldsymbol{\omega}'^T \mathbf{D} \boldsymbol{\omega}' \\ &- 2U(\mathbf{x}) V(\mathbf{x}) \boldsymbol{\omega}^T \mathbf{D} \boldsymbol{\omega}' \\ \hat{\mathbf{L}}^V(\mathbf{x}) \equiv& V(\mathbf{x}) \hat{\mathbf{j}}_\mathbf{x} - \mathbf{D} \nabla_\mathbf{x} V(\mathbf{x}). \end{aligned} \tag{3.7.33}$$

The Dyson series (by adapting Eq. (3.7.26)) based on $\hat{\mathcal{L}}(\mathbf{x}, \boldsymbol{\omega}, \boldsymbol{\omega}')$ for two currents $\mathbf{J}, \mathbf{J}'$ reads

$$\begin{aligned}&\tilde{\mathcal{P}}_t^{\mathbf{J}\mathbf{J}'}(\boldsymbol{\omega}, \boldsymbol{\omega}' | p_\mathrm{s}) = \\ &1 - \int \mathrm{d}^d x_1 \int_0^t \mathrm{d}t_1 \left[\mathrm{i}\boldsymbol{\omega}^T \cdot \hat{\mathbf{L}}^U(\mathbf{x}_1) + \mathrm{i}\boldsymbol{\omega}'^T \cdot \hat{\mathbf{L}}^V(\mathbf{x}_1) + 2U(\mathbf{x}_1) V(\mathbf{x}_1) \boldsymbol{\omega}^T \mathbf{D} \boldsymbol{\omega}' \right] p_\mathrm{s}(\mathbf{x}_1) \end{aligned}$$

J. Phys. A: Math. Theor. 56 155002 (2023)

$$+\int \mathrm{d}^d x_1 \int \mathrm{d}^d x_2 \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \left[\mathrm{i}\boldsymbol{\omega}^T \cdot \hat{\mathbf{L}}^U(\mathbf{x}_2) + \mathrm{i}\boldsymbol{\omega}'^T \cdot \hat{\mathbf{L}}^V(\mathbf{x}_2)\right]$$
$$G(\mathbf{x}_2, t_2 - t_1|\mathbf{x}_1)\left[\mathrm{i}\boldsymbol{\omega}^T \cdot \hat{\mathbf{L}}^U(\mathbf{x}_1) + \mathrm{i}\boldsymbol{\omega}'^T \cdot \hat{\mathbf{L}}^V(\mathbf{x}_1)\right] p_\mathrm{s}(\mathbf{x}_1) + O(\omega^2, \omega'^2). \tag{3.7.34}$$

The expectation value of the product of current components $\langle J_{t,n} J'_{t,m}\rangle_\mathrm{s}$ is given by the terms that are linear in $\omega_n \omega'_m$, i.e., (recall $D_{nm} = D_{mn}$)

$$\langle J_{t,n} J'_{t,m}\rangle_\mathrm{s} = 2tD_{nm}\int \mathrm{d}^d x_1\, U(\mathbf{x}_1)V(\mathbf{x}_1)p_\mathrm{s}(\mathbf{x}_1) + \int_0^t \mathrm{d}t'(t-t')\int \mathrm{d}^d x_1 \int \mathrm{d}^d x_2$$
$$\left[\hat{L}_n^U(\mathbf{x}_2)G(\mathbf{x}_2, t'|\mathbf{x}_1)\cdot \hat{L}_m^V(\mathbf{x}_1)p_\mathrm{s}(\mathbf{x}_1) + \hat{L}_m^V(\mathbf{x}_2)G(\mathbf{x}_2, t'|\mathbf{x}_1)\cdot \hat{L}_n^U(\mathbf{x}_1)p_\mathrm{s}(\mathbf{x}_1)\right]. \tag{3.7.35}$$

We denote by $\hat{=}$ equality up to gradient terms that vanish upon integration to write

$$\begin{aligned}
&\hat{L}_n^U(\mathbf{x}_2)G(\mathbf{x}_2, t'|\mathbf{x}_1)\cdot \hat{L}_m^V(\mathbf{x}_1)p_\mathrm{s}(\mathbf{x}_1) \hat{=} U(\mathbf{x}_2)\hat{\mathrm{j}}_{\mathbf{x}_2,n} G(\mathbf{x}_2, t'|\mathbf{x}_1)\times \\
&\quad \left[V(\mathbf{x}_1)\mathbf{j}_\mathrm{s}(\mathbf{x}_1)p_\mathrm{s}^{-1}(\mathbf{x}_1) - p_\mathrm{s}(\mathbf{x}_1)\mathbf{D}\nabla_{\mathbf{x}_1} p_\mathrm{s}(\mathbf{x}_1)^{-1} - \mathbf{D}\nabla_{\mathbf{x}_1} V(\mathbf{x}_1)\right]_m p_\mathrm{s}(\mathbf{x}_1) \\
&\hat{=} U(\mathbf{x}_2)V(\mathbf{x}_1)\hat{\mathrm{j}}_{\mathbf{x}_2,n}[\mathbf{j}_\mathrm{s}(\mathbf{x}_1)p_\mathrm{s}^{-1}(\mathbf{x}_1) + p_\mathrm{s}(\mathbf{x}_1)\mathbf{D}\nabla_{\mathbf{x}_1} p_\mathrm{s}^{-1}(\mathbf{x}_1)]_m G(\mathbf{x}_2, t'|\mathbf{x}_1)p_\mathrm{s}(\mathbf{x}_1) \\
&= U(\mathbf{x}_2)V(\mathbf{x}_1)\hat{\mathbf{j}}_{\mathbf{x}_2,n}\hat{\mathbf{j}}^\ddagger_{\mathbf{x}_1,m} P_{\mathbf{x}_1}(\mathbf{x}_2, t).
\end{aligned} \tag{3.7.36}$$

Inserting this into Eq. (3.7.35), and relabeling in one term $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$ we obtain for the nm-element of the current covariance matrix

$$\begin{aligned}
\langle J_{t,n} J'_{t,m}\rangle_\mathrm{s} - \langle J_{t,n}\rangle_\mathrm{s}\langle J'_{t,m}\rangle_\mathrm{s} =& 2tD_{nm}\int \mathrm{d}^d x_1\, U(\mathbf{x}_1)V(\mathbf{x}_1)p_\mathrm{s}(\mathbf{x}_1) \\
&+ \hat{\mathcal{I}}^t_{\mathbf{xy}}\left[\hat{\mathrm{j}}_{\mathbf{x}_1,m}\hat{\mathrm{j}}^\ddagger_{\mathbf{x}_2,n} P_{\mathbf{x}_2}(\mathbf{x}_1, t') + \hat{\mathrm{j}}_{\mathbf{x}_2,n}\cdot \hat{\mathrm{j}}^\ddagger_{\mathbf{x}_1,m} P_{\mathbf{x}_1}(\mathbf{x}_2, t')\right].
\end{aligned} \tag{3.7.37}$$

This reproduces and slightly generalizes the main result of Refs. [83, 84] where the diagonal elements ($m = n$) of the covariance matrix were derived. This result for the current covariance matrix and Eq. (3.7.31) for the current-density correlation are the natural generalizations of the density-density covariance Eq. (3.7.32), as described in detail in references [83, 84], with the additional $2tD_{nm}$-term in Eq. (3.7.37) arising from the $(\mathrm{d}\mathbf{W}_\tau)^2$ contribution in $J_{t,n}J'_{t,m}$ manifested in the term $-2U(\mathbf{x})V(\mathbf{x})\boldsymbol{\omega}^T\mathbf{D}\boldsymbol{\omega}'$ in the tilted generator in Eq. (3.7.33). While the density-density covariance Eq. (3.7.32) only depends on integration over all paths from $\mathbf{x}_1$ to $\mathbf{x}_2$ (and vice versa) in time t' via $P_{\mathbf{x}_1}(\mathbf{x}_2, t')$, the current-density correlation Eq. (3.7.31) instead involves $\hat{\mathbf{j}}_{\mathbf{x}_1} P_{\mathbf{x}_2}(\mathbf{x}_1, t')$ and $\hat{\mathbf{j}}^\ddagger_{\mathbf{x}_1} P_{\mathbf{x}_1}(\mathbf{x}_2, t')$ which describe currents at the final- and initial-points, respectively [84]. This notion is further extended in the result Eq. (3.7.37) where $\hat{\mathrm{j}}_{\mathbf{x}_2,n}\hat{\mathrm{j}}^\ddagger_{\mathbf{x}_1,m} P_{\mathbf{x}_1}(\mathbf{x}_2, t')$ corresponds to products of components of displacements along individual trajectories from $\mathbf{x}_1$ to $\mathbf{x}_2$ [83].

J. Phys. A: Math. Theor. 56 155002 (2023)

3.7.4 Conclusion

We employed a Feynman-Kac approach to derive moments and correlations of dynamical functionals of diffusive paths — the time-integrated densities and currents. We presented two different but equivalent approaches to tilting the generator — Itô and functional calculus. These two approaches illustrate how one can freely choose between Itô and functional calculus to derive results on dynamical functionals. In particular, both approaches are accessible without further technical mathematical concepts such as the Cameron-Martin-Girsanov theorem that is often used in the study of tilted generators. Our methodological advance thus provides a flexible repertoire of easily accessible methods that will hopefully prove useful in future studies of related problems.

The derivation of the moments and correlations based on the tilted generator reproduces results with important implications for stochastic thermodynamics and large deviation theory, in particular for the physical and mathematical role of coarse graining [83, 84], and thereby displays how the tilted generator yields results on the statistics of dynamical functionals, even beyond the large deviation limit.

Acknowledgments

Financial support from Studienstiftung des Deutschen Volkes (to C. D.) and the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2 (to A. G.) is gratefully acknowledged.

Data availability statement

No new data were created or analyzed in this study.

J. Phys. A: Math. Theor. **56** 155002 (2023)

Chapter 4

Direct route to thermodynamic uncertainty relations and their saturation

We now come to an explicit application of stochastic time-integrated currents, namely *thermodynamic uncertainty relations* (TURs). We already encountered the TUR for dynamics in a non-equilibrium steady state (NESS) in the publications shown in Secs. 3.4 and 3.5. Here, we give a more comprehensive introduction to TURs (Sec. 4.1), presenting the technical background (Sec. 4.2) and state of the art (Sec. 4.3). We then summarize (Sec. 4.4) and reproduce (Sec. 4.5) original results from a publication, showing how the methods developed in the previous chapter allow to gain a deep insight into the TUR, far beyond the notion of coarse graining.

4.1 Introduction to TURs

The term TURs refers to a class of inequalities for overdamped Langevin dynamics and Markov jump dynamics on a discrete state space. The original TUR was discovered and formulated for non-equilibirum steady-state dynamics in 2015 in Ref. [47], and subsequently proven in Ref. [48]. For steady-state dynamics in continuous space, the TUR states that for an overdamped Langevin dynamics $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$ governed by $\mathrm{d}\mathbf{x}_\tau = \mathbf{a}(\mathbf{x}_\tau)\mathrm{d}\tau + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_\tau$ and for any time-integrated (or time-averaged) current $J_t = \int_{\tau=0}^{\tau=t} \mathbf{U}(\mathbf{x}_\tau)\cdot\circ\mathrm{d}\mathbf{x}_\tau$ where "$\cdot\circ$"

denotes the scalar product and Stratonovich integration, the expected heat Σ_t dissipated into the medium in a time-interval $[0,t]$[1] obeys the inequality

$$\frac{\mathrm{var}(J_t)}{\langle J_t\rangle^2}\times\Sigma_t\geq 2k_{\mathrm{B}}T\,. \tag{4.1.1}$$

It can be seen as an improved version of the second law of thermodynamics, stating that the total entropy production is non-decreasing $\Sigma_t/T=\Delta S_{\mathrm{tot}}\geq 0$, in the sense that for any J_t the zero can be replaced by the superior bound $\Delta S_{\mathrm{tot}}\geq 2k_{\mathrm{B}}\langle J_t\rangle^2/\mathrm{var}(J_t)$. Beyond this improvement of the second law, there are two very intriguing ways to read Eq. (4.1.1), presented in Subsecs. 4.1.1 and 4.1.2. The main power of the inequality lies in its direct experimental accessibility, explained in Subsec. 4.1.3, and its generality, discussed in Subsec. 4.1.4.

4.1.1 The cost of precision

From a conceptual point of view, the TUR (4.1.1) can be read as a "cost of precision" [152], in the sense that a precise stochastic current (that is, a current with a small relative fluctuation $\mathrm{var}(J_t)/\langle J_t\rangle^2$) implies a large energy cost Σ_t. This nicely matches the observation that living systems, that generally require a high precision for functionality, produce much entropy.

To illustrate the cost of precision for a concrete example, it is instructive to consider the thought experiment of a thermodynamic clock [213]. For a simple example consider as a current simply the displacement $J_t=\int_0^t 1\circ\mathrm{d}x_\tau=x_t-x_0$ for the one-dimensional overdamped motion[2] $\mathrm{d}x_\tau=a(x_\tau)\mathrm{d}\tau+\sqrt{2D}\mathrm{d}W_\tau$ on some interval with periodic boundary conditions, illustrated as a circle in Fig. 4.1.1. The position at time τ can then be thought of as a hand of a clock. Assuming steady-state initial conditions, demanding a *precision* that allows for an expected fluctuation of, say one second per hour, i.e., $\mathrm{var}(x_t-x_0)/\langle x_t-x_0\rangle^2\leq 1/3600$, implies that the dissipation is at least $\Sigma_t\geq 7200k_{\mathrm{B}}T$ which we would call the energetic *cost* of the precision.

For comparing this to actual clocks, note that despite the large generality of the TUR (see Subsec. 4.1.4), the assumptions exclude quantum mechanical effects and underdamped motion, which is why this precision limit of clocks is not fully universal. Also note that

[1] The expected heat Σ_t dissipated into the medium equals the product of temperature and entropy production in the medium, and in the steady state also the total entropy production $\Sigma_t=T\Delta S_{\mathrm{med}}([0,t])=T\Delta S_{\mathrm{tot}}([0,t])$, see Sec. 2.9

[2] More generally, the one-dimensional motion in this example may also originate from a projection of a higher-dimensional overdamped motion, see Subsec. 4.1.4.

$7200k_{\mathrm{B}}T$ for a few hundred Kelvins are around 10^{-17} Joules, so for a macroscopic clock this is only a small energy cost.

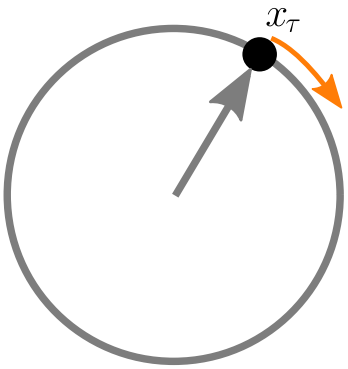

Figure 4.1.1: Sketch of a thermodynamic clock. The orange arrow indicates the direction of the non-equilibrium drift.

Related to the cost of precision, the TUR is also useful for quantifying the efficiency (*quality factor*) of a molecular motor by comparing the minimal cost of precision obtained from the TUR to the actual energy dissipation (see, e.g., [152]).

4.1.2 The inference of dissipation or entropy production

From a more practical point of view, as already pointed out in the previous chapter, for any observed current J_t the inequality (4.1.1) yields a lower bound for Σ_t. This is very useful, since the explicit expressions shown for the total entropy production in Sec. 2.9 require complete knowledge of p_{s}, $\mathbf{j}_{\mathrm{s}}$, and $\mathbf{D}$, or even more details for systems that are not in a steady state. In contrast, inferring a lower bound for Σ_t via the TUR (4.1.1) is directly accessible from measured trajectories without extra knowledge.

Decisive in the inference of dissipation via currents computed from observed trajectories is the choice of the window function $\mathbf{U}$. As found in Ch. 3, for multidimensional space ($d \geq 2$) it is beneficial to choose a function that varies on the correct length scale, i.e., not a delta function or a constant (for the special case of periodic boundaries, see, e.g., Fig. 4.1.1, a constant may also work well). For the sake of inference, the notion of coarse graining developed in Ch. 3 is particularly useful if only some local part of the system is sampled well.

If more than a local part of the system is sampled well, one can combine the different length scales corresponding to the different areas (see Fig. 3.5.6). Building up on the previous insight, which was to choose window functions around points with large currents, it is

expected that the optimal window function will be some local weighting of regions with high currents. We will show precisely how this works in the discussion of the saturation of the TUR, see Subsec. 4.3.2 and in the reproduced publication in Subsec. 4.5.6.

4.1.3 Experimental accessibility

An important point about the TUR (4.1.1) is the direct accessibility of the expressions involving the current from measured trajectories (also called *operational accessibility*, see, e.g., [211]). As already pointed out in Subsec. 3.1, the time-integrated current J_t is directly accessible from a measured trajectory via the Riemann sum in Eq. (2.5.15). Then the ensemble average $\langle J_t \rangle$ and variance $\mathrm{var}(J_t)$ can be directly estimated by sample averages over the measured trajectories. If sufficiently many trajectories are available with a sufficient time resolution, the TUR (4.1.1) can be readily applied to infer a lower bound on Σ_t.

4.1.4 Generality of the TUR

For both applications of the TUR presented in the previous two subsections, a vital point is the generality of the statement. Above, we assumed overdamped Langevin dynamics. As presented in the Introduction in Ch. 1, this assumption is often reasonable. However, due to the limited number of observed degrees of freedom, many observations do not fall into this class because the projection of the full dynamics onto the experimentally observed degrees of freedom may introduce memory in the observed dynamics, yielding more complicated, typically non-Markovian motion. In particular, this observed motion is no longer described by a standard overdamped Lagevin equation. Strikingly, the TUR is still applicable since the definition of the projection can be incorporated in the generalized current, as we show in the following. This largely extends the validity of the TUR, since it then only requires some high-dimensional, unobserved dynamics to be governed by overdamped Langevin motion, which following the lines in the Introduction in Ch. 1 is a very reasonable assumption.

After establishing the validity for such projected Langevin dynamics, we also exclude another source of error, namely the uncontrolled occurrence of different, uncoupled Langevin dynamics in the observation. Consider the scenario that trajectories of some particles are measured, where the particles may belong to different species that obey different dynamics, but the species are not distinguishable in the measurement. Then, we show that if the TUR applies to the individual species, it will also apply to the mixed observation of the indistinguishable species. This means that such a, possibly unavoidable, complication does not impede the application of the TUR.

We now provide short unpublished, possibly original, proofs of the above statements. We here consider the original TUR (4.1.1) but the statements directly extend to the TUR beyond steady-state dynamics that will be shown in Subsec. 4.3.1.

Generalization to projected dynamics

For Markov jump processes, it is well known that the TUR also gives a lower bound for the full dissipation even if only parts of the system are observed, see, e.g., Ref. [132]. For Langevin dynamics, we give a short proof here.

For a projection $\vec{f}\colon\ \mathbb{R}^d \to \mathbb{R}^m$ of the full dynamics in $\mathbb{R}^d$ to the observed dynamics in $\mathbb{R}^m$ with $m \leq d$ (denote m-dimensional vectors with arrows, and matrices and d-dimensional vectors as bold) we have (transforming the integral from $\circ\mathrm{d}\vec{f}(\mathbf{x}_\tau)$ to $\circ\mathrm{d}\mathbf{x}_\tau$ as in [16, Eq. (3.65)])

$$\begin{aligned}\int_0^t \vec{V}(q_\tau)\cdot\circ\mathrm{d}\vec{q}_\tau &= \int_0^t \vec{V}(\vec{f}(\mathbf{x}_\tau))\circ\mathrm{d}\vec{f}(\mathbf{x}_\tau)\\ &= \int_0^t \vec{V}(\vec{f}(\mathbf{x}_\tau))\cdot\mathbf{Jac}_f(\mathbf{x}_\tau)\circ\mathrm{d}\mathbf{x}_\tau\\ &= \int_0^t \mathbf{U}(\mathbf{x}_\tau)\cdot\circ\mathrm{d}\mathbf{x}_\tau \qquad (4.1.2)\end{aligned}$$

where $\mathbf{U} \equiv \vec{V}(\vec{f})\cdot\mathbf{Jac}_f\colon\ \mathbb{R}^d \to \mathbb{R}^d$ consists of $\vec{f}\colon\ \mathbb{R}^d \to \mathbb{R}^m$ and the Jacobian matrix $\mathbf{Jac}_f\colon \mathbb{R}^d \to \mathbb{R}^{m\times d}$. This rewriting in terms of $\mathbf{U}$ shows that the current $\int_0^t \vec{V}(q_\tau)\cdot\circ\mathrm{d}\vec{q}_\tau = \int_0^t \mathbf{U}(\mathbf{x}_\tau)\cdot\circ\mathrm{d}\mathbf{x}_\tau$ is of the form for which TUR in the full, d-dimensional space applies. Therefore, the TUR also holds for an observation of a projected current, but the dissipation in the formula remains the total dissipation of the full dynamics. Thus, it is possible to obtain bounds on the full dissipation from the limited observation. Note, however, that it is of course possible that the TUR only gives very poor bounds, e.g., if the relevant dissipative parts of the system are not observed. In particular, no non-trivial bound for the entropy production can be obtained if the mean of the projected current vanishes.

Generalization to indistinguishable species

Consider distinct, non-interacting species denoted by an index i that follow different steady-state Langevin dynamics, but are experimentally indistinguishable. For each species i, the TUR $\Sigma_t^i \mathrm{var}_i(J_t) \geq 2k_\mathrm{B}T\langle J_t\rangle_i^2$ applies. Assume that in a measurement, the experimentally indistinguishable species are sampled randomly with probability p_i, respectively. If the measurement is repeated very often, we measure a mean current $\langle J_t\rangle = \sum_i p_i\langle J_t\rangle_i$ (where $\langle\cdots\rangle_i$ denotes the average with respect to the i-th species' dynamics), while the

total dissipation per particle of the mixture of species reads $\Sigma_t = \sum_i p_i \Sigma_t^i$. The measured variance obeys[3] $\mathrm{var}(J_t) = \sum_i p_i \langle J_t^2 \rangle_i - \langle J_t \rangle^2 = \sum_i p_i \langle (J_t - \langle J_t \rangle)^2 \rangle_i \geq \sum_i p_i \langle (J_t - \langle J_t \rangle_i)^2 \rangle_i = \sum_i p_i \mathrm{var}_i(J_t)$. We now show that the TUR remains valid in an operationally accessible sense in this scenario, i.e., we show that, we have $\Sigma_t \mathrm{var}(J_t) \geq 2k_\mathrm{B}T \langle J_t \rangle^2$ also for this independent mixture.

This can be derived as follows. From the Cauchy-Schwarz inequality with $\Sigma_t = \sum_i p_i \Sigma_t^i$ and $\sum_i p_i \mathrm{var}_i(J_t) \leq \mathrm{var}(J_t)$ (see above), we have

$$\left[\sum_i p_i \sqrt{\mathrm{var}_i(J_t) \Sigma_t^i} \right]^2 \leq \left[\sum_i p_i \mathrm{var}_i(J_t) \right] \Sigma_t \leq \mathrm{var}(J_t) \Sigma_t \,. \tag{4.1.3}$$

Using $|\langle J_t \rangle| \leq \sum_i p_i \, |\langle J_t \rangle_i|$, averaging over the individual TURs $\sqrt{2k_\mathrm{B}T} \, |\langle J_t \rangle_i| \leq \sqrt{\Sigma_t^i \mathrm{var}_i(J_t)}$, and using Eq. (4.1.3) yields the desired TUR,

$$2k_\mathrm{B}T \langle J_t \rangle^2 \leq \left(\sum_i p_i \sqrt{2k_\mathrm{B}T} \, |\langle J_t \rangle_i| \right)^2 \leq \left[\sum_i p_i \sqrt{\Sigma_t^i \mathrm{var}_i(J_t)} \right]^2 \leq \mathrm{var}(J_t) \Sigma_t \,. \tag{4.1.4}$$

As apparent from the individual steps, the inequality (4.1.4) becomes saturated if all species are the same, and if the saturation criteria for a single species are fulfilled (see Subsecs. 4.3.2 and 4.5.6).

The proof in Eqs. (4.1.3) and (4.1.4) can be performed in an even shorter way using $\sum_i p_i a_i^2 / b_i \geq (\sum_i p_i a_i)^2 / \sum_{i'} p_{i'} b_{i'}$ for $b_i > 0$, which is known as *Sedrakyan's inequality* or *Titu's lemma*, and is a direct consequence of the Cauchy-Schwarz inequality. Together with $\mathrm{var}(J_t) \geq \sum_i p_i \mathrm{var}_i(J_t)$ (see above) and the individual TURs $\mathrm{var}_i(J_t)/2k_\mathrm{B}T \geq \langle J_t \rangle_i^2 / \Sigma_t^i$, this gives

$$\frac{\mathrm{var}(J_t)}{2k_\mathrm{B}T} \geq \sum_i p_i \frac{\mathrm{var}_i(J_t)}{2k_\mathrm{B}T} \geq \sum_i p_i \frac{\langle J_t \rangle_i^2}{\Sigma_t^i} \geq \frac{\left(\sum_i p_i \langle J_t \rangle_i \right)^2}{\sum_{i'} p_{i'} \Sigma_t^{i'}} = \frac{\langle J_t \rangle^2}{\Sigma_t^i} \,. \tag{4.1.5}$$

[3] Here, we use that for random X and constant c we have $\langle (X-c)^2 \rangle \geq \langle (X - \langle X \rangle)^2 \rangle$, which can be shown by minimization of the left-hand side over c, i.e., from $\partial_c \langle (X-c)^2 \rangle = 2(c - \langle X \rangle) = 0$ if, and only if, $\langle X \rangle = c$, and $\partial_c^2 \langle (X-c)^2 \rangle = 2 > 0$. Note that the variances do not simply add up since we are not considering the sum of independent random variables, but instead a random sampling of different random variables.

4.2 Technical background

The technical background required to understand the entropy production and fluctuations of time-integrated currents was already developed in Sec. 2.9 and Ch. 3, respectively. The only remaining technical issue is the proof of the TUR which is sketched in the following.

4.2.1 Derivation of the TUR

While the first proof of the TUR used the framework of large deviation theory [48], later proofs of the TUR and its generalizations employed techniques ranging from bounds to the scaled cumulant generating function [49, 72, 126] to martingale [104] and Hilbert-space [230] techniques. The arguably most insightful proof for the TUR [70,71] employs concepts of information theory and statistics, namely the generalized Cramér-Rao inequality [231, 232]. To state this inequality, we define Fisher information of a random variable X with probability density $p_\theta(\mathbf{x})$ for some parameter θ as $\mathcal{I}(\theta) \equiv \langle[\partial_\theta \log p_\theta(\mathbf{x})]^2\rangle_\theta$ with the notation $\langle\cdots\rangle_\theta \equiv \int \mathrm{d}\mathbf{x} p_\theta(\mathbf{x})\cdots$. This measures the information that the random variable X contains about the parameter θ that is used to model its distribution. The generalized Cramér-Rao inequality [231, 232] then states that for a function $f(X)$ we have

$$\frac{\mathrm{var}_\theta(f(X))}{[\partial_\theta\langle f(X)\rangle_\theta]^2} \times \mathcal{I}(\theta) \geq 1\,. \tag{4.2.1}$$

With a specific choice of the artificial *tilt* of the dynamics parametrized by θ around the value θ_0 corresponding to the actual dynamics, one can map this inequality to the TUR (4.1.1) via $\mathrm{var}_{\theta_0}(f(X)) = \mathrm{var}(J_t)$ and $\partial_\theta\langle f(X)\rangle_\theta|_{\theta=\theta_0} = \langle J_t\rangle$, and by computing the Fisher information of the Onsager-Machlup path measure for the specific tilt $\mathcal{I}(\theta_0) = \Sigma_t/2k_\mathrm{B}T$ [70] (for Markov jump dynamics $\mathcal{I}(\theta_0) \leq \Sigma_t/2k_\mathrm{B}T$, see Subsec. 4.5.12 and [233]). By showing the analogy to the established Cramér-Rao inequality, this derivation yields some insight into the information theoretic foundations of the TUR. However, this proof involves the generalized Cramér-Rao inequality and the Onsager-Machlup path measure under an artificial tilt of the dynamics, and is therefore not a *direct* proof. In the publication in Sec. 4.5 we instead provide a direct proof, solely relying on the stochastic equation of motion and the Cauchy-Schwarz inequality, which has some advantages compared to existing proofs.

4.3 State of the art

4.3.1 Generalizations of the original TUR

The original TUR (4.1.1) applies to steady-state dynamics following overdamped Langevin motion or Markov jump processes. We showed above that Eq. (4.1.1) also applies under projections and for observations of indistinguishable species. Here, we outline several generalizations beyond Eq. (4.1.1), both within and beyond steady-state dynamics. Note that out of the vast literature on this topic, we here only consider genuine *generalizations*, in the sense that we only consider other versions of the TUR that contain to Eq. (4.1.1) as a special case.

Retaining the assumption of steady-state dynamics, a generalization to multidimensional currents is directly possible [70]. More importantly, there is even the possibility to include not only currents, but also densities of the form $\rho_t = \int_0^t V(\mathbf{x}_\tau)\mathrm{d}\tau$ to arrive at the correlation TUR [134]

$$\Sigma_t \,\mathrm{var}(J_t)\left[1-\chi^2_{J\rho}\right] \geq 2\langle J_t\rangle^2, \qquad \chi^2_{J\rho} \equiv \frac{\mathrm{cov}^2(J_t,\rho_t)}{\mathrm{var}(J_t)\mathrm{var}(\rho_t)}\,, \tag{4.3.1}$$

where the coefficient $\chi^2_{J\rho} \in [0,1]$ involving the covariance of current and density generally leads to a tighter inequality. This is an extremely powerful generalization since it is equally operationally accessible as the original TUR (ρ_t is just as easily obtained from observed trajectories as J_t), and the tighter inequality generally leads to better estimation of the dissipation (more on this in Subsec. 4.3.2 and in the publication in Sec. 4.5).

The TUR does not generalize to underdamped dynamics, as shown by a counterexample in Ref. [234]. It is, however, possible to generalize the TUR to overdamped dynamics that is not in a steady state. Note that in this case, Σ_t refers to the product of temperature and total entropy production $\Sigma_t = T\Delta S_{\mathrm{tot}}([0,t]) \neq T\Delta S_{\mathrm{med}}([0,t])$ (only in the steady state we have $\Delta S_{\mathrm{tot}}([0,t]) = \Delta S_{\mathrm{med}}([0,t])$).

For time-homogenous Langevin dynamics (i.e., no explicit time dependence in the drift and diffusion), the generalization beyond steady-state dynamics of the original TUR (4.1.1) with window function $\mathbf{U}(\mathbf{x}_\tau)$ without explicit time dependence reads [126]

$$\Sigma_t \mathrm{var}(J_t) \geq 2(t\partial_t\langle J_t\rangle)^2\,. \tag{4.3.2}$$

This simplifies to Eq. (4.1.1) for steady-state dynamics since there $\langle J_t \rangle$ scales linearly in time, see Eq. (3.2.2).

Finally, for one-dimensional space and additive diffusion, the TUR was extended to Langevin motion with explicitly time-dependent drift and diffusion, assuming a time dependence entering through a protocol speed v (i.e., a drift term $a(x_\tau, v\tau)$), where v has to be varied for experimental accessibility. In this form, the generalized TUR reads [49][4]

$$\Sigma_t \mathrm{var}(J_t^v) \geq 2[(t\partial_t - v\partial_v)\langle J_t^v \rangle]^2 \,. \tag{4.3.3}$$

More details on the different generalizations, and results that go beyond beyond the state of the art, will be presented in the publication reproduced in Sec. 4.5.

4.3.2 Saturation of the TUR

Facing an inequality, one naturally wonders where it becomes saturated, i.e., in which case or limit equality is approached. From a practical point of view, the saturation is important for the best inference of dissipation—the tighter the inequality, the closer the lower bound is to the actual value for the dissipation. We now summarize the state of the art for saturation of TURs. More insight into these statements, and generalizations beyond the state of the art, will be given in the publication reproduced in Sec. 4.5.

In general, saturation of Eq. (4.1.1) only occurs close to equilibrium, i.e., for small values of mean current and dissipation (the variance does not approach zero in equilibrium as we know from the previous chapter, see, e.g., $\Omega \to 0$ in Fig. 3.4.3). Note that we refer to saturation when the ratio of the left and right side of $\Sigma_t \mathrm{var}(J_t) \geq 2\langle J_t \rangle$ approaches 1, i.e., it does not generally suffice if both sides tend to zero, as always becomes the case close to equilibrium. Actual saturation additionally requires the correct choice of $\mathbf{U}$ in the definition of J_t, for details, look into the publication in Subsec. 4.5.6. Intuitively, saturation also requires that all dissipative parts of the system are sampled, since sampling only some parts can always only give lower bounds. Note that the actual knowledge of the correct window function requires knowledge of many details of the system, but since any window function gives a valid lower bound (given sufficient statistics), one can obtain good lower bounds for the total entropy production by guessing or estimating the optimal choice of window function.

[4]Note that the appearance of the result is slightly different compared to [49], since we here consider time-integrated instead of time-averaged currents.

For systems far from equilibrium, saturation for the correct choice of window function can be obtained in the limit of short trajectories [153,154]. More importantly, it was shown how to saturate the steady-state TUR *arbitrarily far from equilibrium* using the correlation TUR (4.3.1) [134]. So far, it remained elusive to saturate TURs beyond steady-state dynamics.

In the case of projected dynamics, see Subsec. 4.1.4, achieving saturation is even harder since the projection $\vec{f}(\mathbf{x}_\tau)$ is typically not known. Moreover, if dissipative parts of the system are hidden by the projection, saturation is generally impossible to achieve.

4.4 Summary of results

Before we turn to the publication in Sec. 4.5, we here outline the contribution beyond the state of the art and literature summarized in the previous sections.

The idea that led to the publication was the following: In the Ch. 3 we derived an exact equality for the variance of time-integrated currents. The TUR is an inequality where the mathematically most difficult part is the variance of time integrated currents. Therefore, we should be able to understand the TUR in terms of the previous methods and results.

This idea led to a rederivation of the TUR based on the same stochastic-calculus approach as pursued in Secs. 3.4-3.6. This shows that the TUR is an inherent consequence of the stochastic equation of motion, and therefore it is not necessary to resort to proofs involving concepts from, e.g., information theory (see Subsec. 4.2.1). The new derivation has several advantages due to its very direct character. The only step that turns equalities into inequalities is a single application of the Cauchy-Schwarz inequality, which allows to systematically investigate the saturation of the TUR based on the saturation of this single calculation step. Moreover, in the new proof, the extension to the correlation TUR becomes trivial and new generalizations come naturally, exceeding the state of the art for multidimensional, multiplicative noise, and most importantly for time-dependent window functions that turn out to be necessary for saturation for transient (i.e., non-stationary) dynamics. Moreover, we extend the correlation TUR to transient dynamics and, for the correct choices of window functions, achieve saturation also for transient dynamics (which was so far limited to steady-state dynamics [134]).

4.5 Direct route to thermodynamic uncertainty relations and their saturation **(Phys. Rev. Lett. 2023)**

This section is a slightly adapted version of the publication C. Dieball and A. Godec, "Direct route to thermodynamic uncertainty relations and their saturation", *Phys. Rev. Lett.* **130**, 087101 (2023).

My personal contribution as first author of the publication was in performing calculations, analyzing the results, and co-writing the manuscript.

Direct route to thermodynamic uncertainty relations and their saturation

Cai Dieball and Aljaž Godec

Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

Abstract

Thermodynamic uncertainty relations (TURs) bound the dissipation in non-equilibrium systems from below by fluctuations of an observed current. Contrasting the elaborate techniques employed in existing proofs, we here prove TURs directly from the Langevin equation. This establishes the TUR as an inherent property of overdamped stochastic equations of motion. In addition, we extend the transient TUR to currents and densities with explicit time dependence. By including current-density correlations we, moreover, derive a new sharpened TUR for transient dynamics. Our arguably simplest and most direct proof, together with the new generalizations, allows us to systematically determine conditions under which the different TURs saturate and thus allows for a more accurate thermodynamic inference. Finally, we outline the direct proof also for Markov jump dynamics.

Phys. Rev. Lett. **130**, 087101 (2023)

4.5.1 Introduction

A defining characteristic of non-equilibrium systems is a non-vanishing entropy production [28, 54, 104, 120, 122, 145, 187, 188] emerging during relaxation [73, 120–124], in the presence of time-dependent (e.g., periodic [49, 167, 211, 212, 235, 236]) driving, or in non-equilibrium steady states (NESS) [34, 47, 48, 83, 84, 106, 125, 127]. A detailed understanding of the thermodynamics of systems far from equilibrium is in particular required for unraveling the physical principles that sustain active, living matter [146, 237–240]. Notwithstanding its importance, the entropy production within a non-equilibrium system beyond the linear response is virtually impossible to quantify from experimental observations, as it requires detailed knowledge about all dissipative degrees of freedom.

A recent and arguably the most relevant method to infer a lower bound on the entropy production in an experimentally observed complex system is via the so-called thermodynamic uncertainty relation (TUR) [67, 68, 83, 84, 132, 133, 153, 154, 201, 241], which relates the

(time-accumulated) dissipation Σ_t to fluctuations of a general time-integrated current J_t. For overdamped systems in a NESS it reads [47, 48]

$$\frac{\Sigma_t}{k_{\mathrm{B}}T} \geq 2\frac{\langle J_t\rangle^2}{\mathrm{var}(J_t)}\,, \tag{4.5.1}$$

with variance $\mathrm{var}(J_t) \equiv \langle J_t^2\rangle - \langle J_t\rangle^2$ and thermal energy $k_{\mathrm{B}}T$, which will henceforth be dropped for convenience and replaced by the convention of energies measured in units of $k_{\mathrm{B}}T$. The TUR may be seen as the natural counterpart of the fluctuation-dissipation theorem [242] or a more precise formulation of the second law [134]. Notably, it may also be interpreted as gauging the "thermodynamic cost of precision" [152], and it was found to limit the temporal extent of anomalous diffusion [214].

Since its original discovery [47] and proof [48] for systems in a NESS, a large number of more or less general variants of the TUR were derived. In particular, for paradigmatic overdamped dynamics and Markov jump processes, such generalized TURs have been found for transient systems (i.e., nonstationary dynamics emerging, e.g., from non-steady-state initial conditions) in the absence [71, 126, 210] and presence of time-dependent driving [49, 211]. Moreover, an extension to state variables (which we will refer to as "densities") instead of currents has been formulated [49], and recently correlations of densities and currents have been incorporated to significantly sharpen and even saturate the inequality for steady-state systems [134]. Note, however, that the validity of the TUR is generally limited to overdamped dynamics, as it was shown to break down in systems with momenta [234].

Many different techniques have been employed to derive TURs, including large deviation theory [48, 132, 162, 242, 243], bounds to the scaled cumulant generating function [49, 72, 126], as well as martingale [104] and Hilbert-space [230] techniques. Most notably, the TUR has been derived as a consequence of the generalized Cramér-Rao inequality [70, 71] which is well known in information theory and statistics. However, while providing valuable insight, the proof via the Cramér-Rao inequality includes quantifying the Fisher information of the Onsager-Machlup path measure [70] and involves a dummy parameter that "tilts" the original dynamics. Thus, it may not be faithfully considered as being direct. In fact, the TUR and its generalizations seem to be an inherent property of overdamped stochastic dynamics and are thus, akin to quantum-mechanical uncertainty, expected to follow directly from the equations of motion.

Here we show that no elaborated concepts beyond the equations of motion are indeed required. Using only stochastic calculus and the well known Cauchy-Schwarz inequality, we prove various existing TURs (including the correlation TUR [134]) for time-homogeneous

Phys. Rev. Lett. **130**, 087101 (2023)

overdamped dynamics in continuous space directly from the Langevin equation. Thereby we both, unify and simplify, proofs of TURs. Moreover, we derive, for the first time, the sharper correlation TUR for transient dynamics without explicit time dependence. This improved TUR can be saturated arbitrarily far from equilibrium for any initial condition and duration of trajectories, which we illustrate with the example of a displaced harmonic trap. Our simple proof offers several advantages and we therefore believe that it deserves attention even in cases that have already been proven before. Most notably, it enables immediate insight into how one can saturate the various TURs and allows for easy generalizations. Beyond the results for overdamped dynamics, we illustrate the analogous direct proof of the steady-state TUR also for Markov jump dynamics.

Phys. Rev. Lett. 130, 087101 (2023)

4.5.2 Setup

We consider d-dimensional[5] time-homogeneous (i.e., coefficients do not explicitly depend on time) overdamped dynamics described by the stochastic differential (Langevin) equation [6, 16]

$$\mathrm{d}\mathbf{x}_\tau = \mathbf{F}(\mathbf{x}_\tau)\mathrm{d}\tau + \boldsymbol{\sigma}(\mathbf{x}_\tau) \circledast \mathrm{d}\mathbf{W}_\tau \,, \tag{4.5.2}$$

where the anti-Itô product $\circledast$ assures thermodynamical consistency in the case of multiplicative noise[6] (i.e., space-dependent $\boldsymbol{\sigma}(\mathbf{x}_\tau)$) [84, 104, 169, 197, 198]. The choice of the product is irrelevant in the case of additive noise $\boldsymbol{\sigma}(\mathbf{x}_\tau) = \boldsymbol{\sigma}$. The increment $\mathrm{d}\mathbf{W}_\tau$ of the Wiener process, has zero mean $\langle \mathrm{d}\mathbf{W}_\tau \rangle = \mathbf{0}$ and is due to its covariance $\langle \mathrm{d}W_{\tau,i}\mathrm{d}W_{\tau',j}\rangle = \delta(\tau-\tau')\delta_{ij}\mathrm{d}\tau\mathrm{d}\tau'$ known as delta-correlated or white noise. The noise amplitude is related to the diffusion coefficient via $\mathbf{D}(\mathbf{x}) \equiv \boldsymbol{\sigma}(\mathbf{x})\boldsymbol{\sigma}(\mathbf{x})^T/2$ where $\boldsymbol{\sigma}$ and $\mathbf{D}$ are $d \times d$ matrices. Let $P(\mathbf{x},\tau)$ be the probability density to find $\mathbf{x}_\tau$ at a point $\mathbf{x}$ given some initial condition $P(\mathbf{x},0)$. Then the instantaneous probability density current $\mathbf{j}(\mathbf{x},\tau)$ is given by

$$\mathbf{j}(\mathbf{x},\tau) = [\mathbf{F}(\mathbf{x}) - \mathbf{D}(\mathbf{x})\nabla]P(\mathbf{x},\tau) \,, \tag{4.5.3}$$

and the Fokker-Planck equation [16, 17] for the time evolution of $P(\mathbf{x},\tau)$ follows from Eq. (4.5.2) and reads [6]

$$\partial_\tau P(\mathbf{x},\tau) = -\nabla \cdot \mathbf{j}(\mathbf{x},\tau) \,. \tag{4.5.4}$$

[5] We consider $\mathbb{R}^d$ or a finite subspace with periodic or reflecting boundary conditions.

[6] For more precise statements in the case of multiplicative noise in multidimensional space see Sec. 2.8, in particular the additional condition in Eq. (2.8.6).

In the special case that $\mathbf{F}(\mathbf{x})$ is sufficiently confining a NESS is eventually reached with invariant density $P_{\rm s}(\mathbf{x}) \equiv P(\mathbf{x}, \tau \to \infty)$ and steady-state current $\mathbf{j}_{\rm s}(\mathbf{x}) \equiv [\mathbf{F}(\mathbf{x}) - \mathbf{D}(\mathbf{x})\nabla]P_{\rm s}(\mathbf{x})$ with $\nabla \cdot \mathbf{j}_{\rm s}(\mathbf{x}) = 0$ [16]. The mean total (medium plus system) entropy production in the time interval $[0, t]$ is given by [28, 54]

$$\Sigma_t = \int \mathrm{d}\mathbf{x} \int_0^t \frac{\mathbf{j}^T(\mathbf{x}, \tau)\mathbf{D}^{-1}(\mathbf{x})\mathbf{j}(\mathbf{x}, \tau)}{P(\mathbf{x}, \tau)} d\tau \,. \tag{4.5.5}$$

Let J_t be a generalized time-integrated current with some vector-valued $\mathbf{U}(\mathbf{x}, \tau)$ defined via the Stratonovich stochastic integral (only for $\mathbf{x}$-dependent $\mathbf{U}$ the convention matters)

$$J_t \equiv \int_{\tau=0}^{\tau=t} \mathbf{U}(\mathbf{x}_\tau, \tau) \cdot \circ \mathrm{d}\mathbf{x}_\tau \,. \tag{4.5.6}$$

Note that for any integrand $\mathbf{U}$ this current and its first two moments are readily obtained from measured trajectories $(\mathbf{x}_\tau)_{0 \le \tau \le t}$. Therefore, a TUR involving such J_t is "operationally accessible." For dynamics in Eq. (4.5.2) the current may be equivalently written as the sum of an Itô integral and a $\mathrm{d}\tau$ integral, $J_t = J_t^{\rm I} + J_t^{\rm II}$, with [84]

$$\begin{aligned} J_t^{\rm I} &\equiv \int_{\tau=0}^{\tau=t} \mathbf{U}(\mathbf{x}_\tau, \tau) \cdot \boldsymbol{\sigma}(\mathbf{x}_\tau) \mathrm{d}\mathbf{W}_\tau \\ J_t^{\rm II} &\equiv \int_0^t \big[\mathbf{U}(\mathbf{x}_\tau, \tau) \cdot \mathbf{F}(\mathbf{x}_\tau) + \nabla \cdot [\mathbf{D}(\mathbf{x}_\tau)\mathbf{U}(\mathbf{x}_\tau, \tau)]\big] \mathrm{d}\tau \\ &\equiv \int_0^t \mathcal{U}(\mathbf{x}_\tau, \tau) \mathrm{d}\tau \,. \end{aligned} \tag{4.5.7}$$

By the zero-mean and independence properties of the Wiener process $\langle J_t^{\rm I} \rangle = 0$ and thus $\langle J_t \rangle = \langle J_t^{\rm II} \rangle = \int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} \mathcal{U}(\mathbf{x}, \tau) P(\mathbf{x}, \tau)$. Integrating by parts and using Eq. (4.5.3) we obtain (see also [84])

$$\langle J_t \rangle = \int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x} \mathbf{U}(\mathbf{x}, \tau) \cdot \mathbf{j}(\mathbf{x}, \tau) \,. \tag{4.5.8}$$

The variance $\mathrm{var}(J_t)$ can in turn be computed from two-point densities [83–86], but is not required to prove TURs.

We now outline our direct proof of TURs. First, we rederive the classical TUR (4.5.1) and its generalization to transients [126], whereby we find a novel correction term that extends the validity of the transient TUR. Next we prove the TUR for densities [49] and thereafter the correlation-improved TUR [134], for the first time also for nonstationary dynamics. Finally, we explain how to saturate the various TURs and illustrate our findings with an

Phys. Rev. Lett. **130**, 087101 (2023)

example. The proof relies solely on the equation of motion Eq. (4.5.2) and implied Fokker-Planck equation (4.5.4), which is why we call the proof "direct."

4.5.3 Direct proof of TURs

The essence of the direct proof is fully contained in the following Eqs. (4.5.9)-(4.5.11). First, we require a scalar quantity A_t with zero mean and whose second moment yields the dissipation defined in Eq. (4.5.5), i.e., $\langle A_t^2\rangle = \Sigma_t/2$ (the factor "$1/2$" is introduced for convenience). Considering the "delta-correlated" property of $\mathrm{d}\mathbf{W}_\tau$ and $\mathbf{D} = \mathbf{D}^T = \boldsymbol{\sigma}(\mathbf{x})\boldsymbol{\sigma}(\mathbf{x})^T/2$ leads to the "educated guess"

$$A_t \equiv \int_{\tau=0}^{\tau=t} \frac{\mathbf{j}(\mathbf{x}_\tau,\tau)}{P(\mathbf{x}_\tau,\tau)} \cdot [2\mathbf{D}(\mathbf{x}_\tau)]^{-1}\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau\,, \tag{4.5.9}$$

where A_t cannot be inferred from trajectories since only $\mathrm{d}\mathbf{x}_\tau$ but not $\mathrm{d}\mathbf{W}_\tau$ is observed. A_t can be understood as the "purely random" part $\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau$ of the increment $\mathrm{d}\mathbf{x}_\tau$ weighted by the local velocity and inverse diffusion coefficient. Because $\langle A_t J_t^{\mathrm{I}}\rangle = \langle J_t\rangle$ and $\langle A_t\langle J_t\rangle\rangle = \langle A_t\rangle\langle J_t\rangle = 0$ we have

$$\langle A_t(J_t - \langle J_t\rangle)\rangle = \langle J_t\rangle + \langle A_t J_t^{\mathrm{II}}\rangle\,, \tag{4.5.10}$$

and the Cauchy-Schwarz inequality $\langle A_t(J_t - \langle J_t\rangle)\rangle^2 \leq \langle A_t^2\rangle \mathrm{var}(J_t)$ further yields

$$\frac{\Sigma_t}{2}\mathrm{var}(J_t) \geq \left[\langle J_t\rangle + \langle A_t J_t^{\mathrm{II}}\rangle\right]^2\,. \tag{4.5.11}$$

Compared to Eq. (4.5.10) the inequality (4.5.11) has the advantage that $\mathrm{var}(J_t)$ is operationally accessible and Σ_t (unlike A_t) has a clear physical interpretation.

To obtain the TUR we are left with evaluating $\langle A_t J_t^{\mathrm{II}}\rangle$, which involves the two-time correlation of $\mathrm{d}\mathbf{W}_\tau$ and $\mathrm{d}\tau'$ integrals in Eqs. (4.5.9) and (4.5.7), respectively. For times $\tau \geq \tau'$, this correlation vanishes due to the independence property of the Wiener process. However, non-trivial correlations occur for $\tau < \tau'$ because the probability density of $\mathbf{x}_{\tau'}$ depends on $\mathrm{d}\mathbf{W}_\tau$. We quantify these correlations including $\mathrm{d}\mathbf{W}_\tau$ by writing $\langle A_t J_t^{\mathrm{II}}\rangle$ as an average over the joint density to be at points $\mathbf{x}, \mathbf{x}+d\mathbf{x}, \mathbf{x}'$ at times $\tau < \tau + d\tau < \tau'$, respectively, and expanding

$$P(\mathbf{x}',\tau'|\mathbf{x}+d\mathbf{x},\tau+d\tau) = P(\mathbf{x}',\tau'|\mathbf{x},\tau) + \mathrm{d}\mathbf{x}\cdot\nabla_{\mathbf{x}}P(\mathbf{x}',\tau'|\mathbf{x},\tau) + \mathcal{O}(d\tau)\,. \tag{4.5.12}$$

Phys. Rev. Lett. 130, 087101 (2023)

Following this approach [84, 85] (or alternatively via Doob conditioning [104, 107, 200] as in Ref. [133]) one can formulate a general calculation rule that in this case reads (for details, see Supplemental Material in Subsec. 4.5.12)

$$\langle A_t J_t^{\mathrm{II}} \rangle = -\int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x}' \mathcal{U}(\mathbf{x}', \tau') \int_0^{\tau'} \mathrm{d}\tau \int \mathrm{d}\mathbf{x} P(\mathbf{x}', \tau' | \mathbf{x}, \tau) \nabla_{\mathbf{x}} \cdot \mathbf{j}(\mathbf{x}, \tau) \,. \tag{4.5.13}$$

For steady-state systems, we have $\nabla \cdot \mathbf{j}(\mathbf{x}, \tau) = \nabla \cdot \mathbf{j}_{\mathrm{s}}(\mathbf{x}) = 0$ and thus $\langle A_t J_t^{\mathrm{II}} \rangle = 0$, such that Eq. (4.5.11) immediately implies the original TUR in Eq. (4.5.1).

To generalize to transients, we use Eq. (4.5.4) $\nabla_{\mathbf{x}} \cdot \mathbf{j}(\mathbf{x}, \tau) = -\partial_\tau P(\mathbf{x}, \tau)$, integrate by parts twice (see Subsec. 4.5.12 for details), and define a second operationally accessible current

$$\widetilde{J}_t \equiv \int_{\tau=0}^{\tau=t} \tau \partial_\tau \mathbf{U}(\mathbf{x}_\tau, \tau) \cdot \circ \mathrm{d}\mathbf{x}_\tau \,, \tag{4.5.14}$$

to obtain

$$\langle A_t J_t^{\mathrm{II}} \rangle = (t\partial_t - 1)\langle J_t \rangle - \langle \widetilde{J}_t \rangle \,. \tag{4.5.15}$$

Thus, we have expressed the correlation $\langle A_t J_t^{\mathrm{II}} \rangle$ in terms of operationally accessible quantities. From this and Eq. (4.5.11), the TUR for general initial conditions and general time-homogeneous Langevin dynamics Eq. (4.5.2) reads

$$\Sigma_t \, \mathrm{var}(J_t) \geq 2 \left[t\partial_t \langle J_t \rangle - \langle \widetilde{J}_t \rangle \right]^2 . \tag{4.5.16}$$

The fact that the TUR for transient dynamics (4.5.16) follows from the original TUR (4.5.1) upon replacing $\langle J_t \rangle \to t\partial_t \langle J_t \rangle$ is well known [71, 210] and was first derived in continuous space in Ref. [126]. However, the novel correction term $\langle \widetilde{J}_t \rangle$ extends the validity of the TUR to currents with an explicit time dependence $\mathbf{U}(\mathbf{x}, \tau)$. We show below and in Fig. 4.5.1 that this additional freedom in choosing $\mathbf{U}$ is crucial for saturating the transient TUR under general conditions. To highlight that end-point derivative $t\partial_t$ and the correction term $\langle \widetilde{J}_t \rangle$ are strictly necessary, we provide explicit counterexamples (see Supplemental Material (SM) in Subsec. 4.5.12).

We note that Eq. (4.5.16) in one-dimensional space and for additive noise can be deduced from restricting the result in [49], where an explicit time dependence was introduced via a speed parameter v, to a time-homogeneous drift, translated to time-integrated currents, and noting that $v\partial_v U(x, v\tau) = \tau \partial_\tau U(x, v\tau)$. The form without the speed parameter has the advantage that the correction term $\langle \widetilde{J}_t \rangle$ is accessible from a single experiment, while the

Phys. Rev. Lett. **130**, 087101 (2023)

∂_v correction requires perturbing the speed of the experiment. However, the result in [49] even holds for an explicitly time-dependent drift.

Notably, generalizing this proof to explicitly time-dependent drift or diffusion, although probably possible, is *not* straightforward because it requires perturbing the dynamics (see [49]), and therefore all relevant information is no longer contained in a single equation of motion.

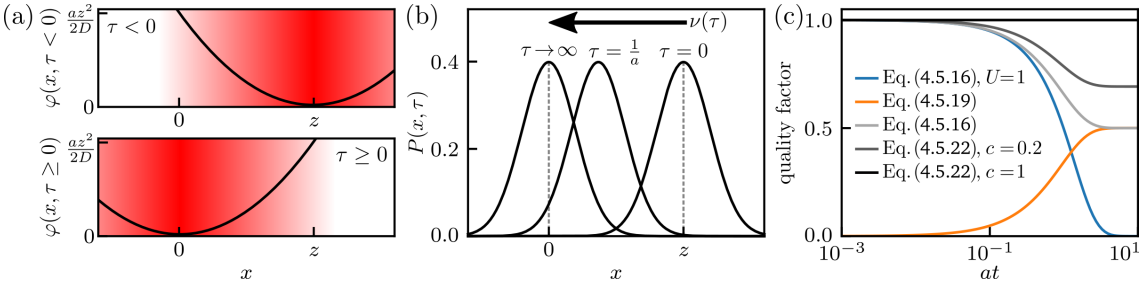

Figure 4.5.1: (a) Brownian particle in a one-dimensional harmonic trap with stiffness a, $\varphi(x,\tau) = a(x-x^0_\tau)^2/2D$ displaced from $x^0_{\tau<0} = z$ to $x^0_{\tau\geq 0} = 0$. Upon being initially equilibrated in $\varphi(x,\tau<0) = a(x-z)^2/2D$ (i.e., from the initial condition $p_0(x) \propto \exp[-a(x-z)^2/2D]$) the particle evolves for $\tau \geq 0$ due to $D\partial_x\varphi(x,\tau\geq 0) = ax$ according to $\mathrm{d}x_\tau = -ax_\tau d\tau + \sqrt{2D}\mathrm{d}W_\tau$ towards an equilibrium $p_{\tau\to\infty}(x) \propto \exp(-ax^2/2D)$. (b) Illustration of the evolution of $P(x,\tau)$ for $z = 5\sqrt{D/a}$. (c) Quality factors defined as the ratio of right- and left-hand side of the TURs as a function of the dimensionless quantity at. All quality factors turn out to be independent of z, D and only depend on a, t through at; explicit analytic expressions are given in Subsec. 4.5.12. Except for $J_t = \int 1 \circ \mathrm{d}x_\tau = x_t - x_0$ (blue line) we always choose the current defined with $U(\tau) = \nu(\tau)$ and density defined with $V(x,\tau) = -x\nu(\tau)$.

4.5.4 TUR for densities

We define general, operationally accessible densities (the term "density" is motivated by the analogy to "current" as e.g., in [43, 83, 84, 86])

$$\begin{aligned}\rho_t &= \int_0^t V(\mathbf{x}_\tau,\tau)\mathrm{d}\tau\,,\\ \widetilde{\rho}_t &\equiv \int_{\tau=0}^{\tau=t} \tau\partial_\tau V(\mathbf{x}_\tau,\tau)\mathrm{d}\tau\,.\end{aligned} \tag{4.5.17}$$

Since in the proof above we did not use the explicit form of $\mathcal{U}$, the density can be treated analogously to J_t in Eq. (4.5.7) by replacing $\mathcal{U} \to V$ and omitting the J^{I}_t term. Analogously to Eqs. (4.5.10) and (4.5.15) we thus obtain

$$\langle A_t(\rho_t - \langle\rho_t\rangle)\rangle = \langle A_t\rho_t\rangle = (t\partial_t - 1)\langle\rho_t\rangle - \langle\widetilde{\rho}_t\rangle\,, \tag{4.5.18}$$

Phys. Rev. Lett. 130, 087101 (2023)

and analogously to Eq. (4.5.11) the transient density-TUR

$$\Sigma_t \mathrm{var}(\rho_t) \geq 2\left[(t\partial_t - 1)\langle \rho_t \rangle - \langle \widetilde{\rho}_t \rangle\right]^2 . \tag{4.5.19}$$

Note that due to the absence of the J_t^{I} term, the right-hand side vanishes in steady-state systems. As in the discussion of Eq. (4.5.16) above, Eq. (4.5.19) is in some sense contained in the results of [49]. However, Eq. (4.5.19) allows for multidimensional space and multiplicative noise, and does not require a variation in protocol speed.

4.5.5 Improving TURs using correlations

It has been recently found [134] that the steady-state TUR can be eminently improved, and even saturated arbitrarily far from equilibrium, by considering correlations between currents and densities as defined in Eq. (4.5.17). To rederive this sharper version, we rewrite Eq. (4.5.11) for the observable $J_t - c\rho_t$ (the constant c is in fact technically redundant since it can be absorbed in the definition of ρ_t)

$$\frac{\Sigma_t}{2}\mathrm{var}(J_t - c\rho_t) \geq \left[\langle J_t \rangle + \langle A_t(J_t^{\mathrm{II}} - c\rho_t)\rangle\right]^2 . \tag{4.5.20}$$

Note that $\mathrm{var}(J_t - c\rho_t) = \mathrm{var}(J_t) + c^2\mathrm{var}(\rho_t) - 2c\,\mathrm{cov}(J_t, \rho_t)$, where cov denotes the covariance. Using the optimal choice $c = \mathrm{cov}(J_t, \rho_t)/\mathrm{var}(\rho_t)$ and recalling that for steady-state systems $\langle A_t(J_t^{\mathrm{II}} - c\rho_t)\rangle = 0$, Eq. (4.5.20) becomes the NESS correlation TUR in [134]

$$\begin{aligned} \Sigma_t\, \mathrm{var}(J_t)\left[1 - \chi^2_{J\rho}\right] &\geq 2\langle J_t \rangle^2, \\ \chi^2_{J\rho} &\equiv \frac{\mathrm{cov}^2(J_t, \rho_t)}{\mathrm{var}(J_t)\mathrm{var}(\rho_t)} . \end{aligned} \tag{4.5.21}$$

Since $\chi^2_{J\rho} \in [0, 1]$, Eq. (4.5.21) is sharper than Eq. (4.5.1) and, as proven in [134] and discussed below, for any steady-state system there exist J_t, ρ_t that saturate this inequality.

Our approach allows to generalize this result to transient dynamics by computing $\langle A_t(J_t^{\mathrm{II}} - c\rho_t)\rangle$ as in Eq. (4.5.15) to obtain from Eq. (4.5.20) the generalized correlation TUR

$$\Sigma_t\, \mathrm{var}(J_t - c\rho_t) \geq 2\left(t\partial_t\langle J_t \rangle - \langle \widetilde{J}_t \rangle - c\left[(t\partial_t - 1)\langle \rho_t \rangle - \langle \widetilde{\rho}_t \rangle\right]\right)^2 . \tag{4.5.22}$$

One could again optimize the left-hand side over c to obtain $\mathrm{var}(J_t - c\rho_t) = \mathrm{var}(J_t)\left[1 - \chi^2_{J\rho}\right]$. However, since here the right-hand side also involves c, this may not be the optimal choice. Thus, it is instead practical to keep c general (or absorb it into ρ_t). The generalized correlation TUR (4.5.22) represents a novel result that sharpens the tran-

Phys. Rev. Lett. **130**, 087101 (2023)

sient TUR in Eq. (4.5.16), and, as we show below and illustrate in Fig. 4.5.1, even allows us to generally saturate the TUR arbitrarily far from equilibrium.

4.5.6 Saturation of TURs

For any choice $\mathbf{U}$ in the definition of J_t in Eq. (4.5.6), the TUR allows to infer a lower bound on the time-accumulated dissipation Σ_t from $\langle J_t \rangle$ and $\mathrm{var}(J_t)$ [67, 68, 83, 84, 132, 153, 154, 201, 241]. The tighter the inequality, the more precise is the lower bound on Σ_t. It is therefore important to understand when the inequality becomes tight or even saturates, i.e., gives equality.

Because of the simplicity and directness of our proof, we can very well discuss the tightness of the bound based on the single application of the Cauchy-Schwarz inequality. As elaborated in the Appendix 4.5.11, this approach reproduces, and extends beyond, numerous existing results on asymptotic and exact saturation of TURs. Most importantly, choosing $\mathbf{U}(\mathbf{x}, \tau) = c'[\mathbf{j}(\mathbf{x}_\tau, \tau)/P(\mathbf{x}_\tau, \tau)] \cdot [2\mathbf{D}(\mathbf{x}_\tau)]^{-1}$ with arbitrary c' and $c\rho_t = J_t^{\mathrm{II}}$ [see Eq. (4.5.7)] gives $J_t - c\rho_t = J_t^{\mathrm{I}} = c' A_t$ which in turn implies equality in the Cauchy-Schwarz argument leading to the correlation TURs Eqs. (4.5.21) and (4.5.22). This directly implies exact saturation of the correlation TURs which was so far achieved only in the steady-state case [134]. Our generalization of the correlation TUR in Eq. (4.5.22) for transient systems therefore allows to saturate a TUR arbitrarily far from equilibrium for any t and for general initial conditions and general time-homogeneous dynamics in Eq. (4.5.2).

4.5.7 Example

To illustrate the novel results in Eqs. (4.5.16), (4.5.19) and (4.5.22) and the new insight into the saturation, we provide an explicit example of transient dynamics in Fig. 4.5.1, that of a Brownian particle in a one-dimensional harmonic potential $\varphi(x, t) = a(x - x_t^0)^2/2D$ displaced from $x_{\tau<0}^0 = z$ to $x_{\tau\geq 0}^0 = 0$, see Fig. 4.5.1a. This setting, illustrated by the color gradient in Fig. 4.5.1a, can easily be realized experimentally using optical tweezers [244–246]. The process features a Gaussian probability density $P(x, \tau)$ with constant variance D/a that moves with a space-independent velocity $\nu(\tau) = j(x, \tau)/P(x, \tau) = -az \exp(-a\tau)$ towards the equilibrium $\propto \exp(-ax^2/2D)$, see Fig. 4.5.1b.

To quantify the tightness of the respective TURs we inspect quality factors – the ratio of the right- and left-hand side of the TUR – shown in Fig. 4.5.1c as a function of the dimensionless quantity at. The blue line represents the transient TUR (4.5.16) for the current $J_t = x_t - x_0$ where $U(x, \tau) = 1$. Since this U does not feature explicit time

Phys. Rev. Lett. 130, 087101 (2023)

dependence, the correction term $\widetilde{J}_t$ does not contribute and the transient TUR from the existing literature [126] applies. The existing (as well as our) results allow to vary the spatial dependence of U but we refrain from considering this for simplicity and since it is not necessary for saturation (i.e., ν, D have no spatial dependence in our example). Due to the novel correction term in Eq. (4.5.16) we may choose a time-dependent U, and following our discussion of the saturation we choose for all following examples J_t with $U(\tau) = c'\nu(\tau)/2D = \nu(\tau)$ (the prefactor c' is arbitrary as it cancels in quality factor) and the corresponding $\rho_t = J_t^{II}$, i.e., with $V(x,\tau) = \mathcal{U}(x,\tau) = -axU(\tau)$, see Eq. (4.5.7). For this choice, we evaluate the transient current [Eq. (4.5.16)] and density-TUR [Eq. (4.5.19)], see light gray and orange line in Fig. 4.5.1c. Moreover, we evaluate the novel generalized correlation TUR (4.5.22) for $c = 0.2$ (dark gray line), where we find that the current TUR is improved by considering correlations with a density, and for $c = 1$ (black line), where we find the expected saturation. This saturation means that the lower bound obtained for Σ_t from this TUR is exactly Σ_t. Note that this exact saturation requires the knowledge of the details of the dynamics for the choice of U, V. However, even with very limited knowledge one can simply consider different guesses or approximations of the optimal U, V and each guess will give a valid lower bound (given sufficient statistics).

Phys. Rev. Lett. **130**, 087101 (2023)

4.5.8 Direct route for Markov jump processes

Beyond overdamped dynamics, one may employ the above direct approach for deriving TURs to Markov jump dynamics on a discrete state-space $\mathcal{N}$ with jump-rates $(r_{xy})_{x,y\in\mathcal{N}}$ and steady-state distribution $(p_x)_{x\in\mathcal{N}}$. To illustrate this generalization, we here provide the proof of the steady-state TUR (4.5.1). Let $\hat{\tau}_x$ denote the (random) time spent in state x and $\hat{n}_{xy}$ the (random) number of jumps from x to y in the time interval $[0,t]$. A general time-accumulated current in a jump process is defined with antisymmetric prefactors $d_{xy} = -d_{yx}$ as the double sum $J \equiv \sum_{x\neq y} d_{xy}\hat{n}_{xy}$. The steady-state dissipation in turn reads $\Sigma \equiv t\sum_{x\neq y} p_x r_{xy} \ln[p_x r_{xy}/p_y r_{yx}]$. Analogously to A_t in Eq. (4.5.9) define

$$A \equiv \sum_{x\neq y} \frac{p_x r_{xy} - p_y r_{yx}}{p_x r_{xy} + p_y r_{yx}} (\hat{n}_{xy} - \hat{\tau}_x r_{xy})\,. \tag{4.5.23}$$

For this choice of A one can check that $\langle A\rangle = 0$, $\langle A^2\rangle \leq \Sigma/2$, and $\langle AJ\rangle = \langle J\rangle$ (a "direct" proof as above follows by analogy of covariance properties of $\partial_t(\hat{n}_{xy} - \hat{\tau}_x r_{xy})$ and $\boldsymbol{\sigma}(\mathbf{x}_t)\mathrm{d}\mathbf{W}_t$,

see Subsec. 4.5.12 for details) which imply, via the Cauchy-Schwarz inequality, equivalently to Eqs. (4.5.10) and (4.5.11) the steady-state TUR for Markov jump processes

$$\langle A(J - \langle J \rangle) \rangle = \langle J \rangle \quad \Rightarrow \quad \frac{\Sigma}{2} \mathrm{var}(J) \geq \langle J \rangle^2 \,. \tag{4.5.24}$$

A discussion of possible generalizations of this proof beyond steady-state dynamics is given in Subsec. 4.5.12.

4.5.9 Conclusion

Using only stochastic calculus and the well known Cauchy-Schwarz inequality, we proved various existing TURs directly from the Langevin equation. This underscores the TUR as an inherent property of overdamped stochastic equations of motion, analogous to quantum-mechanical uncertainty relations. Moreover, by including current-density correlations, we derived a new sharpened TUR for transient dynamics. Based on our simple and more direct proof, we were able to systematically explore conditions under which TURs saturate. The new equality (4.5.10) is mathematically even stronger than TUR (4.5.11). Therefore, it allows us to derive further bounds, e.g., by applying Hölder's instead of the Cauchy-Schwarz inequality which, however, may not yield operationally accessible quantities. Our approach may allow for generalizations to systems with time-dependent driving (see, e.g., [49]) which, however, are not expected to follow anymore directly from a single equation of motion. The novel correction term for currents with explicit time dependence as well as the new transient correlation TUR and its saturation are expected to equally apply to Markov jump processes by generalizing the approach illustrated in Eqs. (4.5.23) and (4.5.24).

Phys. Rev. Lett. **130**, 087101 (2023)

4.5.10 Acknowledgments

We thank David Hartich for insightful suggestions and critical reading of the manuscript. Financial support from Studienstiftung des Deutschen Volkes (to C. D.) and the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2 (to A. G.) is gratefully acknowledged.

4.5.11 Appendix: Saturation of TURs

Thanks to the directness of our proof, we only need to discuss the tightness based on the step from Eq. (4.5.10) to Eq. (4.5.11) where we applied the Cauchy-Schwarz inequality

$\langle A_t(J_t - \langle J_t\rangle)\rangle^2 \leq \langle A_t^2\rangle \mathrm{var}(J_t)$ to the exact Eq. (4.5.10). Thus, the closer A_t and $J_t - \langle J_t\rangle$ are to being linearly dependent (recall that the Cauchy-Schwarz inequality measures the angle φ between two vectors $(\vec{x}\cdot\vec{y})^2 = \vec{x}^2\vec{y}^2\cos^2(\varphi) \leq \vec{x}^2\vec{y}^2$), the tighter the TUR, with saturation for $J_t - \langle J_t\rangle = c' A_t$ for some constant c'. Therefore, the TUR is expected to be tightest for the choice $\mathbf{U}(\mathbf{x},\tau) = c'[\mathbf{j}(\mathbf{x}_\tau,\tau)/P(\mathbf{x}_\tau,\tau)]\cdot[2\mathbf{D}(\mathbf{x}_\tau)]^{-1}$ for which $J_t^{\mathrm{I}} = c' A_t$ (see Eq. (4.5.7)). Note that for NESS this $\mathbf{U}$ becomes time independent with $\mathbf{j}_{\mathrm{s}}(\mathbf{x})/P_{\mathrm{s}}(\mathbf{x})$. This choice is known to saturate the original TUR in Eq. (4.5.1) in the near-equilibrium limit [104]. However, since the full $J_t = J_t^{\mathrm{I}} + J_t^{\mathrm{II}}$ current cannot be chosen to exactly agree with $c' A_t$, equality is generally not reached.

The original TUR (4.5.1) with this choice of $\mathbf{U}(\mathbf{x},\tau)$ was also found to saturate in the short-time limit $t \to 0$ [153, 154]. This result is in turn reproduced with our approach by noting that $J_t^{\mathrm{I}} = c' A_t$ and $\langle A_t J_t^{\mathrm{II}}\rangle = 0$ give $\langle A_t(J_t - \langle J_t\rangle)\rangle^2 = \langle A_t J_t^{\mathrm{I}}\rangle^2 = \langle A_t^2\rangle\langle J_t^{\mathrm{I}2}\rangle$, and in the limit $t \to 0$ the integrals in Eq. (4.5.7) asymptotically scale like a single time step, such that $\langle J_t^{\mathrm{I}2}\rangle \sim (\mathbf{W}_t - \mathbf{W}_0)^2 \sim t$ dominates all $\sim t^{3/2}$, $\sim t^2$ contributions in $\mathrm{var}(J_t)$. In turn, $\langle J_t^{\mathrm{I}2}\rangle \overset{t\to 0}{\to} \mathrm{var}(J_t)$ which yields $\langle A_t(J_t - \langle J_t\rangle)\rangle^2 \overset{t\to 0}{\to} \langle A_t^2\rangle \mathrm{var}(J_t)$. Thus, the Cauchy-Schwarz step from the equality (4.5.10) to the inequality (4.5.11) saturates as $t \to 0$, in turn implying that the TUR saturates.

More recently, it was also found that including correlations [see Eq. (4.5.21) and Ref. [134]] allows us to saturate a sharpened TUR for steady-state systems arbitrarily far from equilibrium for any t, again for the same choice $\mathbf{U}(\mathbf{x},\tau)$ as above. Since our rederivation of the NESS correlation TUR in Eq. (4.5.21) applied the Cauchy-Schwarz inequality to A_t and $J_t - c\rho_t$ we see that choosing $c\rho_t = J_t^{\mathrm{II}}$ yields $J_t - c\rho_t = J_t^{\mathrm{I}} = c' A_t$, such that the application of the Cauchy-Schwarz inequality becomes an equality. That is, the correlation TUR (4.5.21) for this choice of J_t and ρ_t is generally saturated. Notably, this powerful result follows very naturally from the direct proof presented here.

Our generalization of the correlation TUR in Eq. (4.5.22) for transient systems even allows to saturate a TUR (arbitrarily far from equilibrium for any t and) for general initial conditions and general time-homogeneous dynamics in Eq. (4.5.2). This result is strong but obvious, since as for the NESS correlation TUR we can choose J_t and ρ_t such that $J_t - c\rho_t = c' A_t$. Note that it is here crucial that we allowed for an explicit time dependence in $\mathbf{U}$ and V, i.e., that we found new correction terms [terms with tilde in Eqs. (4.5.16), (4.5.19), and (4.5.22)].

Phys. Rev. Lett. **130**, 087101 (2023)

4.5.12 Supplemental Material

In this Supplemental Material, we provide a detailed analysis of the quality (i.e., sharpness) of the distinct versions of thermodynamic uncertainty relation (TUR) applied to the transient example shown above as well as counterexamples underscoring the necessity of the novel versions of the TUR. Moreover, we provide technical details on the direct proof of the TUR as well as a perspective on the extension of the direct proof to Markov-jump processes.

Phys. Rev. Lett. 130, 087101 (2023)

Quality of TURs for displaced harmonic trap

In the following, we derive the quality factors (i.e., the sharpness of the various TURs) shown in Fig. 4.5.1. Consider one-dimensional Brownian motion in a parabolic potential (i.e., a Langevin equation with linear force; known as the Ornstein-Uhlenbeck process) [6] with a Gaussian initial condition x_0 [we denote a normal distribution by $\mathcal{N}(\text{mean, variance})$],

$$\mathrm{d}x_\tau = -ax_\tau \mathrm{d}\tau + \sqrt{2D}\mathrm{d}W_\tau\,, \qquad x_0 \sim \mathcal{N}(z, \sigma_0^2)\,. \tag{4.5.25}$$

Even though this process approaches an equilibrium steady state, for finite times it features transient dynamics if x_0 is not sampled from the steady-state distribution. For any Gaussian initial condition, this process is Gaussian [6]. Therefore, the mean and the variance completely determine the distribution of x_τ. The harmonic potential and the Gaussian initial condition can be realized experimentally by optical tweezers. The mean, variance and covariance are simply obtained as (see, e.g., Appendix F in Ref. [90])

$$\begin{aligned} \langle x_\tau\rangle &= z\mathrm{e}^{-a\tau}\,, \\ \mathrm{var}(x_\tau) \equiv \langle x_\tau^2\rangle - \langle x_\tau\rangle^2 &= \frac{D}{a}\left(1-\mathrm{e}^{-2a\tau}\right) + \sigma_0^2\mathrm{e}^{-2a\tau}\,, \\ \text{For } \tau \geq \tau'\text{:}\quad \mathrm{cov}(x_\tau, x_{\tau'}) \equiv \langle x_\tau x_{\tau'}\rangle - \langle x_\tau\rangle\langle x_{\tau'}\rangle &= \mathrm{e}^{-a(\tau-\tau')}\mathrm{var}(x_{\tau'})\,. \end{aligned} \tag{4.5.26}$$

The Gaussian probability density $P(x,\tau)$ given the initial condition in Eq. (4.5.25) accordingly reads

$$P(x,\tau) = \sqrt{\frac{1}{2\pi\mathrm{var}(x_\tau)}}\exp\left[-\frac{(x-z\mathrm{e}^{-a\tau})^2}{2\mathrm{var}(x_\tau)}\right]. \tag{4.5.27}$$

The local mean velocity $\nu(x,\tau) \equiv j(x,\tau)/P(x,\tau)$ with current $j(x,\tau) \equiv (-ax - D\partial_x)P(x,\tau)$ reads

$$\nu(x,\tau) = -ax + D(x - z\mathrm{e}^{-a\tau})/\mathrm{var}(x_\tau)\,. \tag{4.5.28}$$

For this example we consider the simple case $\sigma_0^2 = D/a$, i.e., we start in the steady-state variance (but as long as $z \neq 0$ not in the steady-state distribution; can be realized by equilibration with optical tweezers at $x = z$ at times $\tau < 0$ and an equally stiff optical trap at position $x = 0$ at times $\tau \geq 0$), for which we obtain the simplified expressions

$$\begin{aligned} \mathrm{var}(x_\tau) &= D/a\,, \\ \text{For } \tau \geq \tau' \text{:} \quad \mathrm{cov}(x_\tau, x_{\tau'}) &= \mathrm{e}^{-a(\tau-\tau')}D/a\,, \\ \nu(x,\tau) &= -az\mathrm{e}^{-a\tau}\,. \end{aligned} \tag{4.5.29}$$

For this initial condition, $P(x,\tau)$ corresponds to a Gaussian distribution of constant variance with mean value $z\mathrm{e}^{-a\tau}$ drifting from z to 0. Since only the mean changes (but the distribution around the mean remains invariant), the local mean velocity $\nu(x,\tau)$ is independent of x [and in fact given by the velocity of the mean $\nu(x,\tau) = \partial_\tau\langle x_\tau\rangle$]. This easily allows to compute the time-accumulated dissipation

$$\Sigma_t = D^{-1}\int_0^t \mathrm{d}\tau \int \mathrm{d}x \langle\nu(x_\tau,\tau)\rangle^2 = \frac{a^2z^2}{D}\int_0^t \mathrm{d}\tau \mathrm{e}^{-2a\tau} = \frac{az^2}{2D}(1-\mathrm{e}^{-2at})\,. \tag{4.5.30}$$

Apart from Σ_t the TURs contain first and second moments of (generalized) currents and densities that we derive for some examples below. Recall the definition of a generalized current (here in one-dimensional space)

$$J_t \equiv \int_{\tau=0}^{\tau=t} U(x_\tau,\tau)\circ \mathrm{d}x_\tau\,. \tag{4.5.31}$$

For simplicity (and since in our example the mean velocity $\nu(\tau)$ is space-independent) we consider only currents without explicit space dependence, i.e., only $U(x_\tau,\tau) = U(\tau)$. In this case, there is no difference between the Stratonovich and Itô interpretation of the integral, i.e., $J_t = \int_{\tau=0}^{\tau=t} U(\tau)\mathrm{d}x_\tau$.

Phys. Rev. Lett. **130**, 087101 (2023)

Current without explicit time dependence

For the simplest case of $U(x,\tau)=1$ we have the displacement current (denote this choice of current by J_t^x)

$$J_t^x \equiv \int_{\tau=0}^{\tau=t} 1 \circ \mathrm{d}x_\tau = x_t - x_0\,. \tag{4.5.32}$$

From Eq. (4.5.29) we obtain

$$\begin{aligned} \langle J_t\rangle &= -z(1-\mathrm{e}^{-at})\,, \qquad t\partial_t\langle J_t\rangle = -zat\mathrm{e}^{-at}\,,\\ \mathrm{var}(J_t) &= \mathrm{var}(x_t)+\mathrm{var}(x_0)-2\mathrm{cov}(x_t,x_0) = \frac{2D}{a}(1-\mathrm{e}^{-at})\,. \end{aligned} \tag{4.5.33}$$

Recalling the expression for the dissipation Eq. (4.5.30) the transient TUR $\Sigma_t\mathrm{var}(J_t)\geq 2[t\partial_t\langle J_t\rangle]^2$ in this example reads $\frac{az^2}{2D}(1-\mathrm{e}^{-2at})\frac{2D}{a}(1-\mathrm{e}^{-at})\geq 2(zat)^2\mathrm{e}^{-2at}$ such that the quality factor $Q_x\in[0,1]$ (ratio of right-hand side and left-hand side; measures sharpness of the inequality) in this example becomes

$$Q_x \equiv \frac{2[t\partial_t\langle J_t\rangle]^2}{\Sigma_t\mathrm{var}(J_t)} = \frac{2(at)^2\mathrm{e}^{-2at}}{(1-\mathrm{e}^{-2at})\,(1-\mathrm{e}^{-at})}\,. \tag{4.5.34}$$

Note that Q_x is independent of z and D and only depends on the dimensionless quantity at. For large values, $at\to\infty$ we have $Q_x\to 0$, see also Fig. 4.5.1.

Currents with explicit time dependence

Saturating a TUR (i.e., obtaining the true dissipation as the lower bound inferred by the TUR) can be achieved for the transient correlation-TUR [Eq. (4.5.22)] by choosing $\mathbf{U}(\mathbf{x},\tau) = c'[\mathbf{j}(\mathbf{x}_\tau,\tau)/P(\mathbf{x}_\tau,\tau)]\cdot[2\mathbf{D}(\mathbf{x}_\tau)]^{-1}$, and the corresponding density $\rho_t = \int_0^t\left[\mathbf{U}(\mathbf{x}_\tau,\tau)\cdot\mathbf{F}(\mathbf{x}_\tau)+\nabla\cdot[\mathbf{D}(\mathbf{x}_\tau)\mathbf{U}(\mathbf{x}_\tau,\tau)]\right]\mathrm{d}\tau$. For this example, the respective current and density (denote this choice by superscript ν) read (Stratonovich convention irrelevant since $\partial_x U(x,\tau)=0$ here)

$$J_t^\nu \equiv \frac{c'}{2D}\int_{\tau=0}^{\tau=t}\nu(\tau)\mathrm{d}x_\tau\,, \qquad \rho_t^\nu \equiv \frac{-ac'}{2D}\int_0^t x_\tau\nu(\tau)\mathrm{d}\tau\,, \tag{4.5.35}$$

with $\nu(\tau)=-az\mathrm{e}^{-a\tau}$, see Eq. (4.5.29). The prefactor c' will equally appear on both sides of the TURs and therefore not change the quality factors. One may set $c'=2D$ as done above, but here we keep it general. Plugging in $\mathrm{d}x_\tau$ from Eq. (4.5.25) we calculate

$$\langle J_t^\nu\rangle = \frac{c'}{2D}\int_0^t\nu(\tau)\langle -ax_\tau\rangle\mathrm{d}\tau = \frac{-azc'}{2D}\int_0^t\nu(\tau)\mathrm{e}^{-a\tau}\mathrm{d}\tau = \frac{(az)^2c'}{2D}\int_0^t\mathrm{e}^{-2a\tau}\mathrm{d}\tau$$

Phys. Rev. Lett. 130, 087101 (2023)

$$= \frac{az^2c'}{4D}(1-\mathrm{e}^{-2at})\,,$$
$$t\partial_t\langle J_t^\nu\rangle = t\frac{(az)^2c'}{2D}\mathrm{e}^{-2at}\,, \qquad \langle\rho_t^\nu\rangle = \langle J_t^\nu\rangle\,, \qquad t\partial_t\langle\rho_t^\nu\rangle = t\partial_t\langle J_t^\nu\rangle\,. \tag{4.5.36}$$

The auxiliary current $\widetilde{J}_t \equiv \int_{\tau=0}^{\tau=t}\tau\partial_\tau U(x_\tau,\tau)\circ\mathrm{d}x_\tau$ and density $\widetilde{\rho}_t$ have due to $\tau\partial_\tau\nu(\tau) = -a\tau\nu(\tau)$ the mean

$$\langle\widetilde{\rho}_t^\nu\rangle = \langle\widetilde{J}_t^\nu\rangle = \frac{-a(az)^2c'}{2D}\int_0^t\tau\mathrm{e}^{-2a\tau}\mathrm{d}\tau = \frac{-a(az)^2c'}{2D}\,\frac{1-\mathrm{e}^{-2at}(1+2at)}{4a^2}$$
$$= \frac{-az^2c'}{8D}[1-(1+2at)\mathrm{e}^{-2at}]\,. \tag{4.5.37}$$

For the variance split $J_t^\nu = J_t^{\mathrm{I}} + J_t^{\mathrm{II}}$ with

$$J_t^{\mathrm{I}} = \frac{c'}{2D}\int_{\tau=0}^{\tau=t}\nu(\tau)\sqrt{2D}\mathrm{d}W_\tau\,, \qquad J_t^{\mathrm{II}} = \rho_t^\nu\,, \tag{4.5.38}$$

and compute

$$\mathrm{var}(\rho_t^\nu) = \mathrm{var}(J_t^{\mathrm{II}}) = \frac{a^2c'^2}{4D^2}\left\langle\left(\int_0^t\nu(\tau)(\mathbf{x}_\tau-\langle\mathbf{x}_\tau\rangle)\mathrm{d}\tau\right)^2\right\rangle$$
$$= \frac{a^2c'^2}{4D^2}2\int_0^t\mathrm{d}\tau\int_0^\tau\mathrm{d}\tau'\,\nu(\tau)\nu(\tau')\mathrm{cov}(x_\tau,x_{\tau'})$$
$$= \frac{a^3z^2c'^2}{2D}\int_0^t\mathrm{d}\tau\int_0^\tau\mathrm{d}\tau'\mathrm{e}^{-a(\tau+\tau')}\mathrm{e}^{-a(\tau-\tau')}$$
$$= \frac{a^3z^2c'^2}{2D}\int_0^t\mathrm{d}\tau\tau\mathrm{e}^{-2a\tau}$$
$$= \frac{c'^2az^2}{8D}\left[1-(1+2at)\mathrm{e}^{-2at}\right]\,. \tag{4.5.39}$$

The cross terms are given by the non-trivial correlations of $\mathrm{d}W_\tau$ and $\mathrm{d}\tau'$ integrals exactly as Eq. (4.5.13), in this example with $\mathcal{U}(x',\tau') = -ax'\nu(\tau')c'/(2D)$, such that

$$\langle J_t^{\mathrm{I}}J_t^{\mathrm{II}}\rangle = -\int_0^t\mathrm{d}\tau'\int\mathrm{d}x'\mathcal{U}(x',\tau')\int_0^t\mathrm{d}\tau\mathbb{1}_{\tau<\tau'}\int\mathrm{d}xP(x',\tau'|x,\tau)\partial_xc'\nu(\tau)P(x,\tau)$$
$$= \frac{c'^2a}{2D}\int_0^t\mathrm{d}\tau'\int_0^{\tau'}\mathrm{d}\tau\nu(\tau)\nu(\tau')\int\mathrm{d}x'x'P(x',\tau'|x,\tau)\partial_xP(x,\tau)\,. \tag{4.5.40}$$

From Eq. (4.5.27) we get

$$\partial_xP(x,\tau) = -(x-z\mathrm{e}^{-a\tau})P(x,\tau)/\mathrm{var}(x_\tau) = -\frac{a}{D}(x-z\mathrm{e}^{-a\tau})P(x,\tau)\,. \tag{4.5.41}$$

Phys. Rev. Lett. **130**, 087101 (2023)

Hence using $z\mathrm{e}^{-a\tau} = \langle x_\tau \rangle$ and $\langle \langle x_{\tau'} \rangle (x_\tau - \langle x_\tau \rangle) \rangle = 0$ we obtain

$$\begin{aligned}\langle J_t^{\mathrm{I}} J_t^{\mathrm{II}} \rangle &= -\frac{c'^2 a^2}{2D^2} \int_0^t \mathrm{d}\tau' \int_0^{\tau'} \mathrm{d}\tau \nu(\tau)\nu(\tau') \int \mathrm{d}x' (x - z\mathrm{e}^{-a\tau}) x' P(x', \tau'; x, \tau) \\ &= -\frac{c'^2 a^2}{2D^2} \int_0^t \mathrm{d}\tau' \int_0^{\tau'} \mathrm{d}\tau \nu(\tau)\nu(\tau') \langle x_{\tau'} (x_\tau - \langle x_\tau \rangle) \rangle \\ &= -\frac{c'^2 a^2}{2D^2} \int_0^t \mathrm{d}\tau' \int_0^{\tau'} \mathrm{d}\tau \nu(\tau)\nu(\tau') \mathrm{cov}(x_\tau, x_{\tau'}) \,. \end{aligned} \tag{4.5.42}$$

Comparison with the third line in Eq. (4.5.39) yields

$$\langle J_t^{\mathrm{I}} J_t^{\mathrm{II}} \rangle = -\mathrm{var}(\rho_t^\nu) \,. \tag{4.5.43}$$

Therefore from $J_t^\nu = J_t^{\mathrm{I}} + J_t^{\mathrm{II}}$ with $\langle J_t^{\mathrm{I}} \rangle = 0$ we obtain

$$\begin{aligned}\mathrm{var}(J_t^\nu) = \langle J_t^{\mathrm{I}2} \rangle + \mathrm{var}(J_t^{\mathrm{II}}) + 2\langle J_t^{\mathrm{I}} J_t^{\mathrm{II}} \rangle &= \frac{c'^2}{2} \Sigma_t - \mathrm{var}(\rho_t^\nu) \\ \overset{\text{Eqs. (4.5.30),(4.5.39)}}{=} \frac{c'^2 a z^2}{8D} \left[1 + (2at - 1)\mathrm{e}^{-2at} \right] \,. \end{aligned} \tag{4.5.44}$$

Since $\rho_t^\nu = J_t^{\mathrm{II}}$ and $\langle J_t^{\mathrm{I}} \rangle = 0$, Eq. (4.5.43) implies

$$\mathrm{cov}(J_t^\nu, \rho_t^\nu) = \mathrm{cov}(J_t^{\mathrm{I}} + J_t^{\mathrm{II}}, J_t^{\mathrm{II}}) = \langle J_t^{\mathrm{I}} J_t^{\mathrm{II}} \rangle + \mathrm{var}(J_t^{\mathrm{II}}) = 0 \,, \tag{4.5.45}$$

which gives

$$\begin{aligned}\mathrm{var}(J_t^\nu - c\rho_t^\nu) &= \mathrm{var}(J_t^\nu) + c^2 \mathrm{var}(\rho_t^\nu) \\ &= \frac{c'^2 a z^2}{8D} \Big(1 + (2at - 1)\mathrm{e}^{-2at} + c^2 \left[1 - (1 + 2at)\mathrm{e}^{-2at} \right] \Big) \,. \end{aligned} \tag{4.5.46}$$

We now have evaluated all expressions entering the various transient TURs for the current J_t^ν and density ρ_t^ν. The quality factors (ratios of right- and left-hand side) Q_J for the transient TUR $\Sigma_t \mathrm{var}(J_t) \geq 2[t\partial_t \langle J_t \rangle - \langle \widetilde{J}_t \rangle]^2$ [Eq. (4.5.16)], Q_ρ for the transient density-TUR $\Sigma_t \mathrm{var}(\rho_t) \geq 2[(t\partial_t - 1)\langle \rho_t \rangle - \langle \widetilde{\rho}_t \rangle]^2$ [Eq. (4.5.19)], and $Q_C(c)$ (function of c) for the transient correlation-TUR $\Sigma_t \mathrm{var}(J_t - c\rho_t) \geq 2(t\partial_t \langle J_t \rangle - \langle \widetilde{J}_t \rangle - c[(t\partial_t - 1)\langle \rho_t \rangle - \langle \widetilde{\rho}_t \rangle])^2$ [Eq. (4.5.22)] after straightforward simplifications read

$$\begin{aligned} Q_J &= \frac{1 - (1 - 2at)\mathrm{e}^{-2at}}{2(1 - \mathrm{e}^{-2at})} \,, \\ Q_\rho &= \frac{1 - (1 + 2at)\mathrm{e}^{-2at}}{2(1 - \mathrm{e}^{-2at})} \,, \end{aligned}$$

Phys. Rev. Lett. 130, 087101 (2023)

$$Q_C(c) = \frac{\left[(1+c)(1-\mathrm{e}^{-2at}) + (1-c)2at\mathrm{e}^{-2at}\right]^2}{2(1-\mathrm{e}^{-2at})\left[(1+c^2)(1-\mathrm{e}^{-2at}) + (1-c^2)2at\mathrm{e}^{-2at}\right]}\,,$$
$$Q_C(1) = 1\,. \tag{4.5.47}$$

These quality factors along with Q_x in Eq. (4.5.34) are depicted in Fig. 4.5.1. They only depend on the dimensionless quantity at but not on other parameters of the process.

Opposed to Q_x in Eq. (4.5.34) these quality factors approach non-zero values as $at \to \infty$, namely $Q_J \to 1/2$, $Q_\rho \to 1/2$, $Q_C(c) \to (1+c)^2/2(1+c^2)$. Interestingly, for this special case, the current and density quality factors add up to one, $Q_J + Q_\rho = 1$. While $\lim_{at\to 0} Q_J = 1$ [as expected in general for such choice of current (see Sec. 4.5.6)], we have $\lim_{at\to 0} Q_\rho = 0$. This is because $at \to 0$ corresponds to the steady-state limit, where the density-TUR becomes the trivial $\Sigma_t \mathrm{var}(\rho_t) \geq 0$.

Recall that $\mathrm{cov}(J_t^\nu, \rho_t^\nu) = 0$ [see Eq. (4.5.45)]. We now see that the choice $c = \mathrm{cov}(J_t, \rho_t)/\mathrm{var}(\rho_t)$, that is optimal in steady-state dynamics (see Subsec. 4.5.5), here gives $c = 0$ although $c = 1$ is optimal, see Eq. (4.5.47). As mentioned in Sec. 4.5.5, this arises because $c = 0$ only optimizes the left-hand side of the corelation-TUR [also seen from Eq. (4.5.45)] but in the generalized (i.e., transient) correlation-TUR the right-hand side also depends on c.

Counterexamples

We showed that the TUR for transient dynamics [Eq. (4.5.16)] reads $\Sigma_t \mathrm{var}(J_t) \geq 2[t\partial_t\langle J_t\rangle - \langle\widetilde{J}_t\rangle]^2$ and thus contains the correction term $-\langle\widetilde{J}_t\rangle$ which contributes if $\mathbf{U}(\mathbf{x},\tau)$ in the current $J_t \equiv \int_{\tau=0}^{\tau=t} U(\mathbf{x}_\tau,\tau)\circ \mathrm{d}x_\tau$ depends on time τ. Based on the setting in the previous section (and Fig. 4.5.1) we here give an explicit counterexample for the TUR without the correction term, i.e., an example where $\Sigma_t \mathrm{var}(J_t) < 2\left[t\partial_t\langle J_t\rangle\right]^2$. This shows that the correction term $-\langle\widetilde{J}_t\rangle$ is indeed necessary, and that the result in Eq. (4.5.16) is valid for a broader class of systems than existing literature [126] by allowing explicit time dependence in $\mathbf{U}(\mathbf{x}_\tau,\tau)$. In addition, we also provide a counterexample to show that the NESS TUR $\Sigma\mathrm{var}(J_t) \geq 2\langle J_t\rangle^2$ does not hold in this transient system. Both counterexamples are shown in Fig. 4.5.2 and the derivations of the respective terms are shown below.

To find an example for which $2\left[t\partial_t\langle J_t\rangle\right]^2 > \Sigma_t \mathrm{var}(J_t)$, we note [recalling Eq. (4.5.8), $\langle J_t\rangle = \int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x}\mathbf{U}(\mathbf{x},\tau)\cdot\mathbf{j}(\mathbf{x},\tau)$] that the term $t\partial_t\langle J_t\rangle = t\int \mathrm{d}\mathbf{x}\mathbf{U}(\mathbf{x},t)\cdot\mathbf{j}(\mathbf{x},t)$ only involves $\mathbf{U}(\mathbf{x},t)$ at the final time but not at any $\tau < t$. In contrast, Σ_t is independent of the choice of $\mathbf{U}$ and $\mathrm{var}(J_t)$ involves $\mathbf{U}(\mathbf{x},\tau)$ at all times. Therefore, examples for $2\left[t\partial_t\langle J_t\rangle\right]^2 > \Sigma_t \mathrm{var}(J_t)$ can be found by making $\mathbf{U}(\mathbf{x},t)$ large compared to $\mathbf{U}(\mathbf{x},\tau)$ at $\tau < t$.

Phys. Rev. Lett. **130**, 087101 (2023)

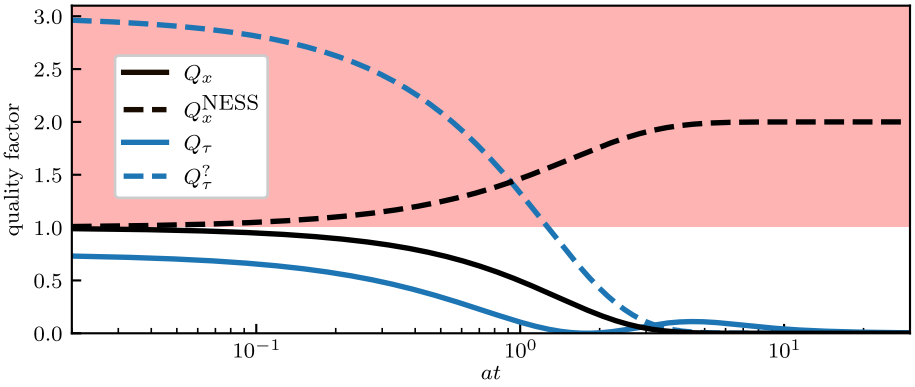

Figure 4.5.2: Quality factors for the displaced harmonic trap for currents $J_t^x \equiv \int_{\tau=0}^{\tau=t} 1 \circ \mathrm{d}x_\tau = x_t - x_0$ (black) and $J_t^\tau \equiv \int_{\tau=0}^{\tau=t} \tau \circ \mathrm{d}x_\tau$ (blue). Quality factors $Q_{x,\tau} \equiv 2[t\partial_t\langle J_t^{x,\tau}\rangle - \langle \widetilde{J}_t^{x,\tau}\rangle]^2/\Sigma_t \mathrm{var}(J_t^{x,\tau}) \leq 1$, see Eqs. (4.5.34) and (4.5.52), for the transient TUR [Eq. (4.5.16)] are shown by solid lines. The dashed black line $Q_x^{\mathrm{NESS}} \equiv 2\langle J_t^x\rangle^2/\Sigma_t \mathrm{var}(J_t^x)$, see Eq. (4.5.53), and dashed blue line $Q_\tau^? \equiv 2[t\partial_t\langle J_t^\tau\rangle]^2/\Sigma_t \mathrm{var}(J_t^\tau)$, see Eq. (4.5.51), intersect the forbidden region $Q > 1$ (red background) which shows that the NESS TUR and the transient TUR without the correction $-\langle \widetilde{J}_t\rangle$ term are not generally valid for transient systems.

We now give an explicit example by choosing a current J_t^τ with linear time dependence in $U(x,\tau) = \tau$ (here one-dimensional),

$$J_t^\tau \equiv \int_{\tau=0}^{\tau=t} \tau \circ \mathrm{d}x_\tau \,. \tag{4.5.48}$$

Note that due to $\partial_x U(x,\tau) = 0$ there is no difference between Stratonovich and Itô integration. We calculate

$$\begin{aligned} \langle J_t^\tau\rangle &= \int_0^t \tau\langle -a x_\tau\rangle \mathrm{d}\tau = -az\int_0^t \tau \mathrm{e}^{-a\tau}\mathrm{d}\tau = -\frac{z}{a}\left[1 - \mathrm{e}^{-at}(1+at)\right] , \\ t\partial_t\langle J_t^\tau\rangle &= \frac{zt}{a}\left[-a(1+at)+a\right]\mathrm{e}^{-at} = -zat^2\mathrm{e}^{-at} \,. \end{aligned} \tag{4.5.49}$$

For the variance compute analogously to the lines leading to Eq. (4.5.44) [with cov from Eq. (4.5.29)]

$$\begin{aligned} \mathrm{var}(J_t^\tau) &= 2D\int_0^t \mathrm{d}\tau\, \tau^2 - a^2\int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau'\, \tau\tau' \mathrm{cov}(x_\tau, x_{\tau'}) \\ &= 2\frac{D}{a^3}\left(\frac{a^3t^3}{3} - \frac{1}{6}\left[a^2t^2(2at-3) - 6\mathrm{e}^{-at}(1+at) + 6\right]\right) \\ &= \frac{D}{a^3}\left[a^2t^2 + 2\mathrm{e}^{-at}(1+at) - 2\right] . \end{aligned} \tag{4.5.50}$$

Phys. Rev. Lett. 130, 087101 (2023)

Since the dissipation Σ_t does not depend on the choice of J_t, U, it is still given by Eq. (4.5.30). The quality factor $Q_\tau^?$ for the TUR $\Sigma_t \mathrm{var}(J_t) \geq 2\left[t\partial_t\langle J_t\rangle\right]^2$ that would hold in the absence of explicit time dependence in U [Eq. (4.5.16) for $\langle\widetilde{J}_t\rangle = 0$; see also Ref. [126]] reads

$$Q_\tau^? = \frac{4a^4t^4\mathrm{e}^{-2at}}{(1-\mathrm{e}^{-2at})\left[a^2t^2+2\mathrm{e}^{-at}(1+at)-2\right]}\,. \tag{4.5.51}$$

Thus we see that as the other quality factors in Eqs. (4.5.34) and (4.5.47), $Q_\tau^?$ only depends on the quantity at. In Fig. 4.5.2 we see that $Q_\tau^? > 1$ for small values of at which breaks the TUR $\Sigma_t \mathrm{var}(J_t) \geq 2\left[t\partial_t\langle J_t\rangle\right]^2$, i.e., this example shows that the correction term $-\langle\widetilde{J}_t\rangle$ in the TUR $\Sigma_t \mathrm{var}(J_t) \geq 2[t\partial_t\langle J_t\rangle - \langle\widetilde{J}_t\rangle]^2$ [Eq. (4.5.16)] is necessary for general validity for currents with explicitly time-dependent $\mathbf{U}$.

For comparison, we also give the correct quality factor Q_τ for the transient TUR [Eq. (4.5.16)] including the correction term $-\langle\widetilde{J}_t\rangle$. Since $\tau\partial_\tau\tau = \tau$ we have $\widetilde{J}_t = J_t$ such that the correct quality factor using Eq. (4.5.49) reads

$$\begin{aligned} Q_\tau &= \frac{2\left(-zat^2\mathrm{e}^{-at} + \frac{z}{a}\left[1-\mathrm{e}^{-at}(1+at)\right]\right)^2}{\frac{az^2}{2D}(1-\mathrm{e}^{-2at})\frac{D}{a^3}\left[a^2t^2+2\mathrm{e}^{-at}(1+at)-2\right]} \\ &= \frac{4\left[1-\mathrm{e}^{-at}(1+at+a^2t^2)\right]^2}{(1-\mathrm{e}^{-2at})\left[a^2t^2+2\mathrm{e}^{-at}(1+at)-2\right]}\,. \end{aligned} \tag{4.5.52}$$

To give a counterexample that shows that NESS TUR $\Sigma_t \mathrm{var}(J_t) \geq 2\langle J_t\rangle^2$ breaks down for this transient system, consider the current $J_t^x = x_t - x_0$ as in Eq. (4.5.32). The quality factor for this example follows from Eq. (4.5.33),

$$Q_x^{\mathrm{NESS}} = \frac{2z^2(1-\mathrm{e}^{-at})^2}{\frac{az^2}{2D}(1-\mathrm{e}^{-2at})\frac{2D}{a}(1-\mathrm{e}^{-at})} = \frac{2(1-\mathrm{e}^{-at})}{1-\mathrm{e}^{-2at}}\,. \tag{4.5.53}$$

Note that for any $y \in [0,1)$ the inequality $(1-y)^2 > 0$ implies $2(1-y) > 1-y^2$. Thus, we have for any value $at > 0$ that $Q_x^{\mathrm{NESS}} > 1$, see also Fig. 4.5.2. This provides (for any $at > 0$) a counterexample against the NESS TUR, i.e., as expected the NESS TUR does not hold for transient dynamics.

Detailed derivation of $\langle A_t^2\rangle = \Sigma_t/2$

Recall the "educated guess" [Eq. (4.5.9)]

$$A_t \equiv \int_{\tau=0}^{\tau=t} \mathbf{a}(\mathbf{x}_\tau,\tau)^T \mathrm{d}\mathbf{W}_\tau\,,$$

Phys. Rev. Lett. **130**, 087101 (2023)

$$\mathbf{a}(\mathbf{x}_\tau,\tau)^T \equiv \frac{\mathbf{j}(\mathbf{x}_\tau,\tau)^T}{P(\mathbf{x}_\tau,\tau)}[2\mathbf{D}(\mathbf{x}_\tau)]^{-1}\boldsymbol{\sigma}(\mathbf{x}_\tau)\,,$$
$$\mathbf{a}(\mathbf{x}_\tau,\tau) = \boldsymbol{\sigma}(\mathbf{x}_\tau)^T[2\mathbf{D}(\mathbf{x}_\tau)]^{-1}\frac{\mathbf{j}(\mathbf{x}_\tau,\tau)}{P(\mathbf{x}_\tau,\tau)}\,, \tag{4.5.54}$$

where A_t and $P(\mathbf{x}_\tau,\tau)$ are scalars, $\mathbf{a}(\mathbf{x},\tau)$, $\mathbf{j}(\mathbf{x}_\tau,\tau)$ and $\mathrm{d}\mathbf{W}_\tau$ are vectors, and $2\mathbf{D}(\mathbf{x}_\tau) = 2\mathbf{D}(\mathbf{x}_\tau)^T = \boldsymbol{\sigma}(\mathbf{x}_\tau)\boldsymbol{\sigma}(\mathbf{x}_\tau)^T$ are matrices (note that $([2\mathbf{D}(\mathbf{x}_\tau)]^{-1})^T = [2\mathbf{D}(\mathbf{x}_\tau)]^{-1}$). Due to the "delta-correlated" noise property $\langle \mathrm{d}W_{\tau,i}\mathrm{d}W_{\tau',j}\rangle = \delta(\tau-\tau')\delta_{ij}\mathrm{d}\tau\mathrm{d}\tau'$ the expectation of the square of the noise-integral A_t is given by [6, 16]

$$\begin{aligned}\langle A_t^2\rangle &= \int_0^t \langle \mathbf{a}(\mathbf{x}_\tau,\tau)^T\mathbf{a}(\mathbf{x}_\tau,\tau)\rangle\,\mathrm{d}\tau \\ &= \int_0^t \left\langle \frac{\mathbf{j}(\mathbf{x}_\tau,\tau)^T}{P(\mathbf{x}_\tau,\tau)}[2\mathbf{D}(\mathbf{x}_\tau)]^{-1}\frac{\mathbf{j}(\mathbf{x}_\tau,\tau)}{P(\mathbf{x}_\tau,\tau)}\right\rangle \mathrm{d}\tau.\end{aligned} \tag{4.5.55}$$

Using that $\langle f(\mathbf{x}_\tau,\tau)\rangle = \int \mathrm{d}\mathbf{x}P(\mathbf{x},\tau)f(\mathbf{x},\tau)$ and comparing to the definition of Σ_t [Eq. (4.5.5)],

$$\Sigma_t = \int \mathrm{d}\mathbf{x}\int_0^t \frac{\mathbf{j}(\mathbf{x},\tau)^T\mathbf{D}^{-1}(\mathbf{x})\mathbf{j}(\mathbf{x},\tau)}{P(\mathbf{x},\tau)}d\tau\,, \tag{4.5.56}$$

we immediately obtain $\langle A_t^2\rangle = \Sigma_t/2$.

Detailed derivation of Eq. (4.5.13)

From the inequality $\Sigma_t \mathrm{var}(J_t) \geq 2\left[\langle J_t\rangle + \langle A_t J_t^{\mathrm{II}}\rangle\right]^2$ [Eq. (4.5.11)] we derive TURs by evaluating the expectation value (define notation $\mathbf{g}(\mathbf{x}_\tau,\tau) \equiv \frac{\mathbf{j}(\mathbf{x}_\tau,\tau)}{P(\mathbf{x}_\tau,\tau)}[2\mathbf{D}(\mathbf{x}_\tau)]^{-1}$)

$$\begin{aligned}\langle A_t J_t^{\mathrm{II}}\rangle &\equiv \left\langle \int_{\tau=0}^{\tau=t}\mathbf{g}(\mathbf{x}_\tau,\tau)\cdot\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau \int_0^t \mathcal{U}(\mathbf{x}_{\tau'},\tau')\mathrm{d}\tau'\right\rangle \\ &= \int_0^t \mathrm{d}\tau' \int_{\tau=0}^{\tau=t}\left\langle \mathbf{g}(\mathbf{x}_\tau,\tau)\cdot\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau\mathcal{U}(\mathbf{x}_{\tau'},\tau')\right\rangle\,.\end{aligned} \tag{4.5.57}$$

We use an approach from Refs. [84, 85] to evaluate this expectation value. For convenience of the reader, we recall this approach here and apply it to this special case.

First, note that for times $\tau \geq \tau'$ this expectation value vanishes due to the independence property of the Wiener process. However, non-trivial contributions occur for $\tau < \tau'$ because the probability density of $\mathbf{x}_{\tau'}$ depends on $\mathrm{d}\mathbf{W}_\tau$. We want to express $\langle A_t J_t^{\mathrm{II}}\rangle$ in terms of integrals over the probability density $P(x,\tau)$ [that contains the information on the initial condition $P(x,0)$] and the conditional density $P(\mathbf{x}',\tau'|\mathbf{x},\tau)$. This would be trivial in the absence of the noise increment $\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau$ by using $\langle \int_0^t V_1(\mathbf{x}_\tau,\tau')\mathrm{d}\tau \int_\tau^t V_2(\mathbf{x}_{\tau'},\tau')\mathrm{d}\tau'\rangle =$

Phys. Rev. Lett. 130, 087101 (2023)

$\int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x}' \int_0^t \mathrm{d}\tau \int_\tau^t \mathrm{d}\tau' V_1(\mathbf{x},\tau) V_2(\mathbf{x}',\tau') P(\mathbf{x}',\tau'|\mathbf{x},\tau) P(\mathbf{x},\tau)$. The critical task is generalizing this to integration involving the noise increment.

For a given point $\mathbf{x}_\tau = \mathbf{x}$ we set $\boldsymbol{\varepsilon} \equiv \boldsymbol{\sigma}(\mathbf{x})\mathrm{d}\mathbf{W}_\tau = \mathcal{O}(\sqrt{\mathrm{d}\tau})$ which has a Gaussian probability distribution $P(\boldsymbol{\varepsilon})$ with zero mean and covariance matrix $2\mathbf{D}(\mathbf{x})\mathrm{d}\tau$. For a given $\boldsymbol{\varepsilon}$, the equation of motion in Itô form implies a position increment $d\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon}) = [\mathbf{F}(\mathbf{x}) + \nabla\cdot\mathbf{D}(\mathbf{x})]\mathrm{d}\tau + \boldsymbol{\varepsilon}$. We now write the average in Eq. (4.5.57) as integrals over the probability density to be at points $\mathbf{x}, \mathbf{x} + d\mathbf{x}_\tau, \mathbf{x}'$ at times $\tau < \tau + d\tau < \tau'$, respectively, i.e., for $\tau < \tau'$

$$\begin{aligned}&\left\langle \mathbf{g}(\mathbf{x}_\tau,\tau)\cdot\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau \mathcal{U}(\mathbf{x}_{\tau'},\tau')\right\rangle \\ &= \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x}' \mathbf{g}(\mathbf{x},\tau)\cdot\boldsymbol{\varepsilon}\, \mathcal{U}(\mathbf{x}',\tau') P(\boldsymbol{\varepsilon}) P(\mathbf{x}',\tau'|\mathbf{x}+d\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon}),\tau+d\tau) P(\mathbf{x},\tau)\,. \end{aligned} \tag{4.5.58}$$

Expanding in small $d\tau$ gives

$$\begin{aligned}&P(\mathbf{x}',\tau'|\mathbf{x}+d\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon}),\tau+d\tau) \\ &= P(\mathbf{x}',\tau'|\mathbf{x},\tau) + \mathrm{d}\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon})\cdot\nabla_{\mathbf{x}} P(\mathbf{x}',\tau'|\mathbf{x},\tau) + \mathcal{O}(d\tau)\,. \end{aligned} \tag{4.5.59}$$

By symmetry only the term of even power $\sim \boldsymbol{\varepsilon}^2$ in $\boldsymbol{\varepsilon}[\mathrm{d}\mathbf{x}_\tau(\mathbf{x},\tau,\boldsymbol{\varepsilon})\cdot\nabla_{\mathbf{x}} P(\mathbf{x}',\tau'|\mathbf{x},\tau)]$ survives the integration over $P(\boldsymbol{\varepsilon})$ and contributes according to the covariance matrix $2\mathbf{D}(\mathbf{x})\mathrm{d}\tau$. Therefore, we arrive at (where $\mathbb{1}_{\tau<\tau'} = 1$ if $\tau < \tau'$ and 0 otherwise)

$$\begin{aligned}&\langle \mathbf{g}(\mathbf{x}_\tau,\tau)\cdot\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau \mathcal{U}(\mathbf{x}_{\tau'},\tau')\rangle \\ &= \mathbb{1}_{\tau<\tau'}\mathrm{d}\tau \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x}' \mathcal{U}(\mathbf{x}',\tau') \mathbf{g}(\mathbf{x},\tau) 2\mathbf{D}(\mathbf{x}) P(\mathbf{x},\tau)\cdot\nabla_{\mathbf{x}} P(\mathbf{x}',\tau'|\mathbf{x},\tau)\,. \end{aligned} \tag{4.5.60}$$

Now we perform an integration by parts in $\mathbf{x}$. The boundary terms vanish in infinite space due to the vanishing of the probability density at $||\mathbf{x}|| \to \infty$, in finite space with reflecting (i.e., zero-flux) boundary conditions they vanish by the divergence theorem, and in a finite system with periodic boundary conditions the boundary terms cancel. Note that finite spatial domains with reflecting boundary conditions are in essence already contained in the infinite-space-case as the limit of a strongly confining potential. Using the symmetry $\mathbf{D}^T(\mathbf{x}) = \mathbf{D}(\mathbf{x})$ we arrive at

$$\begin{aligned}&\langle \mathbf{g}(\mathbf{x}_\tau,\tau)\cdot\boldsymbol{\sigma}(\mathbf{x}_\tau)\mathrm{d}\mathbf{W}_\tau \mathcal{U}(\mathbf{x}_{\tau'},\tau')\rangle \\ &= -\mathbb{1}_{\tau<\tau'}\mathrm{d}\tau \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x}' \mathcal{U}(\mathbf{x}',\tau') P(\mathbf{x}',\tau'|\mathbf{x},\tau) \nabla_{\mathbf{x}}\cdot[2\mathbf{D}(\mathbf{x})\mathbf{g}(\mathbf{x},\tau)P(\mathbf{x},\tau)]\,. \end{aligned} \tag{4.5.61}$$

Phys. Rev. Lett. **130**, 087101 (2023)

Plugging in the explicit form of $\mathbf{g}(\mathbf{x},\tau)$ we finally obtain

$$\begin{aligned}\langle A_t J_t^{\mathrm{II}}\rangle &= -\int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x}' \mathcal{U}(\mathbf{x}',\tau') \int_0^t \mathrm{d}\tau \mathbb{1}_{\tau<\tau'} \int \mathrm{d}\mathbf{x} P(\mathbf{x}',\tau'|\mathbf{x},\tau) \nabla_{\mathbf{x}} \cdot \mathbf{j}(\mathbf{x},\tau) \\ &= -\int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x}' \mathcal{U}(\mathbf{x}',\tau') \int_0^{\tau'} \mathrm{d}\tau \int \mathrm{d}\mathbf{x} P(\mathbf{x}',\tau'|\mathbf{x},\tau) \nabla_{\mathbf{x}} \cdot \mathbf{j}(\mathbf{x},\tau)\,. \end{aligned} \tag{4.5.62}$$

Derivation of Eq. (4.5.15)

We here simplify Eq. (4.5.62) for the case of transient dynamics, i.e., where $\nabla \cdot \mathbf{j}(\mathbf{x},\tau) = -\partial_\tau P(\mathbf{x},\tau)$ does not vanish. An integration by parts in τ with the boundary term $-\int \mathrm{d}\mathbf{x} P(\mathbf{x}',\tau'|\mathbf{x},0)P(\mathbf{x},0) = -P(\mathbf{x}',\tau')$ yields

$$\begin{aligned}&\langle A_t J_t^{\mathrm{II}}\rangle \\ &= \int_0^t \mathrm{d}\tau' \int \mathrm{d}\mathbf{x}' \mathcal{U}(\mathbf{x}',\tau') \left(-P(\mathbf{x}',\tau') - \int \mathrm{d}\mathbf{x} \int_0^t \mathrm{d}\tau P(\mathbf{x},\tau) \partial_\tau \left[\mathbb{1}_{\tau<\tau'} P(\mathbf{x}',\tau'|\mathbf{x},\tau) \right] \right). \end{aligned} \tag{4.5.63}$$

Note that the first term is $-\langle J_t^{\mathrm{II}}\rangle$. Since we consider Markovian systems without explicit time dependence of $\mathbf{F}$ and $\boldsymbol{\sigma}$, we have $\partial_\tau P(\mathbf{x}',\tau'|\mathbf{x},\tau) = \partial_\tau P(\mathbf{x}',\tau'-\tau|\mathbf{x}) = -\partial_{\tau'} P(\mathbf{x}',\tau'-\tau|\mathbf{x}) = -\partial_{\tau'} P(\mathbf{x}',\tau'|\mathbf{x},\tau)$. Using moreover $\int \mathrm{d}\mathbf{x} P(\mathbf{x}',\tau'|\mathbf{x},\tau) P(\mathbf{x},\tau) = P(\mathbf{x}',\tau')$ and $\int_0^t \mathrm{d}\tau \mathbb{1}_{\tau<\tau'} = \tau'$ we obtain, upon integrating by parts with the boundary term entering at $\tau' = t$, and recalling $\langle J_t^{\mathrm{II}}\rangle = \langle J_t\rangle$,

$$\begin{aligned}\langle A_t J_t^{\mathrm{II}}\rangle &= -\langle J_t^{\mathrm{II}}\rangle + \int \mathrm{d}\mathbf{x}' \int_0^t \mathrm{d}\tau' \mathcal{U}(\mathbf{x}',\tau') \partial_{\tau'} \left[\tau' P(\mathbf{x}',\tau') \right] \\ &= (t\partial_t - 1)\langle J_t\rangle - \int \mathrm{d}\mathbf{x}' \int_0^t \mathrm{d}\tau' P(\mathbf{x}',\tau') \tau' \partial_{\tau'} \mathcal{U}(\mathbf{x}',\tau')\,. \end{aligned} \tag{4.5.64}$$

In order to make Eq. (4.5.64) operationally accessible we define a second current

$$\widetilde{J}_t \equiv \int_{\tau=0}^{\tau=t} \tau \partial_\tau \mathbf{U}(\mathbf{x}_\tau,\tau) \cdot \circ \mathrm{d}\mathbf{x}_\tau\,, \tag{4.5.65}$$

where $\langle \widetilde{J}_t\rangle$ is analogously to Eqs. (4.5.7) and (4.5.8) obtained via $\tau\partial_\tau \mathcal{U}$ such that from Eq. (4.5.64) we obtain Eq. (4.5.15), i.e.,

$$\langle A_t J_t^{\mathrm{II}}\rangle = (t\partial_t - 1)\langle J_t\rangle - \langle \widetilde{J}_t\rangle\,. \tag{4.5.66}$$

Phys. Rev. Lett. 130, 087101 (2023)

Extension to Markov jump processes

We consider steady-state Markov jump dynamics. Recall from Subsec. 4.5.8 that currents are defined with antisymmetric prefactors $d_{xy} = -d_{yx}$ as the double sum $J \equiv \sum_{x\neq y} d_{xy}\hat{n}_{xy}$, and that the dissipation reads $\Sigma \equiv t\sum_{x\neq y} p_x r_{xy} \ln[p_x r_{xy}/p_y r_{yx}]$. Moreover, recall the definition [Eq. (4.5.23)]

$$\begin{aligned} A &\equiv \sum_{x\neq y} Z_{xy}(\hat{n}_{xy} - \hat{\tau}_x r_{xy}) \\ Z_{xy} &\equiv \frac{p_x r_{xy} - p_y r_{yx}}{p_x r_{xy} + p_y r_{yx}} \,. \end{aligned} \tag{4.5.67}$$

To complete the proof of the steady-state TUR as outlined in Sec. 4.5.8 we need to show $\langle A\rangle = 0$, $\langle A^2\rangle \leq \Sigma/2$ and $\langle AJ\rangle = \langle J\rangle$. Proving these statements can be performed in complete analogy to the Cramér-Rao proof of the steady-state TUR (the latter can e.g., be found in Ref. [233]) by tilting the rates $r_{xy}(\theta) = r_{xy}\mathrm{e}^{\theta Z_{xy}}$ and identifying $A = \partial_\theta\big|_{\theta=0} \ln \mathcal{P}_\theta$ where $\mathcal{P}_\theta$ is the tilted path measure. Note that $\langle \partial_\theta \ln \mathcal{P}_\theta\rangle = \int \partial_\theta \mathcal{P}_\theta = \partial_\theta 1 = 0$ implies $\langle A\rangle = 0$. Using explicit properties of Z_{xy} we have $\partial_\theta \langle J\rangle_\theta = \langle J\rangle_\theta$ [233] which implies $\langle J\partial_\theta \ln \mathcal{P}_\theta\rangle_\theta = \partial_\theta \int J\mathcal{P}_\theta = \partial_\theta\langle J\rangle_\theta = \langle J\rangle_\theta$. Setting $\theta = 0$ implies $\langle AJ\rangle = \langle J\rangle$. Evaluating the Fisher information $\mathcal{I}(\theta)$ at $\theta = 0$ yields $\mathcal{I}(\theta = 0) = \sum_{x\neq y}(p_x r_{xy} - p_y r_{yx})^2/2(p_x r_{xy} + p_y r_{yx}) \leq \Sigma/2$ [233] which due to $\langle A^2\rangle = \langle (\partial_\theta\big|_{\theta=0} \ln \mathcal{P}_\theta)^2\rangle = \mathcal{I}(\theta = 0)$ concludes the proof.

Although we have completed the proof of the NESS TUR, this is by no means a "direct" proof, and it does not directly generalize in analogy to the generalizations performed above in continuous space. In order to give a genuinely "direct" proof in the sense that it completely follows from the equation of motion (i.e., from the properties of Markovian jumps and from the master equation), to generalize to arbitrary initial conditions, and to incorporate correlations of currents and densities, we consider a direct analogy to the continuous space approach presented above. We now give an outlook on this direct approach. Define $\hat{c}_{xy}(\tau)$ as the random variable representing a jump $x \to y$ at time τ such that $\hat{n}_{xy} = \int_0^t \mathrm{d}\tau \hat{c}_{xy}(\tau)$ and $\hat{\mathbb{1}}_x(\tau)$ as the random variable yielding 1 if the state x is occupied at τ (and 0 otherwise) such that $\hat{\tau}_x = \int_0^t \mathrm{d}\tau \hat{\mathbb{1}}_x(\tau)$. Then define

$$A \equiv \sum_{x\neq y} \int_0^t \mathrm{d}\tau \frac{p_x(\tau) r_{xy} - p_y(\tau) r_{yx}}{p_x(\tau) r_{xy} + p_y(\tau) r_{yx}} \left[\hat{c}_{xy}(\tau) - r_{xy}\hat{\mathbb{1}}_x(\tau)\right] \,, \tag{4.5.68}$$

and split a current $J = \sum_{i\neq j} d_{ij}\hat{n}_{ij} = \sum_{i\neq j} d_{ij}\int_0^t \mathrm{d}\tau \hat{c}_{ij}(\tau)$ (where $d_{ij} = -d_{ji}$) into

$$J = J^{\mathrm{I}} + J^{\mathrm{II}} \,,$$

Phys. Rev. Lett. **130**, 087101 (2023)

$$J^{\mathrm{I}} = \sum_{i\neq j} d_{ij}(\hat{n}_{ij} - r_{ij}\hat{\tau}_i) = \sum_{i\neq j} d_{ij}\int_0^t \mathrm{d}\tau\left[\hat{c}_{ij}(\tau) - r_{ij}\hat{\mathbb{1}}_i(\tau)\right],$$
$$J^{\mathrm{II}} = \sum_{i\neq j} d_{ij}r_{ij}\hat{\tau}_i = \sum_{i\neq j} d_{ij}r_{ij}\int_0^t \mathrm{d}\tau\hat{\mathbb{1}}_i(\tau)\,. \tag{4.5.69}$$

Using this notation, one can directly show that $\langle A\rangle = 0$, $\langle A^2\rangle \leq \Sigma/2$, $\langle AJ^{\mathrm{I}}\rangle = \langle J\rangle$ (analogous to the proof above in continuous space but instead of $\langle\sigma(x)\mathrm{d}W_\tau\rangle = 0$ and $\langle[\sigma(x)\mathrm{d}W_\tau]^2\rangle = 2D(x)\mathrm{d}\tau$ use that given $\hat{\mathbb{1}}_x(\tau) = 1$ the term $[\hat{c}_{xy}(\tau) - r_{xy}]\mathrm{d}\tau$ has zero mean and variance $r_{xy}\mathrm{d}\tau$). Thus, as in Eq. (4.5.11) we have $\Sigma\mathrm{var}(J) \geq 2\left[\langle J\rangle + \langle AJ^{\mathrm{II}}\rangle\right]^2$. Deriving the different TURs then follows in analogy to Subsecs. 4.5.3-4.5.5 by evaluating $\langle AJ^{\mathrm{II}}\rangle$ and accordingly introducing densities. All results, including the discussion of the saturation, would then equally apply to Markov jump processes, which will be addressed in the future.

Phys. Rev. Lett. **130**, 087101 (2023)

Chapter 5

Thermal relaxation asymmetry

The largest part of the thesis, namely the topic of time-integrated densities and currents, is now behind us. In this chapter, we address the topic of thermal relaxation, in particular, the comparison of heating and cooling processes. After introducing and specifying the question, we will reproduce two publications in Secs. 5.4 and 5.5.

5.1 Introduction to thermal relaxation

If we place a system at a certain temperature in an environment with a different temperature, the system will approach the environment's temperature in a process called *thermal relaxation*. The conventional tools to analyze this process are Fourier's law, stating that the transferred heat is proportional to the temperature difference, or the framework of *linear irreversible thermodynamics* [25–27] based on the work of Lars Onsager which was awarded the Nobel Prize in Chemistry in 1968 [247]. Both approaches, however, are inherently linear and do not hold far from thermal equilibrium. Therefore, they are unable to address thermal relaxation beyond the linear regime of small temperature differences.

To go beyond the linear regime, we recall that, in contrast to the conventional approaches, the Langevin equation and the framework of stochastic thermodynamics are valid arbitrarily far from equilibrium. This opens the door to many questions that could not have been studied before. In particular, we want to compare heating and cooling processes to investigate whether (unlike in the linear regime) the sign of the temperature differences matters. The setup to this problem, based on the methods developed in Ch. 2, was first studied in [73]. Consider a Langevin dynamics that tends to equilibrium for long times, and study the relaxation for different initial conditions, reflecting different initial temper-

atures. Specifically, the initial condition will be chosen from the Boltzmann distribution $p(x, t = 0) \propto \exp[-\phi(x)/k_{\mathrm{B}}T_i]$ of the process at an initial temperature T_i.

One important aspect, however, is not a priori clear: How do we quantify the process in time? The problem is that far from equilibrium, there is no reason to assume that the process passes through local equilibrium states at intermediate temperatures, i.e., we will *not* have that $p(x, t) \propto \exp[-\phi(x)/k_{\mathrm{B}}T(t)]$ with temperature $T(t)$. Such an intermediate temperature simply does not exist, which raises the question of how to quantify the time-evolution of the process instead. Of course, the probability density $p(x, t)$ contains the full information on the time evolution, but it does not directly relate to a thermodynamical distance to the target temperature. One could come up with several notions of such distances, but the thermodynamically most natural quantification appears to be in terms of an *excess free energy*, i.e., the free energy of the current state $p(x, t)$ compared to the target equilibrium $p_{\mathrm{s}}(x)$ [73].

5.2 Quantifying thermal relaxation

This section addresses the quantification of relaxation processes far from equilibrium, in particular, thermal relaxation processes. We introduce the excess free energy in Subsec. 5.2.1, and relate it to the concept of non-adiabatic entropy production in Subsec. 5.2.2. Finally, we mention an alternative quantification in Subsec. 5.2.3.

5.2.1 Excess free energy

Consider Langevin dynamics at temperature T in a potential $\phi(\mathbf{x})$ that for long times approaches a Boltzmann-Gibbs equilibrium density $p_{\mathrm{s}}(\mathbf{x}) = \exp[-\phi(\mathbf{x})/k_{\mathrm{B}}T]/Z$, see Subsec. 2.7.1. The equilibrium free energy of such a system reads $F = -k_{\mathrm{B}}T \ln Z$ [28]. Equivalently, we can use the internal energy $U = \langle \phi(\mathbf{x}_t) \rangle_{\mathrm{s}}$ and the Gibbs-Shannon entropy of the system (see Sec. 2.9) to compute the (Helmholtz) free energy of the system from $F = U - TS_{\mathrm{sys}}$, which, by rewriting the Boltzmann-Gibbs form of $p_{\mathrm{s}}(\mathbf{x})$ as $\phi(\mathbf{x}) = -k_{\mathrm{B}}T(\ln[p_{\mathrm{s}}(\mathbf{x})] + \ln Z)$, consistently gives in the equilibrium state for any t

$$\begin{aligned} F &= \langle \phi(\mathbf{x}_t) \rangle_{\mathrm{s}} - TS_{\mathrm{sys}} \\ &= \int \mathrm{d}\mathbf{x}\, p_{\mathrm{s}}(\mathbf{x})\phi(\mathbf{x}) - T\left(-k_{\mathrm{B}} \int \mathrm{d}\mathbf{x}\, p_{\mathrm{s}}(\mathbf{x}) \ln[p_{\mathrm{s}}(\mathbf{x})]\right) \\ &= -k_{\mathrm{B}}T \int \mathrm{d}\mathbf{x}\, p_{\mathrm{s}}(\mathbf{x})(\ln[p_{\mathrm{s}}(\mathbf{x})] + \ln Z) + k_{\mathrm{B}}T \int \mathrm{d}\mathbf{x}\, p_{\mathrm{s}}(\mathbf{x}) \ln[p_{\mathrm{s}}(\mathbf{x})] \\ &= -k_{\mathrm{B}}T \ln Z\,. \end{aligned} \tag{5.2.1}$$

To quantify a free energy for the same dynamics but with different initial condition $p(\mathbf{x}, 0)$, we follow the ansatz from Eq. (5.2.1) but with an arbitrary $p(\mathbf{x}, t)$,

$$\begin{aligned} F(t) &= \langle \phi(\mathbf{x}_t) \rangle - T S_{\mathrm{sys}}(t) \\ &= \int \mathrm{d}\mathbf{x}\, p(\mathbf{x}, t) \phi(\mathbf{x}) + k_{\mathrm{B}} T \int \mathrm{d}\mathbf{x}\, p(\mathbf{x}, t) \ln[p(\mathbf{x}, t)] \\ &= -k_{\mathrm{B}} T \int \mathrm{d}\mathbf{x}\, p(\mathbf{x}, t) (\ln[p_{\mathrm{s}}(\mathbf{x})] + \ln Z) + k_{\mathrm{B}} T \int \mathrm{d}\mathbf{x}\, p(\mathbf{x}, t) \ln[p(\mathbf{x}, t)] \\ &= F + k_{\mathrm{B}} T \int \mathrm{d}\mathbf{x}\, p(\mathbf{x}, t) \ln \left[\frac{p(\mathbf{x}, t)}{p_{\mathrm{s}}(\mathbf{x})} \right] . \end{aligned} \tag{5.2.2}$$

Based on this we define the *excess free energy* as

$$\mathcal{D}(t) \equiv F(t) - F = k_{\mathrm{B}} T \int \mathrm{d}\mathbf{x}\, p(\mathbf{x}, t) \ln \left[\frac{p(\mathbf{x}, t)}{p_{\mathrm{s}}(\mathbf{x})} \right] . \tag{5.2.3}$$

The integral on the right-hand side is known as the *relative entropy* or *Kullback-Leibler divergence* of the probability densities $p(\mathbf{x}, t)$ and $p_{\mathrm{s}}(\mathbf{x})$, and is an established tool in probability and information theory to quantify differences between probability densities [248]. The Kullback-Leibler divergence is not a metric (i.e., does not give a well-defined notion of geometric distance) since it is not symmetric and does not satisfy the triangle inequality. However, as desired, it is zero if $p(\mathbf{x}, t) = p_{\mathrm{s}}(\mathbf{x})$ and positive otherwise. This reflects that $p_{\mathrm{s}}(\mathbf{x})$ minimizes the free energy for Langevin dynamics at temperature T. Indeed, the Fokker-Planck dynamics of $p(\mathbf{x}, t)$ can also be understood as a particular minimization process of the free energy[1], which gives some geometrical interpretation to the relaxation dynamics in the space of probability distributions.

Note that the reasoning in this section remains useful even for relaxation to non-equilibrium steady states (NESS), where $p_{\mathrm{s}}(\mathbf{x})$ is not a genuine Boltzmann-Gibbs density but we nevertheless define an effective potential $\phi(\mathbf{x}) \equiv -k_{\mathrm{B}} T \ln p_{\mathrm{s}}(\mathbf{x})$, see Subsec. 2.7.2. We will refer to Eq. (5.2.3) as a *generalized* excess free energy in the case of relaxation towards a NESS.

Following the arguing above, we will use Eq. (5.2.3) to quantify a thermodynamic distance of $p(\mathbf{x}, t)$ to the steady-state density $p_{\mathrm{s}}(\mathbf{x})$, for both, equilibrium and NESS. Besides the interpretation as a free energy difference as derived above, there is another strong motivation to use exactly this quantification, namely its connection to the entropy production. This is particularly useful to strengthen the thermodynamic justification of Eq. (5.2.3) for NESS, where the interpretation as free energy is not as intuitive as in the equilibrium case.

[1] More precisely, the Fokker-Planck dynamics can be shown to emerge from a *gradient flow in the Wasserstein metric* minimizing the free energy functional $F[p(\cdot)]$ [249].

5.2.2 Connection of excess free energy and entropy production

Relaxation towards equilibrium

We here establish a connection between the excess free energy at time t and the entropy $\Delta S([t,\infty))$ produced in the time-interval $[t,\infty)$, i.e., in the time interval where $p(\mathbf{x},t)$ relaxes towards $p_{\rm s}(\mathbf{x})$. From the first line of Eq. (5.2.2), and using Eq. (2.9.3) for the entropy production in the medium for reversible dynamics, we see that for relaxation towards equilibrium

$$\begin{aligned} F(t)-F(\infty) &= \langle \phi(\mathbf{x}_t)-\phi(\mathbf{x}_\infty)\rangle - T[S_{\rm sys}(t)-S_{\rm sys}(\infty)] \\ &= \left\langle \int_{\tau=t}^{\tau=\infty} -\nabla\phi(\mathbf{x}_\tau)\circ \mathrm{d}\mathbf{x}_\tau \right\rangle + T\Delta S_{\rm sys}([t,\infty)) \\ &= T\Delta S_{\rm med}([t,\infty)) + T\Delta S_{\rm sys}([t,\infty))\,. \end{aligned} \tag{5.2.4}$$

The definition (5.2.3) and Eq. (5.2.4) thus yield for reversible systems

$$\mathcal{D}(t) = T\Delta S_{\rm tot}([t,\infty))\,. \tag{5.2.5}$$

Relaxation towards NESS

For relaxation towards a NESS Eq. (5.2.5) no longer holds. However, for NESS there exists another useful splitting of the total entropy production (i.e., not only the splitting into $\Delta S_{\rm med}$ and $\Delta S_{\rm sys}$), namely into *adiabatic* and *non-adiabatic* contributions [188] (where we define $\boldsymbol{\nu}_{\rm s}(\mathbf{x}) \equiv \mathbf{j}_{\rm s}(\mathbf{x})/p_{\rm s}(\mathbf{x})$ and $\boldsymbol{\nu}(\mathbf{x},t) \equiv \mathbf{j}(\mathbf{x},t)/p(\mathbf{x},t)$ with $\mathbf{j}(\mathbf{x},t) \equiv [\mathbf{a}(\mathbf{x}) - \mathbf{D}\nabla]p(\mathbf{x},t)$, see Eq. (2.6.5)),

$$\begin{aligned} \Delta S_{\rm tot}([t_1,t_2]) &= \Delta S_{\rm non-adiabatic}([t_1,t_2]) + \Delta S_{\rm adiabatic}([t_1,t_2]) \\ \Delta S_{\rm adiabatic}([t_1,t_2]) &\equiv k_{\rm B}\int_{t_1}^{t_2} \mathrm{d}\tau \left\langle \boldsymbol{\nu}_{\rm s}^T(\mathbf{x}_\tau)\mathbf{D}^{-1}\boldsymbol{\nu}_{\rm s}(\mathbf{x}_\tau)\right\rangle \geq 0 \\ \Delta S_{\text{non-adiabatic}}([t_1,t_2]) &\equiv k_{\rm B}\int_{t_1}^{t_2} \mathrm{d}\tau \left\langle [\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)-\boldsymbol{\nu}_{\rm s}(\mathbf{x}_\tau)]^T\mathbf{D}^{-1}[\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)-\boldsymbol{\nu}_{\rm s}(\mathbf{x}_\tau)]\right\rangle \geq 0\,. \end{aligned} \tag{5.2.6}$$

Here, most of the time dependence of the process is contained in the non-adiabatic part, while the adiabatic part only features time dependence through the average $\langle\cdots\rangle$ taken over $p(\mathbf{x},t)$. Note that, unlike the splitting into system and medium entropy production, here, not only the total entropy production but also both individual contributions are always non-negative.

To derive Eq. (5.2.6), note that

$$\begin{aligned}\boldsymbol{\nu}(\mathbf{x},\tau)-\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x}) &= \frac{[\mathbf{a}(\mathbf{x})-\mathbf{D}\nabla]p(\mathbf{x},\tau)}{p(\mathbf{x},\tau)} - \frac{[\mathbf{a}(\mathbf{x})-\mathbf{D}\nabla]p_{\mathrm{s}}(\mathbf{x})}{p_{\mathrm{s}}(\mathbf{x})} \\ &= -\mathbf{D}\nabla\ln\left[\frac{p(\mathbf{x},\tau)}{p_{\mathrm{s}}(\mathbf{x})}\right]. \end{aligned} \tag{5.2.7}$$

Recall that $\Delta S_{\mathrm{tot}}([t_1,t_2]) = k_{\mathrm{B}}\int_{t_1}^{t_2}\mathrm{d}\tau\langle\boldsymbol{\nu}^T(\mathbf{x}_\tau,\tau)\mathbf{D}^{-1}\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)\rangle$, see Eq. (2.9.8), and use $\mathbf{D}=\mathbf{D}^T$, Eq. (5.2.7) and $\nabla\cdot\mathbf{j}_{\mathrm{s}}=0$ to write

$$\begin{aligned}\partial_\tau\big[\Delta S_{\mathrm{tot}}([t_1,\tau]) &- \Delta S_{\text{adiabatic}}([t_1,\tau]) - \Delta S_{\text{non-adiabatic}}([t_1,\tau])\big] \\ &= 2k_{\mathrm{B}}\left\langle[\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)-\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x}_\tau)]^T\mathbf{D}^{-1}\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x}_\tau)\right\rangle \\ &= -2k_{\mathrm{B}}\int\mathrm{d}\mathbf{x}\,p(\mathbf{x},\tau)\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x})\cdot\nabla\ln\left[\frac{p(\mathbf{x},\tau)}{p_{\mathrm{s}}(\mathbf{x})}\right] \\ &= -2k_{\mathrm{B}}\int\mathrm{d}\mathbf{x}\,p_{\mathrm{s}}(\mathbf{x})\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x})\cdot\nabla\left[\frac{p(\mathbf{x},\tau)}{p_{\mathrm{s}}(\mathbf{x})}\right] \\ &= 2k_{\mathrm{B}}\int\mathrm{d}\mathbf{x}\,\frac{p(\mathbf{x},\tau)}{p_{\mathrm{s}}(\mathbf{x})}\nabla\cdot\mathbf{j}_{\mathrm{s}}(\mathbf{x}) \\ &= 0\,. \end{aligned} \tag{5.2.8}$$

This proves the decomposition in Eq. (5.2.6).

For equilibrium states, we have $\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x})=\mathbf{0}$ for all $\mathbf{x}$, such that $\Delta S_{\text{adiabatic}}([t,\infty))=0$ and $\mathcal{D}(t)/T=\Delta S_{\mathrm{tot}}([t,\infty))=\Delta S_{\text{non-adiabatic}}([t,\infty))$. For NESS we have $\Delta S_{\text{adiabatic}}([t,\infty))\neq 0$ such that $\Delta S_{\mathrm{tot}}([t,\infty))\neq\Delta S_{\text{non-adiabatic}}([t,\infty))$ but one can show that

$$\mathcal{D}(t) = T\Delta S_{\text{non-adiabatic}}([t,\infty))\,. \tag{5.2.9}$$

To derive Eq. (5.2.9), i.e., to see that $\Delta S_{\text{non-adiabatic}}([t,\infty))$ as defined in Eq. (5.2.6) agrees with the Kullback-Leibler divergence in Eq. (5.2.3), we write (see also Ref. [188]; in the last step we use $\langle\partial_\tau\ln[p(\mathbf{x},\tau)/p_{\mathrm{s}}(\mathbf{x})]\rangle = \langle\{\partial_\tau p(\mathbf{x},\tau)\}/p(\mathbf{x},\tau)\rangle = \int\mathrm{d}\mathbf{x}\partial_\tau p(\mathbf{x},\tau) = \partial_\tau\int\mathrm{d}\mathbf{x}p(\mathbf{x},\tau)=\partial_\tau 1=0$)

$$\begin{aligned}\partial_\tau\Delta S_{\text{non-adiabatic}}([\tau,\infty))/k_{\mathrm{B}} &\overset{\text{Eq. (5.2.6)}}{=} -\left\langle[\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)-\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x}_\tau)]^T\mathbf{D}^{-1}[\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)-\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x}_\tau)]\right\rangle \\ &\overset{\text{Eq. (5.2.8)}}{=} -\left\langle\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)^T\mathbf{D}^{-1}[\boldsymbol{\nu}(\mathbf{x}_\tau,\tau)-\boldsymbol{\nu}_{\mathrm{s}}(\mathbf{x}_\tau)]\right\rangle \\ &\overset{\text{Eq. (5.2.7)}}{=} -\int\mathrm{d}\mathbf{x}\,p(\mathbf{x},\tau)\boldsymbol{\nu}(\mathbf{x},\tau)\cdot\nabla\ln\left[\frac{p(\mathbf{x},\tau)}{p_{\mathrm{s}}(\mathbf{x})}\right] \\ &\overset{\text{Eq. (2.6.5)}}{=} \int\mathrm{d}\mathbf{x}\,\{\partial_\tau p(\mathbf{x},\tau)\}\ln\left[\frac{p(\mathbf{x},\tau)}{p_{\mathrm{s}}(\mathbf{x})}\right] \end{aligned}$$

$$= \quad \partial_\tau \int \mathrm{d}\mathbf{x}\, p(\mathbf{x},\tau) \ln\left[\frac{p(\mathbf{x},\tau)}{p_\mathrm{s}(\mathbf{x})}\right] . \tag{5.2.10}$$

Using that both, $\Delta S_{\text{non-adiabatic}}([\tau,\infty))$ and $\mathcal{D}(\tau)$, vanish as $\tau \to \infty$, Eq. (5.2.10) implies Eq. (5.2.9) by integrating over τ.

In summary, we see that while $\mathcal{D}(t) = T\Delta S_{\text{tot}}([t,\infty))$ only holds for relaxation towards equilibrium, $\mathcal{D}(t) = T\Delta S_{\text{non-adiabatic}}([t,\infty))$ also holds for relaxation towards NESS. This result strengthens the interpretation of $\mathcal{D}(t)$ as a useful observable to quantify the thermodynamics of relaxation since $\Delta S_{\text{non-adiabatic}}([t,\infty))$ is known to be exactly the part of the entropy production relevant to relaxation processes, while $\Delta S_{\text{adiabatic}}([t,\infty))$ merely constitutes the cost of maintaining the NESS [28,188]. Therefore, we will henceforth use $\mathcal{D}(t)$ to quantify relaxation in both, equilibrium and non-equilibrium steady states.

5.2.3 Thermal kinematics

Above we gave strong arguments for using the (generalized) excess free energy in Eq. (5.2.3) to quantify relaxation, and we do so in the publications reproduced in Secs. 5.4 and 5.5. However, there is one disadvantage of this approach. The Kullback-Leibler divergence is not a true metric since it is not symmetric and does not satisfy the triangle inequality, i.e., it does not represent a geometrical notion of distance. In particular, this has the disadvantage that the initial ($t = 0$) excess free energy of a system at temperature T_1 relaxing towards T_2 is not the same as in the scenario with T_1 and T_2 interchanged (note in particular that this initial excess free energy is *not* the difference in equilibrium free energies, since only the equilibrium free energy at the target temperature appears in Eq. (5.2.3)). This complicates (although not at all prevents) the comparison of heating and cooling processes, see Ref. [73] and following sections. To go beyond the quantification via $\mathcal{D}(t)$, we use information theoretic concepts to develop in Sec. 5.4 an additional, alternative framework, which we call *thermal kinetmatics*, that remedies these shortcomings, though this happens at the cost of losing the thermodynamic notion presented in the preceding paragraphs.

5.3 Summary of results

Before we present the two publications in detail in Secs. 5.4 and 5.5, we briefly outline the results of these publications, and put them into the context of the previous two sections.

Both works are based on Ref. [73], where thermal relaxation towards an equilibrium state was studied. Starting from two different initial temperatures, one hotter and one colder

than the equilibrium temperature, it was found that, beyond the linear regime, heating and cooling become asymmetric and, in particular, heating is always faster than cooling for several systems, including in particular Langevin motion in a harmonic potential.

In the publication reproduced in Sec. 5.4 we experimentally confirm this result. From a theoretical perspective, we also go beyond Ref. [73] by showing that heating is also faster than cooling in the reversed protocol, i.e., when letting the system relax back towards the initial hot and cold temperature. More importantly, we develop the framework of thermal kinematics, see Subsec. 5.2.3, and use it to show that in one-dimensional harmonic potentials, heating is also faster than cooling between any two temperatures. The new theoretical results are also experimentally confirmed.

In the publication reproduced in Sec. 5.5, we study thermal relaxation towards NESS, again for harmonic potentials. We find that in terms of the excess free energy in Eq. (5.2.3), heating is still faster than cooling, i.e., the thermal relaxation asymmetry remains unchanged under non-equilibrium driving. We moreover study several interesting features induced by the non-equilibrium driving, in particular the phenomenon of *accelerated relaxation* and the emergence of counterintuitive rotational motions opposite to the direction of the rotational driving.

5.4 Heating and cooling are fundamentally asymmetric and evolve along distinct pathways (Nat. Phys. 2024)

This section is a slightly adapted version of the publication M. Ibáñez, C. Dieball, A. Lasanta, A. Godec and R. Rica, "Heating and cooling are fundamentally asymmetric and evolve along distinct pathways", *Nat. Phys.* (2024), available under the doi 10.1038/s41567-023-02269-z.

Unlike all other publications included in this thesis, I was not the first author for the manuscript reproduced in this section. My contribution to this manuscript was in developing the theoretical framework, carrying out the proofs, discussing the results, and co-writing the manuscript. I did not perform any measurements on the experiment, and did not contribute significantly to the second 'Supplementary Material' and to the 'Materials and Methods'-section reproduced in Subsecs. 5.4.8 and 5.4.9, respectively.

Heating and Cooling are Fundamentally Asymmetric and Evolve Along Distinct Pathways

Miguel Ibáñez[1,2], **Cai Dieball**[3], **Antonio Lasanta**[2,4,5], **Aljaž Godec**[3] **and Raúl A. Rica**[1,2]

Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

[1]Universidad de Granada, Department of Applied Physics and Research Unit 'Modeling Nature' (MNat), Av. Fuentenueva s.n., 18071 Granada, Spain

[2]Nanoparticles Trapping Laboratory, Universidad de Granada, 18071 Granada, Spain

[3]Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

[4]Universidad de Granada, Department of Algebra, Facultad de Educación, Economía y Tecnología de Ceuta, Universidad de Granada, Cortadura del Valle, s/n, 51001 Ceuta, Spain

[5]Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, 18071 Granada, Spain

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Abstract

According to conventional wisdom, a system placed in an environment with a different temperature tends to relax to the temperature of the latter, mediated by the flows of heat and/or matter that are set solely by the temperature difference. It is becoming clear, however, that thermal relaxation is much more intricate when temperature changes push the system far from thermodynamic equilibrium. Here, we ask whether microscale systems under such conditions heat up faster than they cool down. Using an optically trapped colloidal particle, we experimentally confirm this is indeed generally true. Strikingly, we show with both experiments and theory that between any pair of temperatures, heating is not only faster than cooling, but the respective processes in fact evolve along fundamentally distinct pathways, which we explain with a new theoretical framework we coin "thermal kinematics". Our results profoundly change the view on thermalization at the microscale, and will have a strong impact on energy-conversion applications and thermal management of microscopic devices, in particular in the operation of Brownian heat engines.

5.4.1 Introduction

The basic laws of thermodynamics dictate that any system in contact with an environment eventually relaxes to the temperature of its surroundings as a result of irreversible flows that drive the system to thermodynamic equilibrium. If the difference between the initial temperature of the system and that of the surroundings is small, i.e., the system is initially "close to equilibrium" [25], the relaxation is typically assumed to evolve quasi-statically through local equilibrium states—an assumption that is justifiable only *a posteriori* [25, 250]. However, if the temperature contrast is such that it pushes the system far from equilibrium, the assumption breaks down and the relaxation path is no longer unique, but depends strongly on the initial condition. This gives rise to counterintuitive phenomena, such as anomalous relaxation (also known as the Mpemba effect) [251–257] where, a system reaches equilibrium faster upon a stronger temperature quench, and the so-called Kovacs memory effect [258–262] which features a non-monotonic evolution towards equilibrium.

Intriguingly, thermal relaxation was recently predicted to depend also on the sign of the temperature change. Namely, considering two *thermodynamically equidistant* (TE) temperatures—one higher and the other lower than an intermediate one selected such that the initial free energy difference with the equilibrium state is the same—heating from the colder temperature was predicted to be faster than cooling from the hotter one [73, 75]. This prediction challenges our understanding of non-equilibrium thermodynamics, as it compares reciprocal relaxation processes elusive to classical thermodynamics. The initially hotter system must dissipate into the environment an excess of both, energy and entropy, whereas in the colder system energy and entropy must increase [73, 74]. Moreover, the comparison of heating and cooling provokes an even more fundamental question, namely that of reciprocal relaxation processes between two *fixed* temperatures. According to the "local equilibrium" paradigm [25] the system relaxes quasi-statically and thus traces the same path along reciprocal processes. We show, however, that this is *not* the case: heating and cooling are inherently asymmetric and evolve along distinct pathways.

In this work, we use colloidal particles in temperature-modulated optical traps to interrogate relaxation kinetics upon temperature quenches (see Fig. 5.4.1), and unveil three fundamental asymmetries between heating and cooling. We experimentally confirm the prediction that heating is faster than cooling in three complementary situations, precisely (i) that heating from a colder temperature towards an intermediate target temperature is faster than cooling from the corresponding TE hotter temperature [73]. Unexpectedly, we also show (ii) that the reverse process, i.e., heating from the intermediate temperature to a hotter temperature is faster than cooling to the corresponding TE colder temperature. Most surprisingly, we show (iii) that between a fixed pair of temperatures, heating is faster

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

than the reciprocal cooling. In all cases, we provide mathematical proofs that establish these asymmetries as a general feature of systems with (at least locally) quadratic energy landscapes.

A key result is that the production of entropy within the system during heating is more efficient than heat dissipation during cooling. Asymmetries (ii) and (iii) further imply that the microscopic relaxation paths during heating and cooling are distinct. Moreover, whereas a system prepared at TE temperatures is by construction equally far from equilibrium in terms of free energy, we show that the colder system is in fact statistically farther from equilibrium and yet heating from said colder temperature is faster. Developing a new framework we coin "thermal kinematics" we explain the asymmetry by means of the propagation in the space of probability distributions, which is intrinsically faster during heating.

5.4.2 Heating and cooling at thermodynamically equidistant conditions

Thermal relaxation kinetics beyond the "local equilibrium" regime can be quantified within the framework of Stochastic Thermodynamics [28, 53, 69] which requires the knowledge of statistics of all slow, mesoscopic degrees of freedom. In the present work, where we use a colloidal particle with a diameter of 1 μm in a tightly focused laser (see Fig. 5.4.1), the overdamped regime ensures that only the position has to be analyzed [28,263,264]. Due to the symmetry of the tweezers setup, it suffices to follow a single coordinate of the particle as a function of time, which we denote by x_t. We consider two different initial conditions. By the nature of the setup (see Fig. 5.4.1) x_t is initially in equilibrium in the optical potential $U(x)$ at either the "hot", T_h, or "cold", T_c, temperature, respectively, with a probability density $P^j_{\rm eq}(x) = {\rm e}^{(F_j - U(x))/k_{\rm B}T_j}$ where $F_j \equiv -k_{\rm B}T_j \ln \int_{-\infty}^{\infty} {\rm e}^{-U(x)/k_{\rm B}T_j} dx$ is the equilibrium free energy at temperature T_j. In the following, equilibrium probability densities are denoted by $P^j_{\rm eq}(x)$, where $j = h, w, c$ refers to the bath temperature T_j. Observables with both, subscript and superscript $i, f = h, w, c$, e.g., A^f_i, denote transient observables where the subscript refers to the initial and the superscript to the target state. The state of the system at any time is fully specified by, $P^f_i(x,t)$, the probability density of the particle's position at time t.

We first focus on the forward protocol, where the relaxation occurs at the "warm" temperature T_w and $i = h, c$. The dynamics is ergodic and therefore $P^w_i(x,t)$ relaxes towards $P^w_{\rm eq}(x)$. We use $\langle \dots \rangle^w_i$ to denote averages over $P^w_i(x,t)$ and quantify the instantaneous displacement from the equilibrium distribution $P^w_{\rm eq}(x)$ by means of the *generalized excess free energy* given by [73, 120, 265] $\mathcal{D}^w_i(t) = \langle U(x_t) \rangle^w_i - T_w S^w_i(t) - F_w =$

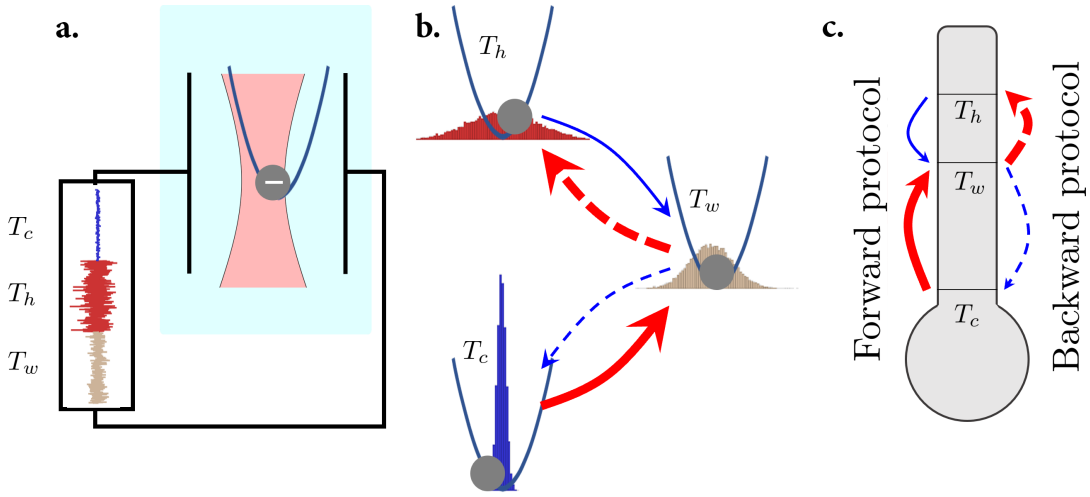

Figure 5.4.1: Setup for probing the heating-cooling asymmetry. a. Schematic representation of the experiment. A charged dielectric microparticle dispersed in water is confined in a parabolic trap generated by a tightly focused infrared laser. Its effective temperature is controlled by an electric field that shakes the particle, mimicking a thermal bath at a higher temperature than the water. An arbitrary signal generator feeds a noisy signal with a Gaussian-white spectrum into a pair of gold microelectrodes immersed in the liquid, thus producing the required electric field. Therefore, the particle exhibits Brownian motion inside the trap, featuring a Gaussian distribution whose variance is determined by the effective temperature. **b.** In experiments, we track the evolution of the position distribution upon quenches of the effective thermal bath during heating (red arrows) and cooling (blue arrows). **c.** Schematic representation of the respective protocols: in the forward protocol, the system is initially prepared at equilibrium with the thermal bath with a temperature higher (T_h) or lower (T_c) than the target (T_w) temperature. T_h and T_c are chosen to be thermodynamically equidistant from T_w with $T_h > T_w > T_c$. During the backward protocol, the system relaxes at the respective thermodynamically equidistant temperatures T_h and T_c, starting from a common initial condition that is the equilibrium at T_w. In a third situation, only two temperatures are compared, considering the evolution of the system upon heating and cooling between them. In **b** and **c** solid and dashed arrows stand for the forward and backward process, respectively, and thick lines indicate faster evolution than thin ones.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

$k_{\rm B}T_w D[P_i^w(x,t)||P_{\rm eq}^w(x)]$ for $i=h,c$, where $S_i^w(t)=-k_{\rm B}\langle \ln P_i^w(x_t,t)\rangle_i^w$ is the Gibbs entropy and $D[P||Q]=\int P\ln(P/Q)dx$ is the relative entropy between the probability distributions P and Q. This generalizes the notion of the canonical free energy $F=U-TS$ to instantaneous non-equilibrium configurations with a probability density $P(x,t)$. Temperatures T_h and T_c are said to be TE from T_w when the initial excess free energies are equal, i.e., $\mathcal{D}_h^w(0)=\mathcal{D}_c^w(0)$ [73]. The unexpected prediction was made in Ref. [73] that $\mathcal{D}_c^w(t)<\mathcal{D}_h^w(t)$ at all times $t>0$. That is, the system heats up to the temperature of its surroundings faster than it cools down. Albeit an asymmetric relaxation is counter-intuitive, our experiments quantitatively corroborate this prediction to be true, see Fig. 5.4.2.

What may be even more surprising is, heating also turns out to be faster along the reversed, *backward* protocol. That is, we prepare the system to be in equilibrium at the "warm" temperature T_w and track the relaxations at T_h and T_c, respectively. Likewise, we quantify the kinetics via the relative entropy $\mathcal{D}_w^i(t)\equiv k_{\rm B}T_w D[P_w^i(x,t)||P_{\rm eq}^w(x)]$, such that $\mathcal{D}_w^h(t)$ and $\mathcal{D}_w^c(t)$ evolve from zero and asymptotically converge to $\mathcal{D}_w^{h/c}(\infty)=\mathcal{D}_{h/c}^w(0)$. Although $\mathcal{D}_w^i(t)$ sensibly quantifies the departure from $P_{\rm eq}^w(x)$, it is strictly speaking *not* an excess free energy, in contrast to $\mathcal{D}_i^w(t)$ defined above, because T_w and $P_{\rm eq}^w(x)$ no longer refer to the target equilibrium, i.e., because $\mathcal{D}_i^w(t)=\langle U\rangle_w^i-T_wS_w^i(t)-F_w\neq\langle U\rangle_w^i-T_iS_w^i(t)-F_w$ where the latter would correspond to an excess free energy. We observe in Fig. 5.4.2 that $\mathcal{D}_w^h(t)>\mathcal{D}_w^c(t)$ for all times $t>0$, i.e., the system heats up to the new equilibrium at T_h faster than it cools back to T_c. This observation is remarkable as it shows that heating is inherently faster than cooling at TE conditions.

To confirm these observations theoretically, we assume that the particle's dynamics evolve in a parabolic potential with stiffness κ, $U(x)=\kappa x^2/2$, according to the overdamped Langevin equation $dx_t=-(\kappa/\gamma)x_t dt+d\xi_t^i$ with friction constant γ given by the Stokes' law $\gamma=6\pi r\eta$ where η is the viscosity of water. The thermal noise $d\xi_t^i$, where $i=h,w,c$ denotes the temperature of the reservoir, vanishes on average and obeys the fluctuation-dissipation theorem $\langle d\xi_t^i d\xi_{t'}^i\rangle=2(k_{\rm B}T_i/\gamma)\delta(t-t')dtdt'$. Under these assumptions we determine TE temperatures T_h and $T_c(T_h)$, i.e., we calculate T_c after we arbitrarily set T_h, (see Eq. (5.4.12) in the Supplemental Material in Subsec. 5.4.7) and $\mathcal{D}_{i/w}^{w/i}(t)$ reads

$$\mathcal{D}_{i/w}^{w/i}(t)=\frac{k_{\rm B}T_w}{2}[\Lambda_{i/w}^{w/i}(t)-1-\ln\Lambda_{i/w}^{w/i}(t)], \tag{5.4.1}$$

where $\Lambda_i^w(t)=1+(T_i/T_w-1)\mathrm{e}^{-2(\kappa/\gamma)t}$ and $\Lambda_w^i(t)=T_i/T_w+(1-T_i/T_w)\mathrm{e}^{-2(\kappa/\gamma)t}$. We consider $\mathcal{D}_i^w(t)$ during the forward and $\mathcal{D}_w^i(t)$ during the backward protocol. According to Eq. (5.4.1), by plotting $\mathcal{D}_{i/w}^{w/i}(t)/k_{\rm B}T_w$ as a function of $\rho=\Lambda_{i/w}^{w/i}(t)$ all data should collapse onto the master curve $f(\rho)=(\rho-1-\ln\rho)/2$, which is indeed what we observe in Fig. 5.4.2.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

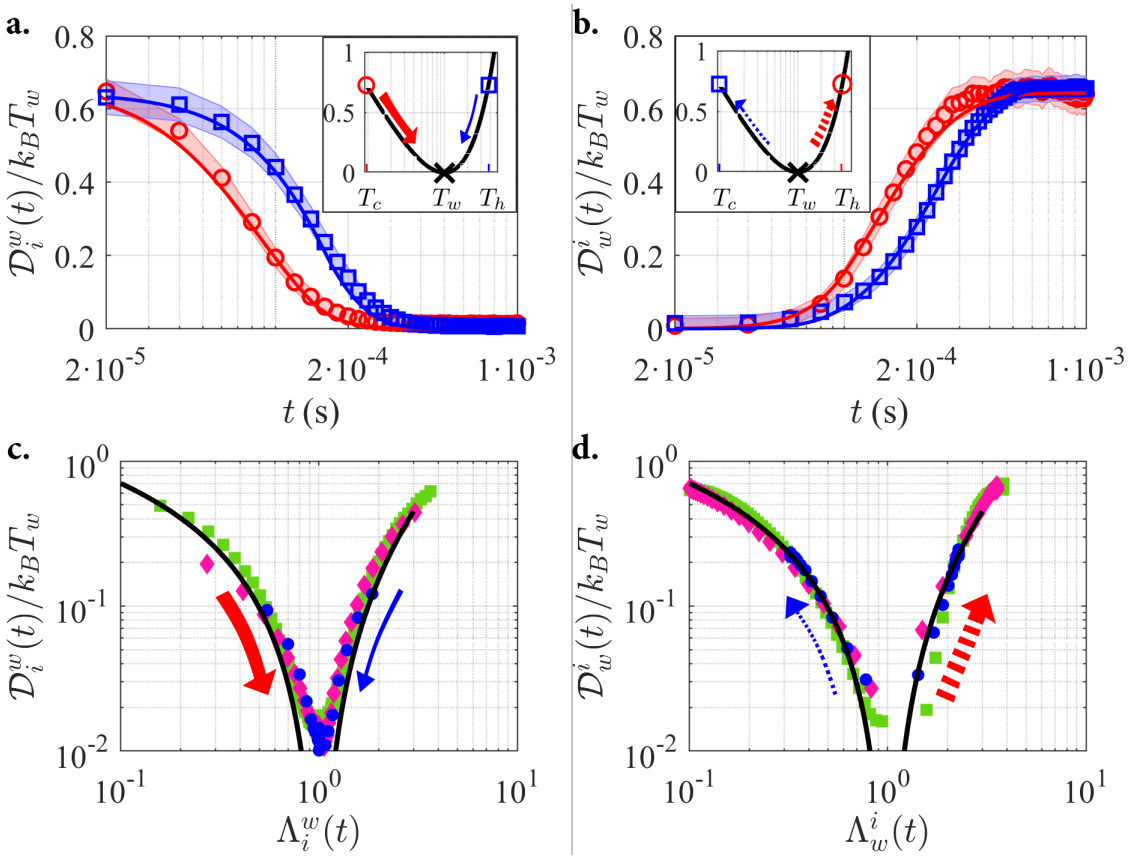

Figure 5.4.2: Experimental evolution of the generalized excess free energy during heating and cooling at thermodynamically equidistant conditions. Panels **a** and **c** correspond to the forward protocol and **b** and **d** to the backward counterpart. Thick, red arrows stand for heating, while blue, thin arrows represent cooling. **a., b.** Time evolution of the generalized excess free energy for a characteristic time $\tau = \gamma/\kappa = 0.1844(3)$ ms, $T_c/T_w = 0.11(1)$, $T_h/T_w = 3.56(1)$. Red circles stand for heating and blue squares stand for cooling. Solid lines correspond to the theoretical predictions without fitting parameters. Insets represent the initial value of the relative entropy $\mathcal{D}_i^w(0)/k_{\mathrm{B}}T_w$ (y-axis) as a function of the temperature (x-axis) on the logarithmic scale. The arrows represent the evolution direction along the master curve $f(\rho) = (\rho - 1 - \ln\rho)/2$. The confidence regions have been estimated by quadratic uncertainty propagation from the standard deviation of the experimental histograms. **c., d.** Generalized excess free energy $\mathcal{D}_{i/w}^{w/i}(t)/k_{\mathrm{B}}T_w$ as a function of $\Lambda_{i/w}^{w/i}(t)$, along the master curve $f(\rho)$, for several different TE conditions. The corresponding time series are included in the Supplemental Material (Fig. S5).

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Having established the validity of the model, we prove (see Theorem 1 in Subsec. 5.4.7) that our observations hold for all TE temperatures and for any κ and γ, i.e.,

$$\mathcal{D}_c^w(t) < \mathcal{D}_h^w(t) \quad \text{and} \quad \mathcal{D}_w^h(t) > \mathcal{D}_w^c(t), \quad \text{for all } 0 < t < \infty. \tag{5.4.2}$$

Our observations in Fig. 5.4.2 and the inequalities (5.4.2) establish rigorously that at TE conditions, heating is faster than cooling.

Notwithstanding, these results are still unsatisfactory for two reasons. First, $\mathcal{D}_w^i(t)$, unlike $\mathcal{D}_i^w(t)$, lacks a consistent thermodynamic interpretation and, in particular, is not an excess free energy. Second, neither the relative entropy $D[P||Q]$ nor $\sqrt{D[P||Q]}$ are a true metric; they are not symmetric, $D[P||Q] \neq D[Q||P]$, and do not satisfy triangle inequalities. The latter in particular implies that while a Pythagorean theorem holds (see [124], and Fig. 5.4.10 in Subsec. 5.4.7 for an experimental validation), $D[P^i_{\rm eq}(x)||P^w_{\rm eq}(x)] \geq D[P^i_{\rm eq}(x)||P^i_w(x,t)] + D[P^i_w(x,t)||P^w_{\rm eq}(x)]$, where we used $P^i_w(x,0) = P^i_{\rm eq}(x)$, the triangle inequality does *not*, i.e., in general $\sqrt{D[P^i_{\rm eq}(x)||P^w_{\rm eq}(x)]} \not\leq \sqrt{D[P^i_{\rm eq}(x)||P^i_w(x,t)]} + \sqrt{D[P^i_w(x,t)||P^w_{\rm eq}(x)]}$. As a result, $\mathcal{D}^{w/i}_{i/w}(t)$ does not measure "distance" from equilibrium and $\partial_t \mathcal{D}^{w/i}_{i/w}(t)$ is therefore not a "velocity", which seems to preclude a kinematic description of relaxation. In the following sections, we show that combining stochastic thermodynamics with information geometry [266] makes it possible to formulate a *thermal kinematics*.

5.4.3 Towards thermal kinematics

Immense progress has been made in recent years in understanding non-equilibrium systems. The discovery of thermodynamic uncertainty relations [47, 48, 87, 134], and "speed limits" [124, 266–268] revealed that the entropy production rate, which quantifies irreversible local flows in the system, universally bounds fluctuations and the rate of change, respectively, in a non-equilibrium system. Closely related is the so-called Fisher information known from information geometry, which quantifies how local flows change in time and allows for defining a *statistical distance* [266, 269–271].

In our context of thermal relaxation, an infinitesimal statistical line element may be defined as follows. Since $D[P^f_i(x,t+dt)||P^f_i(x,t)] = I^f_i(t)dt^2 + \mathcal{O}(dt^3)$ (see Eq. (5.4.22) in Subsec. 5.4.7), where we introduced the Fisher information $I^f_i(t) \equiv \langle(\partial_t \ln P^f_i(x,t))^2\rangle^f_i$, we can define the line element as $dl \equiv \sqrt{D[P^f_i(x,t+dt)||P^f_i(x,t)]} = \sqrt{I^f_i(t)}dt$ and thus $v^f_i(t) \equiv \sqrt{I^f_i(t)}$ is the *instantaneous statistical velocity* of the system [266] relaxing from $P^i_{\rm eq}(x)$ at temperature T_f towards $P^f_{\rm eq}(x)$. The statistical length traced by $P^f_i(x,\tau)$ until

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

time t is $\mathcal{L}_i^f(t) = \int_0^t v_i^f(\tau)d\tau$ and the distance between initial and final states is thus given by $\mathcal{L}_i^f(\infty)$ and does *not* depend on the direction, i.e., $\mathcal{L}_a^b(\infty) = \mathcal{L}_b^a(\infty)$, for two different temperatures T_a and T_b. To establish a kinematic basis for quantifying thermal relaxation kinetics, we define the *degree of completion* $\varphi_{i/w}^{w/i}(t) \equiv \mathcal{L}_{i/w}^{w/i}(t)/\mathcal{L}_i^w(\infty)$, which increases monotonically between 0 and 1.

Assuming that the system evolves according to overdamped Langevin dynamics in a parabolic potential, we find (see Eqs. (5.4.27) and (5.4.28) in Subsec. 5.4.7) $\mathcal{L}_i^w(\infty) = |\ln(T_i/T_w)|/\sqrt{2}$ and

$$\varphi_{i/w}^{w/i}(t) \quad = \quad 1 - \frac{\ln(1 + (T_{i/w}/T_{w/i} - 1)\mathrm{e}^{-2(\kappa/\gamma)t})}{\ln(T_{i/w}/T_{w/i})}. \tag{5.4.3}$$

Moreover, we prove (see Theorem 2 in Subsec. 5.4.7) for any pair of TE temperatures T_h, T_c that $\mathcal{L}_c^w(\infty) > \mathcal{L}_h^w(\infty)$ and yet

$$\varphi_c^w(t) > \varphi_h^w(t) \quad \text{and} \quad \varphi_w^h(t) > \varphi_w^c(t) \quad \text{for all } 0 < t < \infty. \tag{5.4.4}$$

That is, the colder system is statistically farther from equilibrium than the hotter system, but nevertheless, heating is faster than cooling. On the one hand, Eq. (5.4.4) confirms the asymmetry (5.4.2) from a kinematic point of view. On the other hand, it reveals something more striking; during heating in the forward protocol, the system traces a longer path in the space of probability distributions, but it does so faster. The reason lies in the propagation speed $v_{w/i}^{i/w}(t)$ at short times that is intrinsically larger during heating than during cooling. This speed-up is due to the entropy production in the system during heating, being more efficient than heat flow from the system to the environment during cooling (see also [73]). These predictions are fully confirmed by experiments (see Fig. 5.4.3). The results show that an initial overshoot in $v_c^w(t)$ and $v_w^h(t)$ ensures that under TE conditions heating is, in both protocols, at all times faster than cooling according to the inequalities (5.4.4). Since both processes relax to the same equilibrium, $v_c^w(t)$ and $v_w^h(t)$ eventually must cross.

5.4.4 Heating between any pair of temperatures is faster than cooling

We now take our thermal kinematics approach one step further and consider two arbitrary fixed temperatures $T_1 < T_2$ and observe heating, i.e., relaxation at T_2 in a temperature quench from an equilibrium prepared at T_1, and the reverse cooling, i.e., relaxation at T_1 in a temperature quench from the equilibrium at T_2. By construction, the distance between initial and final states along the reciprocal processes is the *same*, $\mathcal{L}_2^1(\infty) = \mathcal{L}_1^2(\infty)$.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

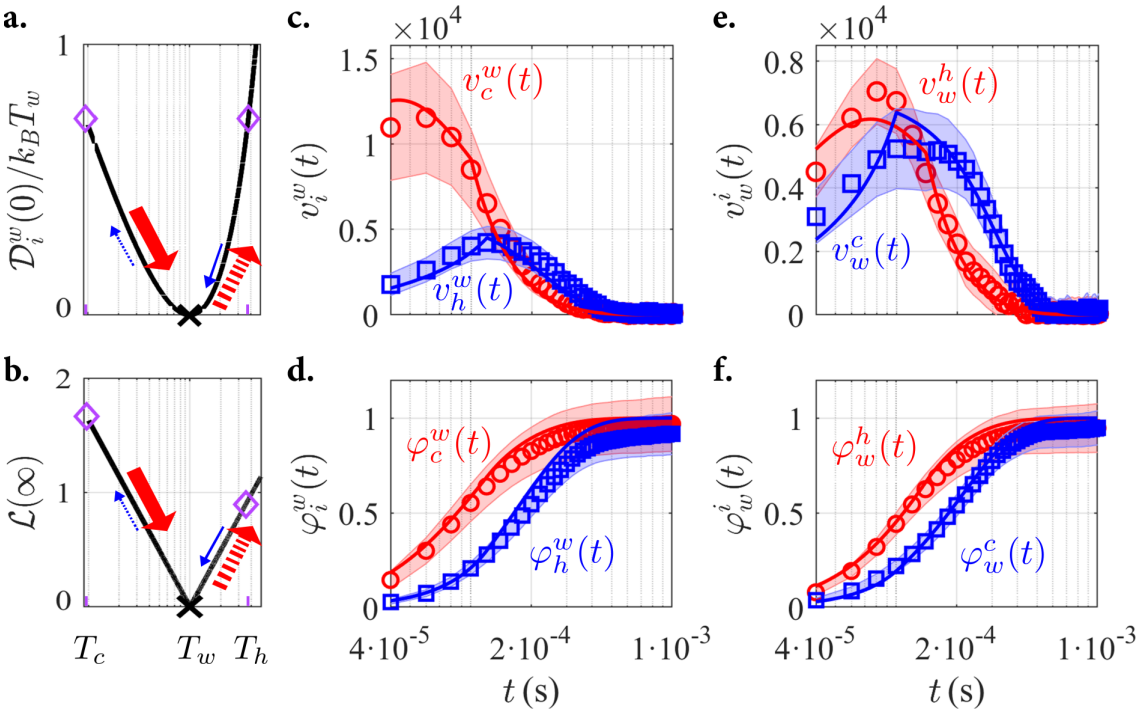

Figure 5.4.3: Thermal kinematics of heating and cooling processes at thermodynamically equidistant conditions. All data corresponds to the series shown in Fig. 5.4.2. As in previous figures, red arrows stand for heating while blue ones represent cooling, solid and dashed arrows refer to forward and backward protocol, respectively, and thicker lines indicate a faster evolution than thin ones. **a.** Initial value of the relative entropy $\mathcal{D}_i^w(0)/k_{\mathrm{B}}T_w$ as a function of the temperature. **b.** Total traversed statistical distance $\mathcal{L}_i^w(\infty) = \mathcal{L}_w^i(\infty) = \mathcal{L}(\infty)$ as a function of the temperature. **c-f** Temporal evolution of the instantaneous statistical velocity $v_i^w(t)$ (**c** and **e**) and the degree of completion $\varphi_i^w(t)$ (**d** and **f**) during the forward (**c** and **d**) and backward (**e** and **f**) protocol. Red circles stand for heating, while blue squares correspond to cooling. Solid lines are theoretical predictions without fitting parameters. Confidence regions have been estimated by quadratic uncertainty propagation from the standard deviation of the experimental histograms.

Nevertheless, according to our model, in particular Eqs. (5.4.3), we have for any $T_1 < T_2$ (see Theorem 3 in Subsec. 5.4.7) that

$$\varphi_1^2(t) > \varphi_2^1(t) \quad \text{for all} \quad 0 < t < \infty. \tag{5.4.5}$$

That is, between any pair of temperatures heating is faster than cooling, which is a much stronger statement. Notice that it was not possible to make such a statement based on the generalized excess free energy, since in this case, the description of the backward process lacked physical consistency. This result highlights that heating and cooling are inherently asymmetric processes, and that this is neither limited to the TE setting nor to strong quenches. The asymmetry (5.4.5) is fully corroborated by experiments (see Fig. 5.4.4). As before, it emerges due to an initial overshoot in $v_1^2(t)$. However, in this case, the difference in velocities implies that the pathway taken during heating is fundamentally different from the pathway followed during the reciprocal cooling process. Note that this goes beyond "hysteresis"-type asymmetries that occur during the heating and cooling of macroscopic systems through phase or glass transitions.

5.4.5 Near equilibrium, heating and cooling become symmetric

Finally, we show that for quenches near equilibrium heating and cooling are indeed almost symmetric, in agreement with linear non-equilibrium thermodynamics [25]. First, for $T_h = (1+\varepsilon)T_w$ with $0 < \varepsilon \ll 1$, we find that $T_c(T_h) = (1 - \varepsilon + \mathcal{O}(\varepsilon^2))T_w$ (see Eq. (5.4.39) in Subsec. 5.4.7), i.e., near equilibrium, TE temperatures are approximately equidistant from the ambient temperature T_w. Second, $\partial_t \mathcal{D}_i^w|_{t=0} = [2 - T_i/T_w - T_w/T_i]k_{\mathrm{B}}T_w\kappa/\gamma$ (see Eq. (5.4.40) in Subsec. 5.4.7) from which for small ε we obtain $(\partial_t \mathcal{D}_c^w|_{t=0} - \partial_t \mathcal{D}_h^w|_{t=0})\gamma/\kappa k_{\mathrm{B}}T_w = \mathcal{O}(\varepsilon^3)$, such that heating and cooling in terms of $\mathcal{D}_i^w$ become symmetric. Third, we find in this limit (see Corollary 4 in Subsec. 5.4.7) that $\varphi_w^i(t) = \varphi_i^w(t) = 1 - \mathrm{e}^{-2(\kappa/\gamma)t} + \mathcal{O}(\varepsilon)$ and

$$\varphi_w^c(t) = \varphi_w^h(t) + \mathcal{O}(\varepsilon) \quad \text{and} \quad \varphi_h^w(t) = \varphi_c^w(t) + \mathcal{O}(\varepsilon). \tag{5.4.6}$$

That is, for near-equilibrium quenches, i.e., for $T_{h,c} - T_w \ll T_w$, heating and cooling in terms of both $\mathcal{D}$ and φ are approximately symmetric, and the asymmetry is thus a genuinely far-from-equilibrium phenomenon, as claimed. Actually, even for values of $(T_{h,c} - T_w)/T_w$ on the order of 1, heating and cooling are virtually symmetric within the experimental error (see Fig. 5.4.6, lower row, and Fig. 5.4.9 in Subsec. 5.4.8).

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

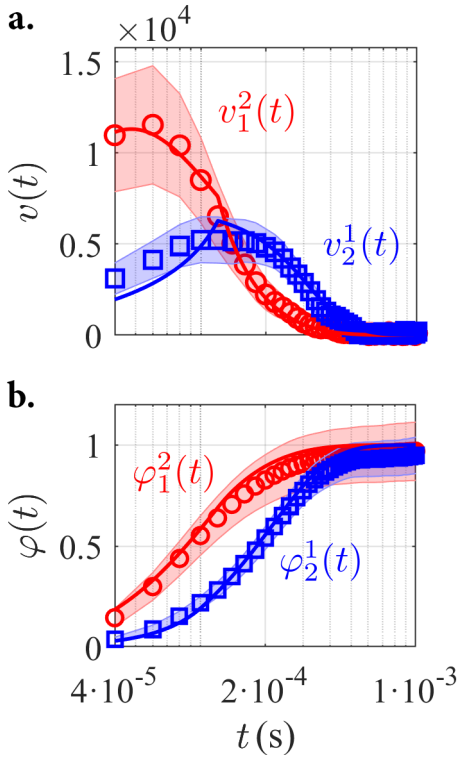

Figure 5.4.4: Thermal kinematics of heating and cooling between any pair of temperatures. The data shown corresponds to $T_1 = 302(3)$ K and $T_2 = 2753(7)$ K, and a characteristic time $\tau = \gamma/\kappa = 0.1844(3)$ ms. **a.** Instantaneous statistical velocity $v(t)$ as a function of time. **b.** Degree of completion $\varphi(t)$ as a function of time. Red circles stand for heating, while blue squares correspond to cooling. Solid lines are theoretical predictions without fitting parameters. The confidence regions have been estimated by quadratic uncertainty propagation from the standard deviation of the experimental histograms.

5.4.6 Discussion

Detailed experiments on colloidal particles corroborated by analytical theory revealed a fundamental asymmetry in thermal relaxation upon a rapid change of temperature; for thermodynamically equidistant temperature quenches as well as between two fixed temperatures, heating is always faster than cooling. Moreover, the microscopic pathways followed by a system during heating and cooling, respectively, are fundamentally different. Therefore, except very near to thermodynamic equilibrium, thermal relaxation, in general, does *not* evolve quasistatically through quasi-equilibria even for systems with a single energy minimum. We, therefore, witness a breakdown of the "near equilibrium" paradigm of classical non-equilibrium thermodynamics [25].

Namely, when the system is brought rapidly out of equilibrium, such as upon a temperature quench, the probability density of the system cannot follow the temperature change quasistatically and a *lag* develops between the instantaneous $P_i^w(x,t)$ and the new equilibrium $P_{\rm eq}^w(x)$ [120]. This lag, which here corresponds to $D[P_i^w(x,t)||P_{\rm eq}^w(x)]$, is nominally smaller during heating than during cooling. This is so because for short times, heating essentially corresponds to a free expansion [73], which is materialized as an overshoot of statistical velocity and is characterized by a smaller dissipated work. The latter in turn bounds from above the maximal lag that can develop [120]. Initial free expansion during heating also explains the faster departure from the initial equilibrium within the backward protocol, as well as for heating and cooling between any two fixed temperatures.

During heating, this initial free expansion leads to a rapid increase in system entropy and therefore, to a decrease in generalized excess free energy $\mathcal{D}(t) = \langle U\rangle(t) - T_w S(t) - F_w$. Cooling is dominated by the decrease in $\langle U\rangle$ which is, however, less efficient than the decay of system entropy during cooling. In turn, this causes the asymmetry for thermodynamically equidistant quenches [73]. Another intuitive argument for faster heating follows from spectral theory. Namely, the spectral representation of localized initial distributions nominally requires more eigenfunctions, in turn implying, since the initial distribution is normalized, a larger weight on the higher, faster-decaying modes [272, 273]. Since the colder system is more localized, one expects a faster relaxation during heating. The faster heating between any two temperatures, established here for the first time, may further be rationalized as follows. Microscopically, diffusion is driven by the collisions with the molecules in the medium. Since heating evolves at a higher temperature of the medium, collisions occur at a higher rate, in turn facilitating faster relaxation [274].

Our work further underscores that there is a fundamental difference between equidistant temperatures, thermodynamically equidistant temperatures, and kinematically equidis-

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

tant temperatures. The existence of a zero temperature and the $\mathrm{e}^{-1/T}$-dependence of Boltzmann-Gibbs equilibrium statistics readily imply that raising and lowering the temperature by the same amount pushes the system differently far from equilibrium. However, even when temperatures are chosen to be thermodynamically equidistant, the colder system is kinematically farther from equilibrium than the hotter one, yet it reaches equilibrium faster.

Thermal relaxation, therefore, seems to be much more complex than originally thought, and our results only scratch at the surface. Relaxation in systems with multiple energy minima [73, 75, 251, 256], time-dependent potentials [49, 263, 264, 275–277], driven relaxation processes [278, 279], and in the presence of time-irreversible, detailed-balance violating dynamics [87, 121, 147, 280] remains poorly understood, and calls for a systematic analysis through the lens of thermal kinematics.

Author contribution

AL and RAR discussed the initial ideas of the project. MI carried out the experiments and analyzed all the data, supervised by RAR. MI, AL and RAR performed preliminary theoretical analysis. CD and AG developed the new theoretical framework and carried out the proofs. MI, AG and RAR prepared the manuscript, with contributions from all authors. All authors contributed to the discussion of the results and contributed to the overall development of the project.

Data and code availability statement

The experimental data presented in this work and the codes used to analyze it will be made publicly available in Zenodo after publication.

Acknowledgements

This work was supported by the projects EQC2018-004693-P, PID2020-116567GB-C22, PID2021-128970OA-I00 and PID2021-127427NB-I00 funded by MCIN/AEI/ 10.13039/501100011033/ FEDER, UE, and by FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades / Proyectos P18-FR-3583 and A-FQM-644-UGR20. M.I. acknowledges financial support from Ministerio de Universidades (Spain) and Universidad de Granada for his FPU grant FPU21/02569. C.D. acknowledges financial support from the Studienstiftung des Deutschen Volkes, A.G. acknowledges financial support from the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

5.4.7 Supplemental Material: Theoretical framework

In this section, we provide all the mathematical proofs referred to in the main text, as well as the general theoretical concepts to fully understand all the phenomena we present in our work.

Fokker-Planck equation for an optically-trapped Brownian particle

A colloidal particle trapped by a highly collimated laser beam is subjected to a three-dimensional quadratic potential, which in the x direction reads $U(x) = \frac{1}{2}\kappa(x - x_{\text{tr}})^2$. Here, x_{tr} denotes the trap center, in which we set the origin of coordinates in the following. The Brownian character of the particle under study makes it necessary to model the multiple collisions of the fluid molecules with it from a stochastic perspective. In the overdamped regime, the differential equation that rules its dynamics is of Langevin type,

$$dx_t = -(\kappa/\gamma)x_t dt + d\xi_t, \tag{5.4.7}$$

with friction constant γ given by the Stokes' law $\gamma = 6\pi r\eta$, where η is the viscosity of water and r the particle radius. The thermal noise $d\xi_t^e$ vanishes on average and obeys the fluctuation-dissipation theorem $\langle d\xi_t^e d\xi_{t'}^e \rangle = 2(k_{\text{B}}T/\gamma)\delta(t - t')dtdt'$.

A stochastic process solution of a Langevin equation can alternatively be studied from its associated probability density function $P(x, t)$ for each time instant $t \geq 0$ [6]. In this case, it verifies the Fokker-Planck equation, which for a Langevin dynamic such as the one described by Eq. (5.4.7) takes the form

$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(\frac{\kappa}{\gamma}xP(x,t)\right) + \frac{k_{\text{B}}T}{\gamma}\frac{\mathbb{P}^2 P(x,t)}{\partial x^2}. \tag{5.4.8}$$

$P(x, t)$ relaxes towards an equilibrium distribution of Boltzmann type, $P_{\text{eq}}(x) = \text{e}^{(F_T - U(x))/k_{\text{B}}T}$, where $F_T \equiv -k_{\text{B}}T \ln \int_{-\infty}^{\infty} e^{-U(x)/k_{\text{B}}T} dx$ is the equilibrium free energy at temperature T. In our case, the parabolic form of the potential gives $P_{\text{eq}}(x)$ the Gaussian character.

The probability density is specified (for a given initial condition) by a single parameter, time t. In our case, the initial distribution is Gaussian and is associated with a different environmental temperature T_0, $P(x, 0) = \text{e}^{\left(F_{T_0} - \frac{1}{2}\kappa x^2\right)/k_{\text{B}}T_0}$. The parameters that charac-

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

terize it, the average position $\langle x_t \rangle$ and variance $(\Delta x_t)^2 \equiv \langle (x_t - \langle x_t \rangle)^2 \rangle$, then depend on time and satisfy the pair of differential equations

$$\begin{cases} \frac{d}{dt}\langle x_t \rangle = -\frac{\kappa}{\gamma}\langle x_t \rangle, \\ \frac{d}{dt}(\Delta x_t)^2 = -2\frac{\kappa}{\gamma}(\Delta x_t)^2 + \frac{2k_\mathrm{B}T}{\gamma}, \end{cases} \tag{5.4.9}$$

with the corresponding initial conditions $\langle x_0 \rangle = 0$ and $(\Delta x_0)^2 = \langle x_0^2 \rangle = k_\mathrm{B}T_0/\kappa$. The average position therefore remains constantly zero, and its variance results in

$$\begin{aligned} (\Delta x_t)^2 = \langle x_t^2 \rangle &= \langle x_0^2 \rangle \mathrm{e}^{-2(\kappa/\gamma)t} + \int_0^t \frac{2k_\mathrm{B}T}{\gamma} \mathrm{e}^{-2(\kappa/\gamma)(t-s)}\, ds \\ &= \frac{k_\mathrm{B}T}{\kappa}\left[1 + (T_0/T - 1)\mathrm{e}^{-2(\kappa/\gamma)t}\right]. \end{aligned} \tag{5.4.10}$$

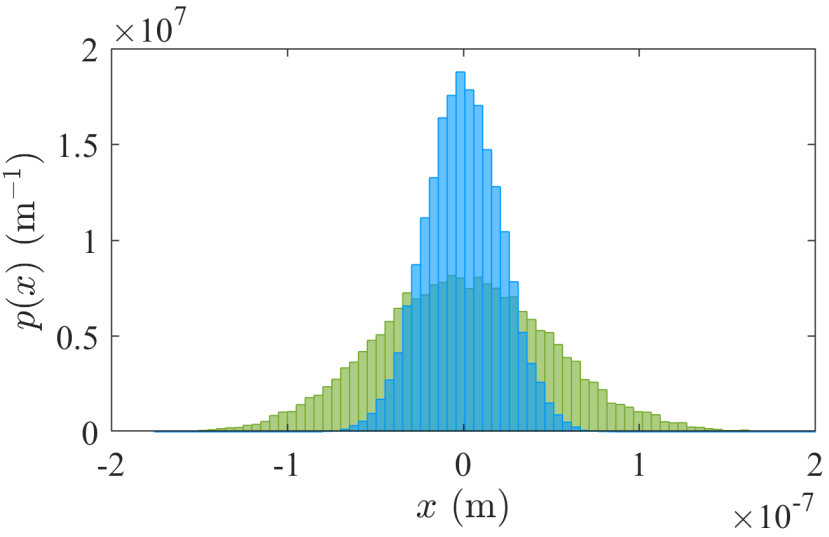

Figure 5.4.5: Histograms of the particle position for hot (green) and cold (blue) temperatures.

As the statistics is at all times a zero-mean Gaussian and its variance may be fully specified by some *instantaneous temperature* $\tilde{T} \equiv T/T_w$ we may, without loss of generality, consider a "quasi temperature" θ as parameterizing the distribution in all physical and abstract, information-geometric situations, i.e., $P(x, \theta)$. In the physical evolution of the system we have $\theta \equiv t$ which we will write simply as $P(x, t)$. In the general and all other parametrizations we will use $P(x, \theta)$.

The relative entropy between two normal, univariate probability distributions P and Q of zero mean and variances $(\Delta x)_P^2$ and $(\Delta x)_Q^2$ takes the form

$$D[P||Q] = \int dx P(x) \ln \frac{P(x)}{Q(x)} = \frac{1}{2}\left(\frac{(\Delta x)_P^2}{(\Delta x)_Q^2} - 1 - \ln \frac{(\Delta x)_P^2}{(\Delta x)_Q^2} \right) = \frac{1}{2}\left(\Lambda - 1 - \ln \Lambda\right). \tag{5.4.11}$$

When considering $D[P(x,t)||P_{\rm eq}(x)]$ then $\Lambda = \Lambda(t) = 1 + (T_0/T - 1)\mathrm{e}^{-2(\kappa/\gamma)t}$, while for $D[P(x,t)||P(x,0)]$ we have $\Lambda(t) = T/T_0 + (1 - T/T_0)\mathrm{e}^{-2(\kappa/\gamma)t}$. The excess of generalized free energy is thus defined as $\mathcal{D}_i^w(t) \equiv k_{\rm B} T_w D[P(x,t)||P_{\rm eq}(x)]$ in the forward protocol, where $T_i = T_{c,h}$ denotes the initial temperatures and T_w the target one.

Conversely, $\mathcal{D}_w^i(t) \equiv k_{\rm B} T_w D[P(x,t)||P(x,0)]$ in the backward protocol, which does not correspond to a free energy since T_w and $P(x,0)$ no longer refer to the target equilibrium.

Thermodynamically equidistant temperature quenches

A pair of temperatures T_h and T_c are said to be thermodynamically equidistant from T_w when the the initial excess free energies are equal, i.e., $\mathcal{D}_h^w(0) = \mathcal{D}_c^w(0)$. Under the assumptions on the equations of motion specified in the manuscript, the thermodynamically equidistant temperatures are given by T_h and $T_c(T_h)$ [73]

$$T_c(T_h) = -T_w W_0\left(-\frac{T_h}{T_w} \mathrm{e}^{-T_h/T_w} \right), \tag{5.4.12}$$

where $W_0(z)$ is the principal branch of the Lambert-W function [281].

Relaxation asymmetry in the backward protocol

The thermal relaxation asymmetry in the forward protocol is proven in [73]. We now prove the thermal asymmetry in the backward protocol.

Theorem 1: *For any pair TE temperatures T_c and T_h relative to T_w and for all times $t > 0$ we have*

$$\Delta D(t) \equiv \frac{1}{k_{\rm B} T_w}(\mathcal{D}_w^h(t) - \mathcal{D}_w^c(t)) > 0. \tag{5.4.13}$$

That is, during the backward protocol heating is faster than cooling.

Proof: We define

$$\phi \equiv \mathrm{e}^{-2(\kappa/\gamma)t} \in (0,1], \tag{5.4.14}$$

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

such that $\Lambda(t) = \Lambda_w^{c,h}(t)$ takes the expression previously introduced,

$$\Lambda_w^{c,h}(t) \equiv \tilde{T}_{c,h}(1-\phi) + \phi, \tag{5.4.15}$$

where here and in the following we denote $\tilde{T}_i \equiv T_i/T_w$. We now show Eq. (5.4.13), i.e., that for all $\phi \in (0,1]$ we have

$$2\Delta D(t) = \Lambda_w^h - \Lambda_w^c + \ln\left(\frac{\Lambda_w^c}{\Lambda_w^h}\right) \; 0. \tag{5.4.16}$$

First we rewrite this as

$$\begin{aligned} 2\Delta D(\phi) &= (\tilde{T}_h - \tilde{T}_c)(1-\phi) + \ln\left(\frac{\Lambda_w^c}{\Lambda_w^h}\right) \\ &= (\tilde{T}_h - \tilde{T}_c)(1-\phi) + \ln\left(\frac{\tilde{T}_c}{\tilde{T}_h}\right) + \ln\left(\frac{\Lambda_w^c/\tilde{T}_c}{\Lambda_w^h/\tilde{T}_h}\right). \end{aligned} \tag{5.4.17}$$

Note that for equidistant temperature quenches $\ln(\tilde{T}_c/\tilde{T}_h) = \tilde{T}_c - \tilde{T}_h$ which upon introducing further notation gives

$$\begin{aligned} C \equiv \frac{1-\tilde{T}_c}{\tilde{T}_c} > 0, \qquad H \equiv \frac{\tilde{T}_h - 1}{\tilde{T}_h} > 0, \quad H < 1, \\ 2\Delta D(\phi) = \ln\left(\frac{\Lambda_w^c/\tilde{T}_c}{\Lambda_w^h/\tilde{T}_h}\right) - \phi(\tilde{T}_h - \tilde{T}_c) = \ln\left(\frac{1+C\phi}{1-H\phi}\right) - \phi\left(\frac{1}{1-H} - \frac{1}{1+C}\right). \end{aligned} \tag{5.4.18}$$

Now let us take the derivative:

$$\begin{aligned} \frac{d}{d\phi} 2\Delta D(\phi) &= \frac{\left(\frac{C}{1-H\phi} + \frac{H(C\phi+1)}{(1-H\phi)^2}\right)(1-H\phi)}{C\phi+1} + \frac{1}{C+1} - \frac{1}{1-H} \\ &= \frac{(C+H)\,(\phi-1)\,(CH\phi + CH - C + H)}{(C+1)\,(H-1)\,(C\phi+1)\,(H\phi-1)}. \end{aligned} \tag{5.4.19}$$

Note that for $\phi \in (0,1)$ this expression will be 0 if and only if $CH\phi + CH - C + H = 0$. Since this is a linear equation, we conclude that *there is at most one* $\phi \in (0,1)$ *such that* $\frac{d}{d\phi}\Delta D(\phi) = 0$. This means that there is at *most one local extremum of* $\Delta D(\phi)$ *for* $\phi \in (0,1)$.

Note also that $\tilde{T}_h\tilde{T}_c < 1$, which follows from the fact that $\tilde{T}_c = -W_0(-\tilde{T}_h \mathrm{e}^{-\tilde{T}_h})$ with $W_0(z) \in [-1,0]$ defined for $z \in (-\mathrm{e}^{-1}, 0)$ is the principal branch of the Lambert-W function [73], and from Lemma 3.4 in [282].

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

We thus obtain

$$\begin{aligned}
2\Delta D(\phi \to 0) &\overset{\text{Eq. (5.4.18)}}{=} \ln(1) - 0(\tilde{T}_h - \tilde{T}_c) = 0, \\
2\Delta D(\phi \to 1) &\overset{\text{Eq. (5.4.16)}}{=} 1 - 1 + \ln(1) = 0 \\
\frac{d}{d\phi} 2\Delta D(\phi \to 0) &\overset{\text{Eq. (5.4.19)}}{=} H + C + \frac{1}{1+C} - \frac{1}{1-H} = \frac{\tilde{T}_h - 1}{\tilde{T}_h} + \frac{1-\tilde{T}_c}{\tilde{T}_c} + \tilde{T}_c - \tilde{T}_h \\
&= \frac{(\tilde{T}_h - \tilde{T}_c)(1 - \tilde{T}_c\tilde{T}_h)}{\tilde{T}_h\tilde{T}_c} > 0.
\end{aligned} \tag{5.4.20}$$

This implies that $\Delta D(\phi)$ for $\phi \in [0,1]$ starts at $\Delta D(0) = 0$, initially increases, and ends at $\Delta D(1) = 0$. The fact that there is at most one local extremum of $\Delta D(\phi)$ for $\phi \in (0,1)$ (as noted above from Eq. (5.4.19)), implies that $\Delta D(\phi)$ never crosses 0. Hence, $\Delta D(\phi) > 0$ for all $\phi \in [0,1]$, i.e., we have proven the relaxation asymmetry in the backward protocol, Eq. (5.4.13).

Information geometry and thermal kinematics

To provide intuition about, and to generally formulate the question of, the heating-cooling asymmetry in the three addressed situations, we now turn to the *thermal kinematics* of the relaxation process. For this we need to define distance, speed, and progression during the respective relaxation processes.

We now introduce the Fisher information $I(\theta)$ [266, 270] in our (physical) case defined as

$$I(\theta) \equiv \int_{-\infty}^{\infty} dx \frac{(\partial_\theta P(x,\theta))^2}{P(x,\theta)}. \tag{5.4.21}$$

The geometric interpretation of the Fisher information follows from defining a statistical line element ds as $ds^2 \equiv I(\theta)d\theta^2$, hence ds may be considered as a dimensionless distance between probability densities $P(x,\theta)$ and $P(x,\theta+d\theta)$. To understand this, we note that

$$\begin{aligned}
& D(P(x,\theta+d\theta)||P(x,\theta)) = \int dx P(x,\theta+d\theta) \ln \frac{P(x,\theta+d\theta)}{P(x,\theta)} \\
= & \int dx [P(x,\theta) + \partial_\theta P(x,\theta) d\theta + \mathcal{O}(d\theta^2)] \ln\left(1 + \frac{\partial_\theta P(x,\theta)}{P(x,\theta)} d\theta + \mathcal{O}(d\theta^2)\right) \\
= & \partial_\theta \int dx P(x,\theta) d\theta + \int dx \frac{(\partial_\theta P(x,\theta))^2}{P(x,\theta)} d\theta^2 + \mathcal{O}(d\theta^3) \\
= & 0 + I(\theta) d\theta^2 + \mathcal{O}(d\theta^3),
\end{aligned} \tag{5.4.22}$$

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

where in the third line we used $\ln(1+x) = x + \mathcal{O}(x^2)$ and in the fourth that the first term in the third line vanishes because $\int dx P(x,\theta) = 1$, $\forall\theta$. The *statistical length* may thus be defined as

$$\mathcal{L}(\theta_{\rm i},\theta_{\rm f}) = \int_{\theta_{\rm i}}^{\theta_{\rm f}} d|\theta|\sqrt{I(\theta)} \equiv \int_{\theta_{\rm i}}^{\theta_{\rm f}} d|\theta| v(\theta), \tag{5.4.23}$$

where in the last step we introduced the *local speed* $v(\theta)$. The expression $\mathcal{L}(\theta_{\rm i},\theta_{\rm f})$ measures the length of the path traced by the probability density under a change of the parameter from $\theta_{\rm i}$ to $\theta_{\rm f}$. Note that $\mathcal{L}(\theta_{\rm i},\theta_{\rm f})$ has, in contrast to the relative entropy that measures the thermodynamic displacement, all the properties of a path length: it satisfies the triangle inequality and is invariant under monotonic path reparametrizations.

In addition, to quantify the progression along the path $\{\theta_s\}_{\rm i\le s\le f}$ we further introduce the *degree of completion* as

$$\varphi(s) \equiv \frac{\mathcal{L}(\theta_{\rm i},\theta_s)}{\mathcal{L}(\theta_{\rm i},\theta_{\rm f})} \in [0,1]. \tag{5.4.24}$$

Thermal kinematics along thermodynamically equidistant quenches

In the case of the physical time evolution, an elementary calculation shows that

$$\begin{aligned} I^{\rm F}(t;\tilde{T}) &= \frac{1}{2}\left(\frac{\frac{d}{dt}\langle x^2(t)\rangle_{\rm F}}{\langle x^2(t)\rangle_{\rm F}}\right)^2 = \frac{2}{[1+{\rm e}^{2t}/(\tilde{T}-1)]^2}, \\ I^{\rm B}(t;\tilde{T}) &= \frac{1}{2}\left(\frac{\frac{d}{dt}\langle x^2(t)\rangle_{\rm B}}{\langle x^2(t)\rangle_{\rm B}}\right)^2 = \frac{2}{[1+{\rm e}^{2t}\tilde{T}/(1-\tilde{T})]^2}, \end{aligned} \tag{5.4.25}$$

where time is measured in units of γ/κ and the superscripts F and B denote the forward ($\tilde{T}\to 1$) and backward ($1\to\tilde{T}$) quench protocols. This implies the respective local speeds

$$v^{\rm F}(t;\tilde{T}) = \frac{\sqrt{2}}{|1+{\rm e}^{2t}/(\tilde{T}-1)|}, \quad v^{\rm B}(t;\tilde{T}) = \frac{\sqrt{2}}{|1+{\rm e}^{2t}\tilde{T}/(1-\tilde{T})|}, \tag{5.4.26}$$

and distances traversed until time t

$$\begin{aligned} \mathcal{L}^{\rm F}(t;\tilde{T}) &= {\rm sign}(\ln(\tilde{T}))\,\sqrt{2}[\ln\tilde{T} - \ln(1+(\tilde{T}-1){\rm e}^{-2t})], \\ \mathcal{L}^{\rm F}(\infty) &= \mathcal{L}(\tilde{T},1) = {\rm sign}(\ln(\tilde{T}))\,\sqrt{2}\ln\tilde{T}, \\ \mathcal{L}^{\rm B}(t;\tilde{T}) &= {\rm sign}(\ln(\tilde{T}))\,\sqrt{2}\ln({\rm e}^{-2t}+\tilde{T}(1-{\rm e}^{-2t})), \\ \mathcal{L}^{\rm B}(\infty) &= \mathcal{L}(1,\tilde{T}) = {\rm sign}(\ln(\tilde{T}))\,\sqrt{2}\ln\tilde{T}, \end{aligned} \tag{5.4.27}$$

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

finally yielding also the degrees of completion

$$\begin{aligned}\varphi^{\mathrm{F}}(t;\tilde{T}) &= 1-\frac{\ln(1+(\tilde{T}-1)\mathrm{e}^{-2t})}{\ln\tilde{T}},\\ \varphi^{\mathrm{B}}(t;\tilde{T}) &= \frac{\ln(\mathrm{e}^{-2t}+\tilde{T}(1-\mathrm{e}^{-2t}))}{\ln\tilde{T}}=1+\frac{\ln(1-(\tilde{T}-1)\mathrm{e}^{-2t}/\tilde{T})}{\ln\tilde{T}}.\end{aligned} \tag{5.4.28}$$

At this point we make the interesting observation:

Theorem 2: *(i) For any pair of thermodynamically equidistant temperatures $\tilde{T}^{+}>1>\tilde{T}^{-}$, the total distance $\mathcal{L}(0,\tilde{T})$ defined in Eq. (5.4.27) obeys $\mathcal{L}(1,\tilde{T}^{-})>\mathcal{L}(1,\tilde{T}^{+})$. Moreover, (ii) for any pair of thermodynamically equidistant temperatures $\tilde{T}^{+}>\tilde{T}^{-}$ we have that during the forward protocol*

$$\varphi^{\mathrm{F}}(t;\tilde{T}^{-})>\varphi^{\mathrm{F}}(t;\tilde{T}^{+})\quad\forall t\in(0,\infty), \tag{5.4.29}$$

and during the backward protocol

$$\varphi^{\mathrm{B}}(t;\tilde{T}^{+})>\varphi^{\mathrm{B}}(t;\tilde{T}^{-})\quad\forall t\in(0,\infty). \tag{5.4.30}$$

That is, the progress during both forward and backward processes is faster during heating for all times and thermodynamically equidistant quenches.

Remark: Theorem 2 implies that whereas the hot and cold system are initially thermodynamically equidistant from the equilibrium, the colder system is initially statistically farther away from equilibrium than the hotter system.

Moreover, the fact that the colder system is statistically farther from equilibrium than the corresponding thermodynamically equidistant hotter system shows that the statistical length of the path to equilibrium does not play an important role in the relaxation asymmetry. Namely, in the forward protocol heating is faster despite the longer path taken, whereas in the backward protocol heating is faster while the path taken is shorter. Conversely, the local speed $v(t)=\sqrt{I(t)}$, which initially is *always* faster during heating than during cooling, ensures that heating is, at all times, faster than cooling. This provides a kinematic explanation of the fact that the entropy production in the system during heating initially evolves as essentially "free diffusion" (and is not hampered by the drift imposed by the potential), whereas the heat flow during cooling is always (including at short times) opposed by the reduction of entropy in the system [73].

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Proof: Part (i): By the definition in Eq. (5.4.27) we have for $t = 0$

$$\mathcal{L}(1,\tilde{T}^-) - \mathcal{L}(1,\tilde{T}^+) = -\sqrt{2}\ln\tilde{T}^-\tilde{T}^+ > -\sqrt{2}\ln 1 > 0, \quad \forall \tilde{T}^+, \tilde{T}^-(\tilde{T}^+), \tag{5.4.31}$$

where in the second step we used $\tilde{T}^-\tilde{T}^+ \leq 1$ (equality holding only when $\tilde{T}^- = \tilde{T}^+ = 1$) (see Lemma 3.4. in [281] and recall that $\tilde{T}^- = -W_0(-\tilde{T}^+ \mathrm{e}^{-\tilde{T}^+})$ in terms of the principal branch of the Lambert-W function $W_0(z)$).

Part (ii): We set $\Delta\varphi^{\mathrm{F}}(t) \equiv \varphi^{\mathrm{F}}(t;\tilde{T}^-) - \varphi^{\mathrm{F}}(t;\tilde{T}^+)$ and $\Delta\varphi^{\mathrm{B}}(t) \equiv \varphi^{\mathrm{B}}(t;\tilde{T}^+) - \varphi^{\mathrm{B}}(t;\tilde{T}^-)$. We have trivially that $\Delta\varphi^{\mathrm{F,B}}(0) = \Delta\varphi^{\mathrm{F,B}}(\infty) = 0$. Let $\delta^\pm \equiv \tilde{T}^\pm - 1$ as well as $0 \leq x \equiv \mathrm{e}^{-2t} \leq 1$. For convenience and without any loss of generality, we consider $\varphi^{\mathrm{F,B}}$ as a function of x rather than t since $x(t)$ is monotone in t with $x \to 0 \iff t \to \infty$ and $x \to 1 \iff t \to 0$. We first prove the statement for the forward and then for the backward protocol. Since $\Delta\varphi^{\mathrm{F,B}}(x)$ is continuously differentiable, we only need to show that $\Delta\varphi^{\mathrm{F,B}}(t)$ has a single maximum and that it is positive somewhere on $t \in (0,\infty)$ (i.e., on $x \in (0,1)$).

Forward protocol.—We first carry out the proof for the forward protocol. First, we show that $\Delta\varphi^{\mathrm{F}}(x) > 0$ for $x = \varepsilon \ll 1$. Expanding around $x = 0$, i.e., $\Delta\varphi^{\mathrm{F}}(\varepsilon) = K^\dagger(\delta^+,\delta^-)\varepsilon + \mathcal{O}(\varepsilon^2)$, we find

$$\begin{aligned} K^\dagger(\delta^+,\delta^-) &= \left(\frac{\delta^+}{\ln(1+\delta^+)} - \frac{\delta^-}{\ln(1+\delta^-)}\right) \\ &= \frac{\delta^-\ln(1+\delta^+) - \delta^+\ln(1+\delta^-)}{\ln(1+\delta^+)[-\ln(1+\delta^-)]} \\ &> [-\ln(1+\delta^-)]\frac{\delta^+ - \ln(1+\delta^+)}{\ln(1+\delta^+)[-\ln(1+\delta^-)]} > 0, \end{aligned} \tag{5.4.32}$$

where in the second line we made the denominator positive and in the third line we used twice the concavity of the logarithm implying $\ln(1+x) \leq x$, $\forall x > -1$, i.e., first $\delta^- > \ln(1+\delta^-)$ and in the last step $\delta^+ > \ln(1+\delta^+)$, which completely proves that $\Delta\varphi^{\mathrm{F}}(\varepsilon) > 0$.

The final step is to show that $\Delta\varphi^{\mathrm{F}}(x)$ has a single stationary point on $x \in (0,1)$. We do this by direct computation, i.e., the only solution of $\frac{d}{dx}\Delta\varphi^{\mathrm{F}}(x) = 0$ on the interval $x \in (0,1)$ is

$$x^*_{\mathrm{F}} = \frac{1-C}{C\delta^+ - \delta^-} \quad \text{where} \quad C \equiv \frac{\delta^-}{\delta^+}\frac{\ln(1+\delta^+)}{\ln(1+\delta^-)}, \tag{5.4.33}$$

which must be a maximum as $\Delta\varphi^{\mathrm{F}}(\varepsilon) > 0$ and $\Delta\varphi^{\mathrm{F}}(1-\varepsilon) > 0$ and because $\Delta\varphi^{\mathrm{F}}(x)$ is continuously differentiable on $x \in (0,1)$. This proves the theorem for the forward protocol.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Backward protocol.—Proceeding as above, we show that $\lim_{\varepsilon\to 0}\Delta\varphi^{\mathrm{B}}(1-\varepsilon)>0$. First, by Taylor expanding we find from the second line in Eq. (5.4.28)

$$\Delta\varphi^{\mathrm{B}}(1-\varepsilon)=K^{\dagger}(\delta^{+},\delta^{-})\varepsilon+\mathcal{O}(\varepsilon^{2}), \tag{5.4.34}$$

which already proves $\lim_{\varepsilon\to 0}\Delta\varphi^{\mathrm{B}}(1-\varepsilon)>0$ because we showed in Eq. (5.4.32) that $K^{\dagger}(\delta^{+},\delta^{-})>0$.

The final step is to show that $\Delta\varphi^{\mathrm{B}}(x)$ has a single stationary point on $x\in(0,1)$. We do this by showing that the only solution of $\frac{d}{dx}\Delta\varphi^{\mathrm{B}}(x)=0$ on the interval $x\in(0,1)$ is

$$x^{*}_{\mathrm{B}} \;=\; \frac{C(1+\delta^{+})-(1+\delta^{-})}{C\delta^{+}-\delta^{-}}, \tag{5.4.35}$$

with C defined in Eq. (5.4.33). x^{*}_{B} must be a maximum as $\Delta\varphi^{\mathrm{B}}(\varepsilon)>0$ and $\Delta\varphi^{\mathrm{B}}(1-\varepsilon)>0$ and because $\Delta\varphi^{\mathrm{B}}(x)$ is continuously differentiable on $x\in(0,1)$. This proves the theorem also for the backward protocol.

Thermal kinematics in heating and cooling processes between a single pair of temperatures

Finally, we prove that between any pair of temperatures $\tilde{T}$ and 1 heating is always faster. In particular, we prove

Theorem 3: *Given any pair of temperatures $\tilde{T}\neq 1$ and 1 (denote $\tilde{T}>1$ by $\tilde{T}^{+}$ and $\tilde{T}<1$ by $\tilde{T}^{-}$), the heating process between the two temperatures is always faster than the corresponding cooling processes. Precisely,*

$$\begin{aligned}
\varphi^{\mathrm{F}}(t;\tilde{T}^{-}) &> \varphi^{\mathrm{B}}(t;\tilde{T}^{-}) \quad \forall t\in(0,\infty),\\
\varphi^{\mathrm{F}}(t;\tilde{T}^{+}) &< \varphi^{\mathrm{B}}(t;\tilde{T}^{+}) \quad \forall t\in(0,\infty).
\end{aligned} \tag{5.4.36}$$

Proof: Adopting the previous notation, we have

$$\begin{aligned}
\varphi^{\mathrm{F}}(x;\tilde{T}^{\pm})-\varphi^{\mathrm{B}}(x;\tilde{T}^{\pm}) &\equiv \Delta_{\pm}(x)=\frac{\ln(1+\delta^{\pm})-\ln(1+\delta^{\pm}x)-\ln(1+\delta^{\pm}-\delta^{\pm}x)}{\ln(1+\delta^{\pm})}\\
&= -\frac{1}{\ln(1+\delta^{\pm})}\ln\left(\frac{(1+\delta^{\pm}-\delta^{\pm}x)(1+\delta^{\pm}x)}{1+\delta^{\pm}}\right)\\
&\equiv \frac{-\ln\Xi^{\pm}(x)}{\ln(1+\delta^{\pm})},
\end{aligned} \tag{5.4.37}$$

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

where in the last line we defined $\Xi^{\pm}(x)$. We further have

$$\Xi^{\pm}(x) = 1 + \frac{x(1-x)(\delta^{\pm})^2}{1+\delta^{\pm}} > 1 \quad \forall x \in (0,1), \tag{5.4.38}$$

and thus $-\ln \Xi^{\pm}(x) < 0$. Moreover, we have $\ln(1+\delta^-) < 0$ while $\ln(1+\delta^+) > 0$ and therefore $\Delta_-(x) > 0$ while $\Delta_+(x) < 0$, which proves the theorem.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Heating and cooling near equilibrium

Near equilibrium, that is, for $T_+ = T_w(1+\varepsilon)$ with $0 < \varepsilon \ll 1$ we have, using Eq. (5.4.12)

$$\begin{aligned} T_-((1+\varepsilon)T_w) &= -W_0(-(1+\varepsilon)\mathrm{e}^{-(1+\varepsilon)})T_w \\ &= -W_0(-\mathrm{e}^{-1} + \mathrm{e}^{-1}\varepsilon^2/2 + \mathcal{O}(\varepsilon^3))T_w \\ &= -\left(-1 + \varepsilon + \varepsilon^2/3 + \mathcal{O}(\varepsilon^3)\right) T_w \\ &= (1 - \varepsilon[1-\varepsilon/3])T_w + \mathcal{O}(\varepsilon^3), \end{aligned} \tag{5.4.39}$$

as stated in Subsec. 5.4.5. In passing from the second to the third line, we used the Taylor series of $W_0(z)$ around the branch point at $-\mathrm{e}^{-1}$ (see Eq. (4.22) in [281]), i.e., $W_0(-\mathrm{e}^{-1} + z) = -1 + \sqrt{2\mathrm{e}z} + \mathcal{O}(z)$. Therefore, close to equilibrium thermodynamically equidistant temperatures are to linear order in the deviation ε also equidistant.

For the Langevin dynamics Eqs. (5.4.7) we can write the mean energy as $\langle U(x_t)\rangle = \kappa(\Delta x_t)^2/2$ with variance $(\Delta x_t)^2$ as in Eq. (5.4.10). Taking the time-derivative, we obtain $\partial_t \langle U(x_t)\rangle_i^w|_{t=0} = -k_\mathrm{B}T_w\kappa(T_i/T_w - 1)/\gamma$. For this Gaussian process, the time derivative of the system entropy $S(t) = -k_\mathrm{B}\int P_i^w(x,t)\ln P_i^w(x,t)dx$ is calculated as $\partial_t S_i^w(t)|_{t=0} = -k_\mathrm{B}\kappa(1 - T_w/T_i)\gamma$. Adding the two contributions, we obtain for the change in excess free energy that

$$\partial_t \mathcal{D}_i^w(t)|_{t=0} = \partial_t \left[\langle U(x_t)\rangle_i^w - T_w S_i^w(t)\right]|_{t=0} = k_\mathrm{B}T_w\frac{\kappa}{\gamma}\left[2 - \frac{T_i}{T_w} - \frac{T_w}{T_i}\right]. \tag{5.4.40}$$

Corollary 4: *To linear order in ε heating and cooling between thermodynamically equidistant as well as between any fixed pair of temperatures are symmetric. Precisely,*

$$\begin{aligned} \varphi_w^\mathrm{F}(t; 1-\varepsilon[1-\varepsilon/3]) &= \varphi^\mathrm{F}(t; 1+\varepsilon) + \mathcal{O}(\varepsilon), \\ \varphi^\mathrm{B}(t; 1+\varepsilon) &= \varphi_w^\mathrm{B}(t; 1-\varepsilon[1-\varepsilon/3]) + \mathcal{O}(\varepsilon), \end{aligned} \tag{5.4.41}$$

as well as

$$\begin{aligned}\varphi_w^{\mathrm{B}}(t;1+\varepsilon) &= \varphi_w^{\mathrm{F}}(t;1+\varepsilon)+\mathcal{O}(\varepsilon),\\ \varphi_w^{\mathrm{F}}(t;1-\varepsilon[1-\varepsilon/3]) &= \varphi_w^{\mathrm{B}}(t;1-\varepsilon[1-\varepsilon/3])+\mathcal{O}(\varepsilon).\end{aligned} \tag{5.4.42}$$

Proof: Inserting $T_+ = T_w(1+\varepsilon)$ and $T_- = T_w(1-\varepsilon[1-\varepsilon/3])$ into Eqs. (5.4.28) and Taylor expanding around $\varepsilon = 0$, we find

$$\begin{aligned}\varphi^{\mathrm{F}}(t;1+\varepsilon) &= 1-\mathrm{e}^{-2(\kappa/\gamma)t}-\varepsilon\,\mathrm{e}^{-3(\kappa/\gamma)t}\sinh((\kappa/\gamma)t)+\mathcal{O}(\varepsilon^2),\\ \varphi^{\mathrm{F}}(t;1-\varepsilon[1-\varepsilon/3]) &= 1-\mathrm{e}^{-2(\kappa/\gamma)t}+\varepsilon\,\mathrm{e}^{-3(\kappa/\gamma)t}\sinh((\kappa/\gamma)t)+\mathcal{O}(\varepsilon^2),\\ \varphi^{\mathrm{B}}(t;1+\varepsilon) &= 1-\mathrm{e}^{-2(\kappa/\gamma)t}+\varepsilon\,\mathrm{e}^{-3(\kappa/\gamma)t}\sinh((\kappa/\gamma)t)+\mathcal{O}(\varepsilon^2),\\ \varphi^{\mathrm{B}}(t;1-\varepsilon[1-\varepsilon/3]) &= 1-\mathrm{e}^{-2(\kappa/\gamma)t}-\varepsilon\,\mathrm{e}^{-3(\kappa/\gamma)t}\sinh((\kappa/\gamma)t)+\mathcal{O}(\varepsilon^2),\end{aligned} \tag{5.4.43}$$

which proves the corollary.

5.4.8 Supplemental Material: Further results

We include in this section a set of additional results that support the main findings presented in the core part of our work. They have been obtained from three distinct experimental settings, which cover different stiffness coefficients (and, subsequently, different characteristic times of the optical trap) and bath temperatures, see Figs. 5.4.6-5.4.9.

The first two series correspond to situations that are remarkably far from equilibrium, where the asymmetric character is evident. The initial conditions associated with the third one, however, are closer to equilibrium, and the quasi-symmetry involving heating and cooling is also confirmed.

Experimental check of the Pythagorean inequality

As we mentioned in the main part and had been proven in Ref. [124], the relative entropy satisfies what is called the Pythagorean inequality when it tracks a pure thermal relaxation process, such as the ones covered throughout our work. More specifically, if a trapped, Brownian particle suffers an instantaneous temperature quench that links an initial state $P(x,0)$ with a target state $P(x,\infty)$ (reached when $t\to\infty$), then

$$D[P(x,0)||P(x,\infty)] \geq D[P(x,0)||P(x,t)] + D[P(x,t)||P(x,\infty)]. \tag{5.4.44}$$

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

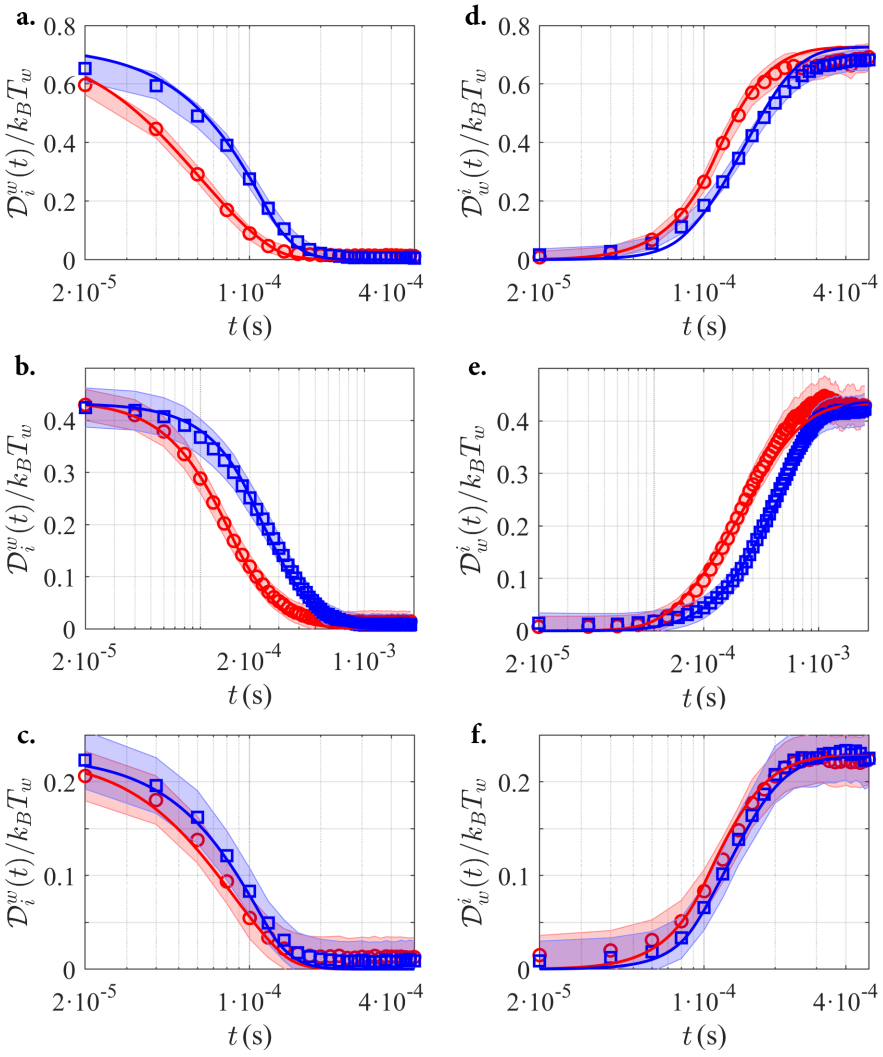

Figure 5.4.6: Temporal evolution of the generalized excess free energy as a function of time during heating and cooling at thermodynamically equidistant conditions. Left column corresponds to the forward protocol, while the right column corresponds to its backward counterpart. Red circles stand for heating and blue squares stand for cooling. Solid lines are theoretical predictions without fitting parameters. The confidence regions have been estimated by quadratic uncertainty propagation from the standard deviation of the experimental histograms. *Upper row:* time evolution for a characteristic time $\tau = \gamma/\kappa = 0.099(1)$ ms, $T_c/T_w = 0.09(1)$, $T_h/T_w = 3.78(1)$. *Middle row:* time evolution for a characteristic time $\tau = \gamma/\kappa = 0.539(2)$ ms, $T_c/T_w = 0.18(1)$, $T_h/T_w = 2.95(1)$. *Lower row:* time evolution for a characteristic time $\tau = \gamma/\kappa = 0.099(1)$ ms, $T_c/T_w = 0.32(1)$, $T_h/T_w = 2.28(1)$.

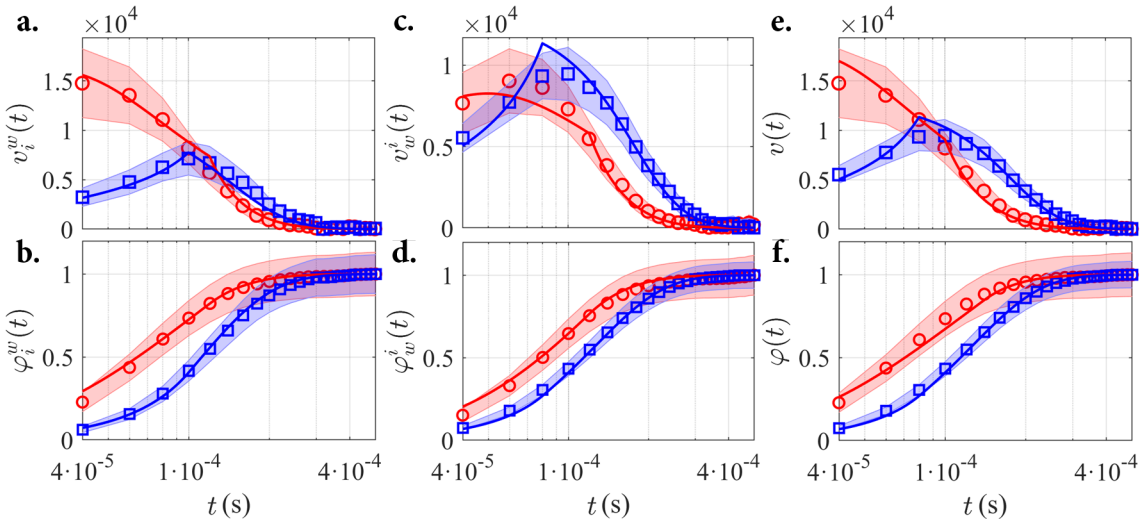

Figure 5.4.7: Temporal evolution of the statistical velocity (a., c., e.) and the degree of completion (b., d., f.) along the three considered situations. All graphics correspond to a characteristic time $\tau = \gamma/\kappa = 0.099(1)$ ms, $T_c = 302(2)$ K, $T_w = 3192(4)$ K, $T_h = 12080(14)$ K. Red circles stand for heating, while blue squares correspond to cooling. Solid lines are theoretical predictions without fitting parameters. The confidence regions have been estimated by quadratic uncertainty propagation from the standard deviation of the experimental histograms. **a., b.** Temporal evolution along the thermodynamically equidistant situation, forward protocol. **c., d.** Temporal evolution along the thermodynamically equidistant situation, backward protocol. **e., f.** Temporal evolution during heating and cooling between temperatures T_c and T_w.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

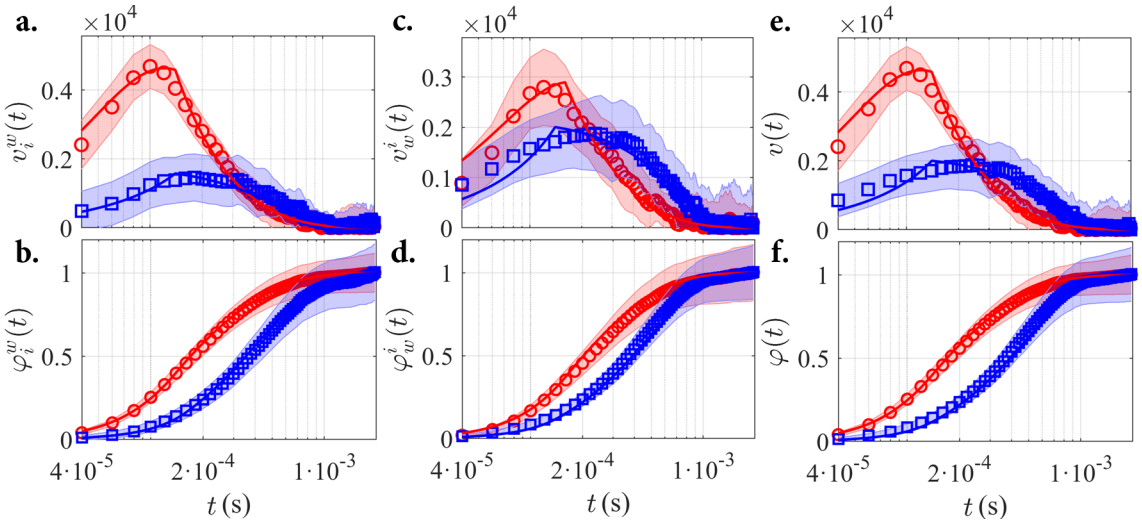

Figure 5.4.8: Temporal evolution of the statistical velocity (a., c., e.) and the degree of completion (b., d., f.) along the three considered situations. All graphics correspond to a characteristic time $\tau = \gamma/\kappa = 0.539(2)$ ms, $T_c = 302(2)$ K, $T_w = 1639(1)$ K, $T_h = 4828(6)$ K. Red circles stand for heating, while blue squares correspond to cooling. Solid lines are theoretical predictions without fitting parameters. The confidence regions have been estimated by quadratic uncertainty propagation from the standard deviation of the experimental histograms. **a., b.** Temporal evolution along the thermodynamically equidistant situation, forward protocol. **c., d.** Temporal evolution along the thermodynamically equidistant situation, backward protocol. **e., f.** Temporal evolution during heating and cooling between temperatures T_c and T_w.

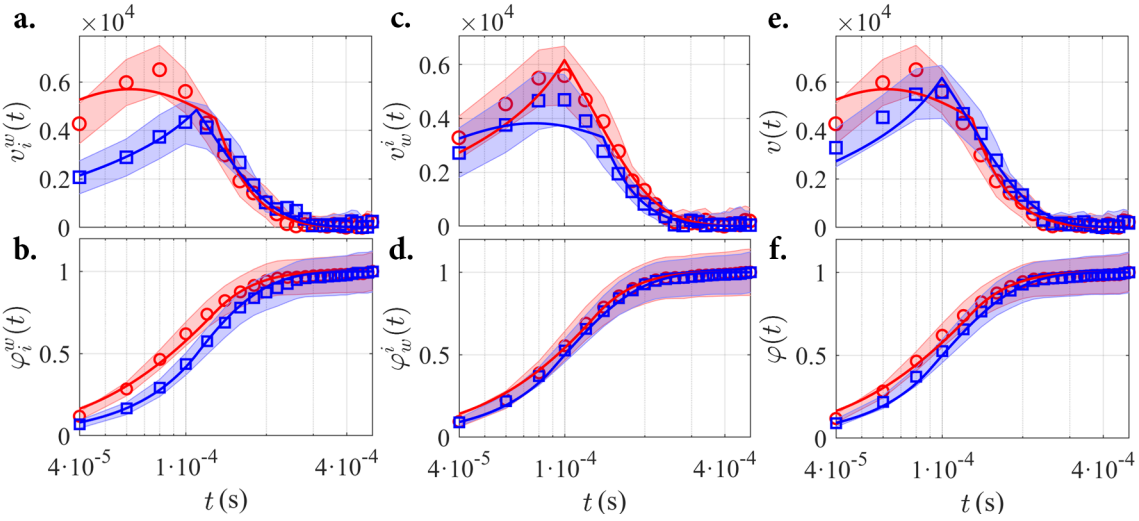

Figure 5.4.9: Temporal evolution of the statistical velocity (a., c., e.) and the degree of completion (b., d., f.) along the three considered situations. All graphics correspond to a characteristic time $\tau = \gamma/\kappa = 0.099(1)$ ms, $T_c = 302(2)$ K, $T_w = 935(5)$ K, $T_h = 2131(7)$ K. Red circles stand for heating, while blue squares correspond to cooling. Solid lines are theoretical predictions without fitting parameters. The confidence regions have been estimated by quadratic uncertainty propagation from the standard deviation of the experimental histograms. **a., b.** Temporal evolution along the thermodynamically equidistant situation, forward protocol. **c., d.** Temporal evolution along the thermodynamically equidistant situation, backward protocol. **e., f.** Temporal evolution during heating and cooling between temperatures T_c and T_w.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

As it is stated in the reference cited above, it establishes a tighter bound to the entropy production than the one imposed by the second law of thermodynamics, since $s_{[0,t]} \equiv D[P(x,0)||P(x,\infty)] - D[P(x,t)||P(x,\infty)]$ is the entropy production until time $t \geq 0$.

Furthermore, it trivially leads to a bound for the relaxation time: $t \geq D[P(x,0)||P(x,t)]/\bar{s}$, where $\bar{s} \equiv s_{[0,t]}/t$ is the average entropy production. Put in this way, the previous inequality establishes that large entropy production rates are necessary for shorter relaxation times, which cannot be arbitrarily small.

We address here the experimental check of the Pythagorean inequality, Eq. (5.4.44), by studying two particular thermal relaxation processes, one corresponding to a heating process and the other corresponding to a cooling one. As a matter of closure, we choose two initial conditions for which their corresponding states are thermodynamically equidistant with respect to the target, but the validity of the inequality (5.4.44) does not depend at all on this fact. Fig. 5.4.10 shows the values of $D[P(x,0)||P(x,t)]$, $D[P(x,t)||P(x,\infty)]$ and the sum of both magnitudes. In order for the inequality to be fulfilled, the latter must be below the value $D[P(x,0)||P(x,\infty)]$ for all times, which has to coincide with the initial value of $D[P(x,t)||P(x,\infty)]$ and the final value of $D[P(x,0)||P(x,t)]$.

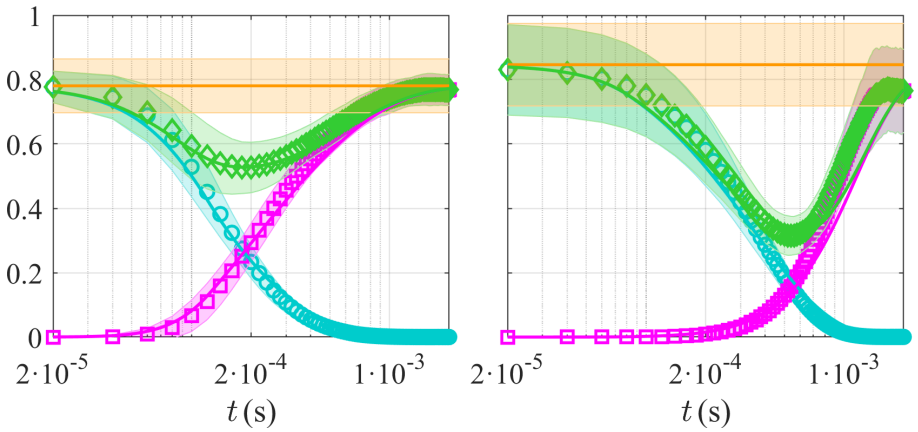

Figure 5.4.10: Experimental check of the Pythagorean inequality. Each plot shows the quantities $D[P(x,t)||P(x,\infty)]$ (cyan, circles), $D[P(x,0)||P(x,t)]$ (magenta, squares) and the sum of both (green, diamonds). In order for the inequality to be fulfilled, the latter must be below the orange, horizontal line $D[P(x,0)||P(x,\infty)]$. Solid lines are theoretical predictions without fitting parameters. Both graphics correspond to a characteristic time $\gamma/\kappa = 0.96(3)$ ms and relative temperatures $T_0/T_f = 0.0841$ (left panel, heating) and $T_0/T_f = 4.1020$ (right panel, cooling).

In view of the previous figure, it is noticeable that the magnitude $D[P(x,0)||P(x,t)] + D[P(x,t)||P(x,\infty)]$ reaches a minimum value which is higher (and closer to the limit value $D[P(x,0)||P(x,\infty)]$) during the heating process than during its cooling counterpart when

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

thermodynamically equidistant quenches are considered. We have checked that this is actually the general case (though we do not include a mathematical proof here as it goes far away from the main objectives of our work), a fact that may support the already proven fact that heating is always faster than cooling.

5.4.9 Materials and methods

Experimental set-up

Silica microspheres of diameter $d = 1$ μm are dispersed in deionized and filtered water at a concentration of a few spheres per ml. The dispersion is injected into a custom-made electrophoretic fluid chamber where two parallel, gold-coated wires act as a pair of parallel electrodes. A white, Gaussian noise signal composed of a sample of 500,000 independent Gaussian values of zero mean and unit variance is generated by an arbitrary waveform generator (RSPro, RSDG2082X, 80MHz bandwidth), thus producing a voltage output signal $V(t)$ that is fed into the pair of electrodes. The repetition frequency of the cited signal is chosen to minimize correlations and the time of noise persistence is just 0.04ms, which is much shorter than any characteristic time of the studied processes (these cover a range between 0.1 and 0.5 ms).

Our experiment uses an optical tweezers setup, which consists of a commercial device from JPK-Bruker (NanoTracker 2). The optical potential is generated by a near-infrared laser with a wavelength of 1064 nm, which is used for both trapping and detection. The forward scattered light is analyzed using a quadrant position detector (QPD) at a rate of 50 kHz. The trap stiffness depends linearly on the laser power and can be adjusted by controlling the latter with the software provided by the manufacturer.

The noisy signal is also modulated by a square signal with a period of 10 ms (see Fig. 5.4.11). This period is chosen to be a multiple of the sampling time of 0.02 ms for accurate data processing. The heights of the square signal are set to ensure equidistant quenches between the heating and cooling processes. As only the ratio between initial and final temperatures is important, it is not necessary to set the high level of the heating process equal to the low level of the cooling process as long as the initial values of the relative entropy coincide for both heating and cooling.

In Fig. 5.4.11, we show an example of the evolution in time of the variance of the modulated noisy signal for both heating and cooling processes. There is a "ramp" of finite slope and length $t_r = 0.10$ ms linking its initial and final values, instead of an instantaneous change. However, this does not prevent the studied phenomena from occurring, as our results seem

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

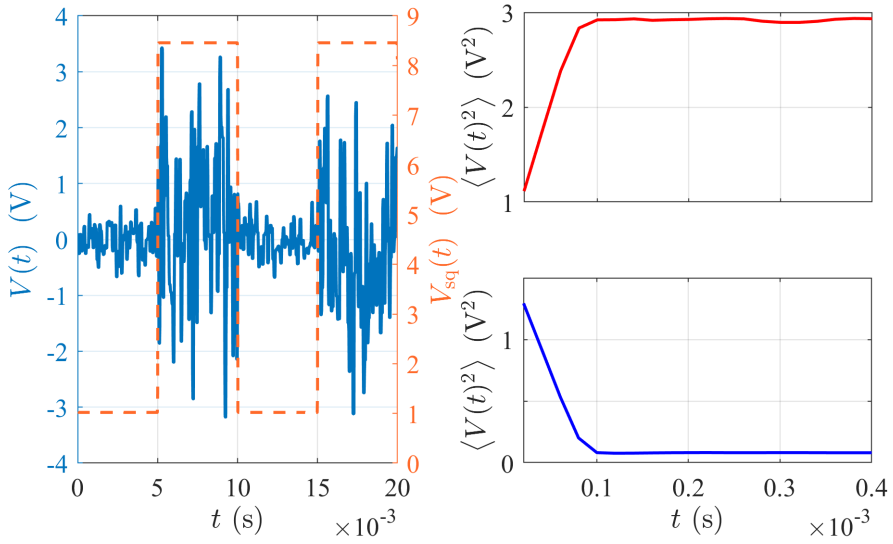

Figure 5.4.11: White, Gaussian noise modulated by a square signal. *Left panel:* modulated voltage output signal $V(t)$ (solid line, left axis) and square signal $V_{\text{sq}}(t)$ (dashed line, right axis) as a function of time. *Right panel:* evolution of the modulated signal variance $\langle V(t)^2 \rangle$ as a function of time for heating (upper panel) and cooling (lower panel).

to corroborate despite the fact that the ramp time is, in some situations, of the order of the characteristic time of the considered relaxation processes.

Effective heating by means of an external random force

The applied noisy electric field $d\xi_t^{\text{ext}}$ mimics an additional, independent source of thermal fluctuations in the particle dynamics, due to the presence of electric charges all over its surface [283]. When it is present, the position x_t along the considered direction obeys the Langevin equation $\gamma dx_t = -\kappa x_t dt + d\xi_t^{\text{th}} + d\xi_t^{\text{ext}}$, where $d\xi_t^{\text{th}}$ is of the same type as the one introduced in Eq. (5.4.7). If the spectrum of the external noise is also white and of amplitude σ^2, then the particle is subjected to an effective white noise $d\xi_t^{\text{eff}}$ of zero mean and autocorrelation $\langle d\xi_t^{\text{eff}} d\xi_{t'}^{\text{eff}} \rangle = 2\gamma k_{\text{B}} T_{\text{eff}} \delta(t - t') dt dt'$, where

$$T_{\text{eff}} = T + \frac{\sigma^2}{2k_{\text{B}}\gamma}. \tag{5.4.45}$$

Hence the position of the particle along the selected axis fluctuates corresponding to an effective temperature T_{eff} which is always higher than the temperature of the thermal bath and without affecting dissipation at all.

In our case, $\sigma^2 = q^2 V_0^2 / d^2$, with V_0^2 the variance of the noisy voltage output. The previous expression shows that the effective temperature increases linearly with V_0^2, whose slope

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

can be estimated from several realizations of the particle position for different amplitudes. Fig. 5.4.12 shows an example of two power spectral densities (PSDs) of the position of the sphere, with and without the electric field, as well as an example of a calibration curve T vs. V_0^2.

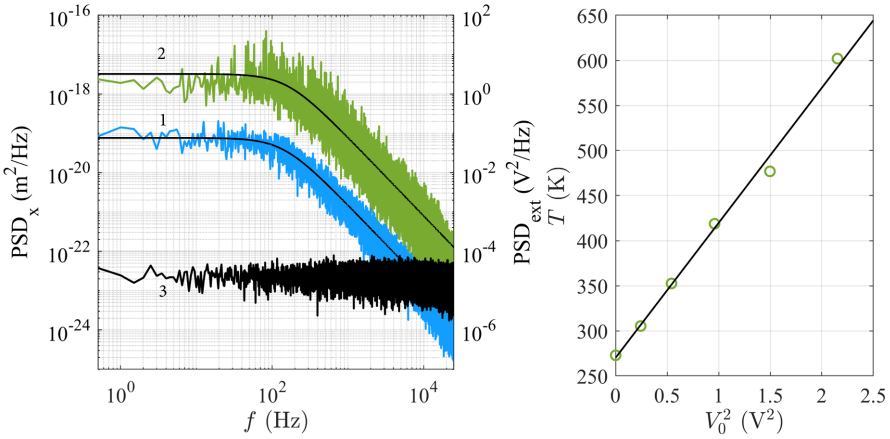

Figure 5.4.12: Effective heating by means of a noisy electric field. *Left panel:* PSD of the position of a trapped microparticle without (1) and with (2) the noisy electric field. Solid, black lines correspond to the fitting curves. The PSD of the output voltage signal is also included (3). *Right panel:* effective temperatures as a function of the squared voltage amplitude. The solid line corresponds to the linear fitting, according to Eq. (5.4.45).

Data analysis

We measure a collection of $M = 24,000$ trajectories, each one composed by values of the position at $N = 250$ different time instants, $\boldsymbol{x}^{(k)} = (x_1^{(k)}, \ldots, x_N^{(k)})$, $k = 1, \ldots, M$. From the M trajectories, the N histograms (one for each time instant) are calculated and properly normalized, $\{p(x, t_j)\}_{j=1}^{N}$. A total of $N_b = 80$ equidistantly distributed bins and separated by a fixed length $\Delta x = (\max_{jk} x_j^k - \min_{jk} x_j^k)/N_b$ have been chosen.

The values of the relative entropy calculated throughout our work are computed directly from the definition,

$$D[P||Q] = \int dx\, P \log \frac{P}{Q}. \tag{5.4.46}$$

The previous integral is discretized as usual from the histograms:

$$D[P||Q] \approx \Delta x \sum_{k=1}^{N_b} P(x_k) \log \frac{P(x_k)}{Q(x_k)}. \tag{5.4.47}$$

The convention $0 \log 0 = 0$ is properly employed. Finally, in order to smooth out the results, the presented series are estimated from an average of ten different temporal series.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

The number of counts n each bin of the histogram contains are assumed to follow a Poissonian distribution, since the spatial interval Δx associated with each bin is small and the total number N of measured values of the position is high enough. Consequently, the uncertainty associated with each of the heights can be fairly well approximated by $\sqrt{n}$, as it corresponds to the standard deviation of the cited probability distribution. Subsequent uncertainties are estimated by quadratic propagation.

With regard to Fisher information, we estimate it from its definition, which directly involves the relative entropy: $D[P(x,t+\Delta t)||P(x,t)] = I(t)\Delta t^2 + \mathcal{O}(\Delta t^3)$. An appropriate choice of Δt must be made to reduce numerical noise and to smooth the resulting curves. In our case, we choose $\Delta t = 0.04$ms. We further take an average of ten different temporal series.

Finally, to calculate the statistical length $\mathcal{L}(t)$, we perform a first-order numerical algorithm to calculate the corresponding integrals. If $t = n\Delta t,\ n \geq 0$ is a multiple of the sampling time and $v(t) \equiv \sqrt{I(t)}$ denotes the instantaneous velocity introduced in the main part,

$$\mathcal{L}(t) \approx \Delta t \sum_{i=0}^{n-1} v(i\Delta t). \tag{5.4.48}$$

Transient evolution during a ramp thermal protocol

As we showed earlier, the thermal quench our particle underwent was not completely instantaneous. The noise variance (as shown in Fig. 5.4.11) exhibited a ramp connecting its initial and final values. Although the time interval for this ramp is expected to be small when compared to the relaxation time $\tau = \gamma/\kappa$, this may not be the case for higher values of the cut-off frequency.

We denote the time of the ramp as t_r, and the parameter $\delta \equiv t_r/\tau$ determines the strength of the driving protocol into the thermal relaxation processes being studied (see Fig. 5.4.13). In this section, we will show that from the time instant in which the bath reaches its target temperature, the evolution law for $\langle x_t^2 \rangle$ coincides in shape with that corresponding to an instantaneous quench.

To this end, let us consider a non-instantaneous quench protocol determined by the bath temperature profile

$$T(t) = \begin{cases} T_0 + \lambda t,\ 0 \leq t \leq t_r; \\ T_f,\ t \geq t_r, \end{cases} \tag{5.4.49}$$

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

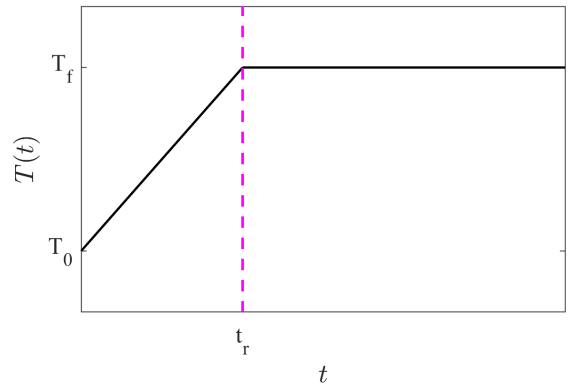

Figure 5.4.13: Scheme of the thermal protocol specified in equation (5.4.49)

where $\lambda \equiv (T_f - T_0)/t_r$ and $T_{0,f}$ the initial and final temperatures, respectively. If the initial distribution follows a Gaussian profile

$$P(x,0) = \sqrt{\frac{\kappa}{2\pi k_{\mathrm{B}} T_0}} e^{-\kappa x^2/2k_{\mathrm{B}}T_0}, \tag{5.4.50}$$

the solution of the Fokker-Planck equation

$$\partial_t P = \frac{\kappa}{\gamma}\partial_x(xP) + \frac{k_{\mathrm{B}}T(t)}{\gamma}\partial_x^2 P \tag{5.4.51}$$

also follows a Gaussian profile of zero mean and variance

$$\langle x_t^2 \rangle = \frac{k_{\mathrm{B}}T_0}{\kappa} e^{-2(\kappa/\gamma)t} + 2\int_0^t ds\, \frac{k_{\mathrm{B}}T(t)}{\gamma} e^{-2(\kappa/\gamma)(t-s)}. \tag{5.4.52}$$

The preceding integral takes a different expression whether the considered time instant t is greater or lower than t_r. On the one hand, when $0 \leq t \leq t_r$ it takes the form

$$\langle x_t^2 \rangle = \frac{k_{\mathrm{B}}T_f}{2\delta\kappa}\left[\left(1-\tilde{T}\right)e^{-2(\kappa/\gamma)t} + \tilde{T} - 1 + 2\delta\tilde{T} + 2\frac{\kappa}{\gamma}\left(1-\tilde{T}\right)t\right]. \tag{5.4.53}$$

where we define $\tilde{T} \equiv T_0/T_f$. On the other hand, when $t \geq t_r$

$$\langle x_t^2 \rangle = \frac{k_{\mathrm{B}}T_f}{\kappa}\left[1 + \frac{e^{2\delta}-1}{2\delta}\left(\tilde{T}-1\right)e^{-2(\kappa/\gamma)t}\right]. \tag{5.4.54}$$

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

The previous expression thus takes the same form as the one corresponding to an instantaneous quench, but with a parameter that acts as a relative temperature,

$$\tilde{T}_{\text{eff}} \equiv \frac{e^{2\delta}-1}{2\delta}(\tilde{T}-1)+1. \tag{5.4.55}$$

All the theoretical curves plotted over the experimental data shown all over our work assume the previous expression for $\langle x_t^2 \rangle$.

In order to accurately observe the phenomena addressed in our study, the closest situation to ideality is to maintain the value of δ at or below one, as this aligns with the original theoretical model and ensures that the effects of the temperature ramp do not overpower the observations. However, our findings indicate that the phenomenology studied is relatively robust to the presence of a non-zero response time in the generator, and our theoretical predictions are undoubtedly substantiated by our experimental results even when δ exceeds unity (see Subsec. 5.4.8). Further examination of all the phenomenology derived from the protocol analyzed in this section would be necessary to fully understand the limits of the validity of those predictions.

Nat. Phys. (2024)—10.1038/s41567-023-02269-z

5.5 Asymmetric thermal relaxation in driven systems: Rotations go opposite ways **(Phys. Rev. Res. 2023)**

This section is a slightly adapted version of the publication C. Dieball, G. Wellecke and A. Godec, "Asymmetric thermal relaxation in driven systems: Rotations go opposite ways", *Phys. Rev. Res.* **5**, L042030 (2023).

My personal contribution as first author of the publication was in performing calculations and simulations, analyzing the results, and co-writing the manuscript.

Asymmetric Thermal Relaxation in Driven Systems: Rotations go Opposite Ways

Cai Dieball[1], Gerrit Wellecke[1,2] and Aljaž Godec[1]

[1]Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

[2]Present address: Theory of Biological Fluids, Max Planck Institute for Dynamics and Self-Organization, Göttingen 37077, Germany

Abstract

It was predicted and recently experimentally confirmed that systems with microscopically reversible dynamics in locally quadratic potentials warm up faster than they cool down. This thermal relaxation asymmetry challenged the local-equilibrium paradigm valid near equilibrium. Because the intuition and proof hinged on the dynamics obeying detailed balance, it was not clear whether the asymmetry persists in systems with irreversible dynamics. To fill this gap, we here prove the relaxation asymmetry for systems driven out of equilibrium by a general linear drift. The asymmetry persists due to a non-trivial isomorphism between driven and reversible processes. Moreover, rotational motions emerge that, strikingly, occur in opposite directions during heating and cooling. This highlights that noisy systems do not relax by passing through local equilibria.

Phys. Rev. Res. **5**, L042030 (2023)

5.5.1 Introduction

According to the laws of thermodynamics, systems in contact with a thermal environment evolve to the temperature of their surroundings in the process called *thermal relaxation* [25]. Relaxation close to equilibrium may be explained by linear response theory, conceptually based on Onsager's regression hypothesis [26, 27, 250]. That is, relaxation from a temperature quench is indistinguishable from the decay of a spontaneous thermal fluctuation at equilibrium [26, 27, 250]. Analogous results were meanwhile formulated also for relaxation near non-equilibrium steady states [121, 280, 284]. Beyond the linear regime, however, the regression hypothesis and perturbative arguments fail.

Important advances have been made in understanding relaxation beyond the linear regime addressing hydrodynamic limits [173, 285], barrier crossing in driven systems [286, 287], memory effects [288–297], far-from-equilibrium fluctuation-dissipation theorems [150, 298], optimal heating/cooling protocols [299], anomalous relaxation also

known as the *Mpemba effect* [251–257, 300] and its isothermal analogue [301], the Kovacs effect [260, 261], and dynamical phase transitions [302–310]. Important advances further include transient thermodynamic uncertainty relations [49, 71, 87, 126, 210, 211], speed limits [124, 266, 268, 311, 312], and analyses of relaxation from the viewpoint of information geometry [88, 124, 266].

A particularly striking feature of relaxation was unraveled with the discovery of the asymmetry between heating and cooling from thermodynamically equidistant temperature quenches [73]. That is, it was found that systems with locally quadratic energy landscapes and microscopically reversible dynamics heat up faster than they cool down. Later works expanded on this result [74–76]. The asymmetry was recently quantitatively confirmed by experiments [88].

The asymmetry emerges because the entropy production within the system during heating is more efficient than heat dissipation into the environment during cooling [73]. In turn, close to equilibrium they become equivalent and symmetry is restored [73, 88]. An even deeper understanding of the asymmetry was recently achieved by means of "thermal kinematics" [88]. However, both the reasoning and the proof of the asymmetry [73, 88, 313] seem to hinge on the reversibility of the dynamics. Therefore, the persistence of the asymmetry in systems driven into non-equilibrium steady states (NESS) was unexpected. In particular, a non-conservative force profoundly changes relaxation behavior [164, 204, 314, 315] even near stable fixed points [122] and in systems with linear drift [316], and may thus a priori also break the asymmetry.

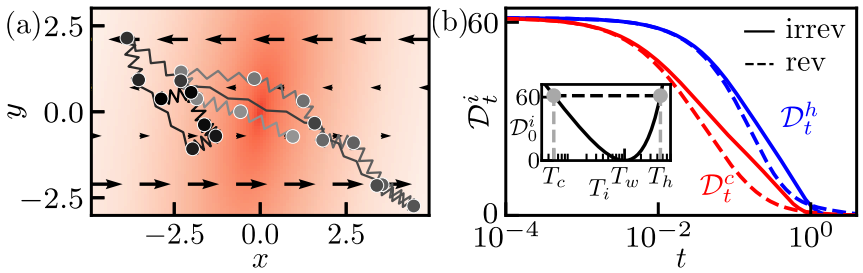

Figure 5.5.1: (a) Configuration of a harmonically confined (color gradient) Rouse polymer with $N = 20$ beads in 3d with hydrodynamic interactions and internal friction subject to a shear flow (arrows) in the x-y-plane drawn from the NESS with covariance $\mathbf{\Sigma}_{\mathrm{s},w}$ (see Subsec. 5.5.8 for parameters); a projection onto the x-y-plane is shown. (b) The corresponding free energy difference $\mathcal{D}_t^i$ in Eq. (5.5.5) during heating from T_c (red) and cooling from T_h (blue) with (solid lines) and without (dashed lines) irreversible shear flow. The shear changes $\mathcal{D}_t^i$, but the thermal relaxation asymmetry $\mathcal{D}_t^c < \mathcal{D}_t^h$ for $t > 0$ remains valid. Inset: Temperatures T_i before the quench are chosen thermodynamically equidistant, i.e., $\mathcal{D}_0^c = \mathcal{D}_0^h$.

Phys. Rev. Res. **5**, L042030 (2023)

Here, we investigate the speed and asymmetry of thermal relaxation to an NESS. As a paradigmatic example we first consider a harmonically confined Rouse polymer with hydrodynamic interactions and internal friction driven by shear flow (see Fig. 5.5.1), and demonstrate that heating is faster than cooling. Next we provide a systematic analysis of relaxation under broken detailed balance and explain under which conditions heating and cooling both become faster. Finally, we prove that all ergodic systems with a linear drift, including those driven arbitrarily far from equilibrium and displaying rotational motions, heat up faster than they cool down. In this regime, the notion of a local effective non-equilibrium temperature is nominally impossible. Our proof, which exploits dual-reversal symmetry, unravels a non-trivial isomorphism between reversible and driven systems. Finally, we find a new unexpected facet of the relaxation asymmetry—rotational motions occur in opposite directions during heating and cooling, respectively.

Phys. Rev. Res. **5**, L042030 (2023)

5.5.2 Setup and motivating example

The relaxation asymmetry was originally proven for reversible diffusions in locally quadratic energy landscapes, as well as their low-dimensional projections [73, 313]. It states that such systems, when quenched from thermodynamically equidistant (TED) temperatures T_h, T_c to an ambient temperature T_w with $T_c < T_w < T_h$, heat up faster than they cool down. In quantitative terms, the generalized excess free energy in units of $k_\mathrm{B}T_w$ [120, 122, 265, 317] or non-adiabatic entropy production [188, 318] (i.e., the relative entropy in units of k_B [248] between the instantaneous $P_i^w(\mathbf{x},t)$ and stationary $p_\mathrm{s}^w(\mathbf{x})$ probability density at T_w with $i = h, c$)

$$\mathcal{D}_t^i \equiv \mathcal{D}_\mathrm{KL}[P_i^w(\mathbf{x},t)||p_\mathrm{s}^w(\mathbf{x})] \equiv \int \mathrm{d}\mathbf{x} P_i^w(\mathbf{x},t) \ln \frac{P_i^w(\mathbf{x},t)}{p_\mathrm{s}^w(\mathbf{x})}, \tag{5.5.1}$$

is always smaller during heating [73, 313]. That is, $\mathcal{D}_t^c < \mathcal{D}_t^h$ for all $t > 0$ and all TED T_h and T_c.

In a strict sense, the asymmetry is to be understood as a statement about linearized drift around a local minimum in some high-dimensional energy landscape [73]; counterexamples for diffusion in rugged landscapes [73] and for small quenches also in sufficiently anharmonic wells [74] are known. The generalization to driven systems therefore involves a linear drift that, however, does not derive from a potential and breaks detailed balance. Our main result is the discovery and proof (see last section) of the asymmetry $\mathcal{D}_t^c < \mathcal{D}_t^h$ in driven systems.

Consider a d-dimensional system evolving according to the overdamped Langevin equation [6, 14]

$$\mathrm{d}\mathbf{x}_t = -\mathbf{A}\mathbf{x}_t \mathrm{d}t + \boldsymbol{\sigma}_i \mathrm{d}\mathbf{W}_t, \tag{5.5.2}$$

with square drift and noise-amplitude matrices, $\mathbf{A}$ and $\boldsymbol{\sigma}_i$, respectively. In terms of the friction matrix $\boldsymbol{\gamma}$, given by Stokes' law, the positive definite diffusion matrix reads $\mathbf{D}_i \equiv \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T/2 = k_\mathrm{B} T_i \boldsymbol{\gamma}^{-1}$ and thus depends linearly on temperature T_i. The external force $\mathbf{F}(\mathbf{x})$ yields a T_i-independent drift $-\mathbf{A}\mathbf{x} = \boldsymbol{\gamma}^{-1}\mathbf{F}(\mathbf{x})$, where $\mathbf{A}$ is generally non-symmetric but confining, i.e., the eigenvalues of $\mathbf{A}$ have positive real parts. Thus, $\mathbf{x}_t$ is ergodic but irreversible with zero-mean Gaussian NESS density $p_\mathrm{s}^i(\mathbf{x}) = (2\pi)^{-d/2} \det[\boldsymbol{\Sigma}_{\mathrm{s},i}]^{-1/2} \exp[-\mathbf{x}^T \boldsymbol{\Sigma}_{\mathrm{s},i}^{-1} \mathbf{x}/2]$ where the covariance $\boldsymbol{\Sigma}_{\mathrm{s},i}$ obeys the Lyapunov equation (see Supplemental Material in Subsec. 5.5.8)

$$\mathbf{A}\boldsymbol{\Sigma}_{\mathrm{s},i} + \boldsymbol{\Sigma}_{\mathrm{s},i}\mathbf{A}^T = 2\mathbf{D}_i = 2k_\mathrm{B} T_i \boldsymbol{\gamma}^{-1}, \tag{5.5.3}$$

and thus depends linearly on the temperature T_i. Eq. (5.5.3) implies for all T_i the decomposition into reversible $-\mathbf{A}_\mathrm{rev}\mathbf{x} \equiv \mathbf{D}_i \nabla \ln p_\mathrm{s}^i(\mathbf{x}) = -\mathbf{D}_i \boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\mathbf{x}$ and irreversible $-\mathbf{A}_\mathrm{irr}\mathbf{x} \equiv (-\mathbf{A} + \mathbf{A}_\mathrm{rev})\mathbf{x} = -\boldsymbol{\alpha}_i \boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\mathbf{x}$ drift [84], where $\boldsymbol{\alpha}_i^T = -\boldsymbol{\alpha}_i$ is an antisymmetric matrix[2].

We focus on temperature quenches—instantaneous changes of the environmental temperature at fixed drift. The thermodynamics of relaxation upon a quench $T_i \to T_w$ is fully specified by $\mathcal{D}_t^i$, as the adiabatic entropy production (housekeeping heat divided by T_w) [188] merely embodies the cost of maintaining the NESS [28] and thus need not be considered. Therefore, TED temperatures $T_{h,c}$ correspond to $\mathcal{D}_0^h = \mathcal{D}_0^c$ and are equal to those of a reversible system at the same T_w [73].

Since the initial condition is a zero-mean Gaussian with $\boldsymbol{\Sigma}_i^w(0) = \boldsymbol{\Sigma}_{\mathrm{s},i}$, the probability density is Gaussian for all times with $\boldsymbol{\Sigma}_i^w(t) \equiv \langle \mathbf{x}_t \mathbf{x}_t^T \rangle_i^w - \langle \mathbf{x}_t \rangle_i^w \langle \mathbf{x}_t^T \rangle_i^w$ given by (see Subsec. 5.5.8)

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Sigma}_i^w(t) = -\mathbf{A}\boldsymbol{\Sigma}_i^w(t) - \boldsymbol{\Sigma}_i^w(t)\mathbf{A}^T + 2\mathbf{D}_w \\ &\Rightarrow \boldsymbol{\Sigma}_i^w(t) = \boldsymbol{\Sigma}_{\mathrm{s},w} + \mathrm{e}^{-\mathbf{A}t}\left[\boldsymbol{\Sigma}_{\mathrm{s},i} - \boldsymbol{\Sigma}_{\mathrm{s},w}\right]\mathrm{e}^{-\mathbf{A}^T t}, \end{aligned} \tag{5.5.4}$$

[2] As $\boldsymbol{\Sigma}_{\mathrm{s},i}$ is invertible and symmetric Eq. (5.5.3) $\boldsymbol{\alpha}_i = (\mathbf{A} - \mathbf{D}_i\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1})\boldsymbol{\Sigma}_{\mathrm{s},i} = -\boldsymbol{\Sigma}_{\mathrm{s},i}(\mathbf{A}^T - \boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\mathbf{D}_i) = -\boldsymbol{\alpha}_i^T$. In fact $\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\mathbf{x}$ and $\boldsymbol{\alpha}_i\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\mathbf{x}$ are orthogonal since their scalar product yields an antisymmetric quadratic form $\mathbf{x}^T\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\boldsymbol{\alpha}_i\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\mathbf{x} = 0$ [122].

Phys. Rev. Res. **5**, L042030 (2023)

where $\langle\cdot\rangle_i^w$ denotes the average over all paths $\mathbf{x}_t$ at temperature T_w evolving from $p_{\rm s}^i(\mathbf{x})$. Note that $\mathbf{\Sigma}_{{\rm s},i} = T_i\mathbf{\Sigma}_{{\rm s},w}/T_w$ [see Eq. (5.5.3)]. Introducing $\delta\tilde{T}_i \equiv T_i/T_w - 1$, the generalized excess free energy reads (see Subsec. 5.5.8)

$$\mathcal{D}_t^i = \frac{1}{2}\delta\tilde{T}_i\,{\rm tr}\mathbf{X}(t) - \frac{1}{2}\ln\det\left[\mathbb{1} + \delta\tilde{T}_i\,\mathbf{X}(t)\right], \tag{5.5.5}$$

where we introduced the $d\times d$ matrix

$$\mathbf{X}(t) \equiv {\rm e}^{-\mathbf{A}t}\mathbf{\Sigma}_{{\rm s},w}{\rm e}^{-\mathbf{A}^T t}\mathbf{\Sigma}_{{\rm s},w}^{-1}, \tag{5.5.6}$$

which via Eq. (5.5.5) fully describes relaxation dynamics.

As a paradigmatic example for such processes, we consider a harmonically confined Rouse polymer with N beads experiencing hydrodynamic interactions [319, 320] and internal friction [321–324] subject to a shear flow, which was investigated experimentally in [189–192, 325–329]. For a representative configuration of the NESS ensemble, see Fig. 5.5.1a. One may also consider colloidal particles in the presence of non-conservative optical forces [330]. The effect of these forces is included in the $3N\times 3N$ drift matrix $\mathbf{A}$ and $3N\times 3N$ noise amplitude $\boldsymbol{\sigma}_i$ (see Subsec. 5.5.8). Evaluating $\mathcal{D}_t^i$ for the heating and cooling processes upon quenches from TED temperatures T_h and T_c we find $\mathcal{D}_t^c < \mathcal{D}_t^h$ for all $t > 0$. That is, heating is faster than cooling (the red line in Fig. 5.5.1b is at all times below the blue line). This agrees with the relaxation asymmetry predicted [73] and experimentally verified [88] in reversible systems, and provokes the question if this holds for any linear driving.

5.5.3 Systematics of breaking detailed balance

We now systematically assess the influence of non-equilibrium drifts on relaxation upon a temperature quench. As shown above, any linear drift $\mathbf{A}$ for $i = c, w, h$ decomposes as

$$\mathbf{A} = (\mathbf{D}_i + \boldsymbol{\alpha}_i)\mathbf{\Sigma}_{{\rm s},i}^{-1} \quad \text{with} \quad \boldsymbol{\alpha}_{\rm i}^{\rm T} = -\boldsymbol{\alpha}_{\rm i}. \tag{5.5.7}$$

Thus, by choosing any antisymmetric matrix $\boldsymbol{\alpha}_i$ we alter the NESS current as well as $\mathbf{X}(t)$, but neither $\mathbf{\Sigma}_{{\rm s},i}$ nor $p_{\rm s}(\mathbf{x})$. We can thus directly compare an NESS with the corresponding reversible system $\boldsymbol{\alpha}_i = \mathbf{0}$ with the same steady state. Note that such a direct comparison is not given in the example in Fig. 5.5.1, since the shear flow alters $\mathbf{\Sigma}_{{\rm s},i}$ as it is not of the form $\boldsymbol{\alpha}_i\mathbf{\Sigma}_{{\rm s},i}^{-1}$ with $\boldsymbol{\alpha}_i^T = -\boldsymbol{\alpha}_i$ (see Subsec. 5.5.8 for details about the consistent comparison of equilibrium versus non-equilibrium).

We now consider the influence of the non-equilibrium driving. For linear drift, the relaxation is governed by the eigenvalues of $\mathbf{A}$ [331, 332]. Since $\mathbf{\Sigma}_{{\rm s},i}$ is, by definition, symmetric

Phys. Rev. Res. 5, L042030 (2023)

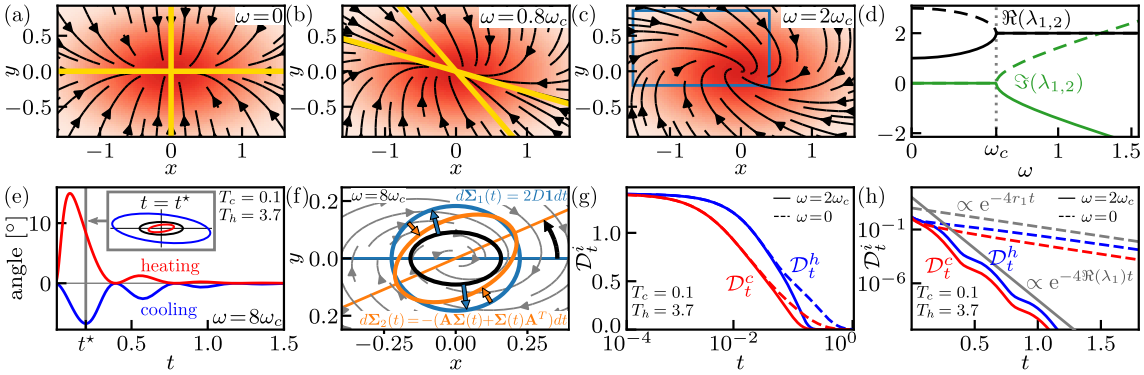

Figure 5.5.2: (a-c) Steady-state density $p_{\rm s}^w(\mathbf{x})$ (color gradient) and streamlines of the drift field $-\mathbf{A}\mathbf{x}$ for a 2d motion in Eq. (5.5.2) with $\boldsymbol{\sigma}_w = \sqrt{2}\mathbb{1}$ and drift matrix $\mathbf{A}$ with elements $A_{jj} = r_j$ with $r_1 = 1$, $r_2 = 3$, $A_{jk} = (-1)^j \omega r_k$ for $j,k \in \{1,2\}$, with ω in units of $\omega_c \equiv |r_2 - r_1|/2\sqrt{r_1 r_2}$. Real eigendirections (yellow) only exist for $\omega \leq \omega_c$. (d) Real and imaginary parts of eigenvalues of $\mathbf{A}$ as a function of ω. At $\omega = \omega_c$ the eigenvalues coincide and eigendirections (yellow lines in b,c) merge, i.e., $\mathbf{A}$ is not diagonalizable. For $\omega > \omega_c$ the eigenvalues are complex. (e) Angle between the covariance matrices $\boldsymbol{\Sigma}_i^w(t)$ and $\boldsymbol{\Sigma}_{{\rm s},w}$. (f) Explanation of the counter-intuitive opposing (effective) rotations at small times during heating from $T_c/T_w = 0.1$. The change $\mathrm{d}\boldsymbol{\Sigma}(t)$ in Eq. (5.5.4) starting from the initial $\boldsymbol{\Sigma}_{{\rm s},i}$ (black ellipse) for $\mathrm{d}t = 0.05$ split into diffusive (yielding the blue ellipse) and drift along the grey streamlines (yielding orange ellipse) contributions. (g-h) $\mathcal{D}_t^i$ for heating and cooling with and without driving on logarithmic-linear and linear-logarithmic scales. The driven system relaxes faster at large t as predicted from the eigenvalues in (e). Grey lines in (h) show the limiting relaxation rates for long times, $\mathrm{e}^{-4r_1 t}$ (dashed line) and $\mathrm{e}^{-4\Re(\lambda_1)t}$ (solid line).

with positive eigenvalues, we can find a matrix $\boldsymbol{\beta} = \boldsymbol{\beta}^T$ such that[3] $\boldsymbol{\beta}^2 \equiv \boldsymbol{\Sigma}_{{\rm s},i}^{-1}$. Thus, the matrix $\boldsymbol{\beta}\mathbf{D}_i\boldsymbol{\Sigma}_{{\rm s},i}^{-1}\boldsymbol{\beta}^{-1} = \boldsymbol{\beta}\mathbf{D}_i\boldsymbol{\beta} = \boldsymbol{\beta}\boldsymbol{\sigma}_i(\boldsymbol{\beta}\boldsymbol{\sigma}_i)^T/2$ is symmetric which alongside $\det(\boldsymbol{\beta}\boldsymbol{\sigma}_i) \neq 0$ implies that $\mathbf{D}_i\boldsymbol{\Sigma}_{{\rm s},i}^{-1}$ is diagonalizable with positive eigenvalues[4]. Therefore, in the absence of driving, $\mathbf{A} = \mathbf{D}_i\boldsymbol{\Sigma}_{{\rm s},i}^{-1}$ expectedly has strictly positive eigenvalues reflecting a monotonous relaxation to equilibrium.

Once we include driving $\boldsymbol{\alpha}_w \neq 0$ in the steady-state-preserving form Eq. (5.5.7), the spectrum may or may not become complex depending on the detailed form of $\boldsymbol{\alpha}_w$, see e.g., Fig. 5.5.2a-d. Complex eigenvectors imply that eigendirections where the drift points "straight" towards $\mathbf{0}$ cease to exist, see Fig. 5.5.2a-c. This happens already at arbitrarily small driving if level sets of $p_{\rm s}(\mathbf{x})$ are (hyper)spherical. If some eigenvalues are on the threshold of becoming complex (branching point ω_c in Fig. 5.5.2d), $\mathbf{A}$ may become

[3] From the orthogonal diagonalization $\mathbf{O}\boldsymbol{\Sigma}_{{\rm s},i}^{-1}\mathbf{O}^T = \mathrm{diag}(s_j)$, we define $\boldsymbol{\beta} \equiv \mathbf{O}^T \mathrm{diag}\sqrt{s_j}\mathbf{O}$.

[4] Any matrix of the form $\mathbf{M}\mathbf{M}^T$ is symmetric, and therefore diagonalizable, with real non-negative eigenvalues, since $\mathbf{M}\mathbf{M}^T\mathbf{v} = \lambda\mathbf{v}$ implies $\lambda = \mathbf{v}^T\mathbf{M}\mathbf{M}^T\mathbf{v}/\mathbf{v}^T\mathbf{v} = (\mathbf{M}^T\mathbf{v})^T\mathbf{M}^T\mathbf{v}/\mathbf{v}^T\mathbf{v} \geq 0$.

Phys. Rev. Res. **5**, L042030 (2023)

non-diagonalizable. In terms of the minimal 2d example in Fig. 5.5.2 we have that $\mathbf{A}$ is non-diagonalizable when $\omega = \pm\omega_c$ (see Fig. 5.5.2d).

An interesting consequence of driving is that the different dimensions no longer decouple as they do under detailed balance (see Fig. 5.5.2a). This means that the d-dimensional Langevin equation (5.5.2) cannot be decomposed into 1d equations and that rotational dynamics may emerge. In the particular case of temperature quenches, we find that driving causes a time-dependent rotation of the level sets of $P_i^w(\mathbf{x}, t)$, see Fig. 5.5.2e. In agreement with the opposite sings of $T_i - T_w$ in Eq. (5.5.4), these rotations occur in opposite directions during heating and cooling, which is a striking new feature of the relaxation asymmetry. The asymmetry implies that thermal relaxation must not be understood as passing through local equilibria at intermediate (effective) temperatures [25], since this would imply a symmetric relaxation independent of the sign of the temperature quench. Moreover, the rotation in opposite directions emphasizes that heating and cooling here evolve along very distinct pathways in the space of probability distributions (see also [88]).

While the initial rotation during cooling follows the direction of driving, most surprisingly the effective rotations during heating initially oppose the direction of the driving (see Fig. 5.5.2e). This effect can be traced to the interplay of ("Trotterized" [333]) diffusion and drift during individual small time increments, see Fig. 5.5.2f. During heating for an increment $\mathrm{d}t$ diffusion alone propagates the black to the more circular blue ellipse. The subsequent drift along the elliptical streamlines propagates this blue ellipse to the orange ellipse that is, however, effectively rotated in the direction opposite to the drift (for further details see Subsec. 5.5.8).

Phys. Rev. Res. 5, L042030 (2023)

5.5.4 Accelerated relaxation

Before proving the relaxation asymmetry we discuss the acceleration of relaxation via driving [164, 204, 315, 316, 316]. We therefore focus on the real part of the eigenvalues, which determines the relaxation time-scales. Upon a change of basis we find $\widetilde{\mathbf{A}} \equiv \boldsymbol{\beta}\mathbf{A}\boldsymbol{\beta}^{-1} = \boldsymbol{\beta}\mathbf{D}_i\boldsymbol{\beta} + \boldsymbol{\beta}\boldsymbol{\alpha}_i\boldsymbol{\beta}$ where $(\boldsymbol{\beta}\boldsymbol{\alpha}_i\boldsymbol{\beta})^T = -\boldsymbol{\beta}\boldsymbol{\alpha}_i\boldsymbol{\beta}$. Then, for any complex eigenvalue λ of $\widetilde{\mathbf{A}}$ with eigenvector $\mathbf{v} \neq 0$ we may write $2\Re(\lambda)\mathbf{v}^\dagger\mathbf{v} = (\lambda + \lambda^\dagger)\mathbf{v}^\dagger\mathbf{v} = \mathbf{v}^\dagger(\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^\dagger)\mathbf{v} = 2\mathbf{v}^\dagger\boldsymbol{\beta}\mathbf{D}_i\boldsymbol{\beta}\mathbf{v}$, where $\dagger$ denotes the Hermitian adjoint. Decomposing $\mathbf{v}$, $\mathbf{v}^\dagger$ in the orthonormal eigenbasis of $\boldsymbol{\beta}\mathbf{D}_i\boldsymbol{\beta}$ with eigenvalues $0 < \mu_1 \leq \cdots \leq \mu_d$, we have with $c_j \in \mathbb{C}$

$$\Re(\lambda) = \frac{\mathbf{v}^\dagger\boldsymbol{\beta}\mathbf{D}_i\boldsymbol{\beta}\mathbf{v}}{\mathbf{v}^\dagger\mathbf{v}} = \frac{\sum_{j=1}^d c_j^\dagger c_j \mu_j}{\sum_{j=1}^d c_j^\dagger c_j} \in [\mu_1, \mu_d]. \tag{5.5.8}$$

This means that the real parts of the eigenvalues in the presence of driving remain not only positive, as required for the existence of a steady state, but even remain in the interval $[\mu_1, \mu_d]$. Thus, Eq. (5.5.8) states that the smallest real part of eigenvalues of $\mathbf{A}$ under driving obeys $\Re(\lambda_1) \geq \mu_1$. Note that $\Re(\lambda_1)$ typically[5] sets the slowest relaxation rate [331, 332]. Since $\Re(\lambda_1)$ increases (or does not decrease) upon driving, the latter typically enhances relaxation on long timescales, as already shown in [316].

Driving also affects the adiabatic entropy production. This effect, however, scales trivially, as the adiabatic entropy production increases with increasing $\boldsymbol{\alpha}_i$ according to $\boldsymbol{\alpha}_i^T \mathbf{D}_i^{-1} \boldsymbol{\alpha}_i$, see Subsec. 5.5.8. Hence, there is no direct connection between faster relaxation and steady-state dissipation, as the influence of driving on the eigenvalues is specific. For example, the acceleration in $d = 2$ saturates [see $\Re(\lambda_1)$ in Fig. 5.5.2d]. More drastically, multiplying $\boldsymbol{\alpha}_i$ by a factor larger than 1 in $d = 3$ may decrease $\Re(\lambda_1)$ [316].

We see from Eq. (5.5.6) that $\mathbf{X}(t) \sim \mathrm{e}^{-2\Re(\lambda_1)t}$ for long times and therefore $\mathcal{D}_t^i \sim \mathrm{e}^{-4\Re(\lambda_1)t}$ (see Subsec. 5.5.8 and Fig. 5.5.2g-h). The statement "accelerated relaxation", $\Re(\lambda_1) \geq \mu_1$, means that both, heating and cooling will at long times be faster. In general the difference between heating and cooling upon driving can become larger or smaller than for reversible dynamics with the same $\boldsymbol{\Sigma}_{\mathrm{s},i}$, but as we now prove heating is always faster than cooling.

5.5.5 Proof of relaxation asymmetry in driven systems

We now prove the relaxation asymmetry for the dynamics in Eq. (5.5.2), i.e., $\Delta\mathcal{D}_t \equiv \mathcal{D}_t^h - \mathcal{D}_t^c > 0$ for all $t > 0$. By Eq. (5.5.6)

$$\Delta\mathcal{D}_t = \frac{\delta\tilde{T}_h - \delta\tilde{T}_c}{2}\,\mathrm{tr}\mathbf{X}(t) - \frac{1}{2}\ln\frac{\det\left[\mathbb{1} + \delta\tilde{T}_h\,\mathbf{X}(t)\right]}{\det\left[\mathbb{1} + \delta\tilde{T}_c\,\mathbf{X}(t)\right]}. \tag{5.5.9}$$

To prove the asymmetry we must understand the properties of $\mathbf{X}(t)$, which is T_i-independent. Using the steady-state Lyapunov equation (5.5.3) we can rewrite $\mathbf{X}(t)$ as

$$\mathbf{X}(t) = \mathrm{e}^{-\mathbf{A}t}\mathrm{e}^{-\mathbf{A}_{-\boldsymbol{\alpha}}t}, \tag{5.5.10}$$

where $\mathbf{A}_{-\boldsymbol{\alpha}} \equiv (\mathbf{D}_w - \boldsymbol{\alpha}_w)\boldsymbol{\Sigma}_{\mathrm{s},w}^{-1}$ is the driving-reversed version of $\mathbf{A}$ as in Eq. (5.5.7). This form is reminiscent of the dual-reversal symmetry [83, 84, 133, 176] stating that time-reversal in non-equilibrium steady states requires concurrent current reversal. Eq. (5.5.10) is illustrated in Fig. 5.5.3a. The proof again requires changing the basis via $\boldsymbol{\beta}$ as

[5] Unless the initial distribution has only a negligible projection onto the slowest modes.

Phys. Rev. Res. **5**, L042030 (2023)

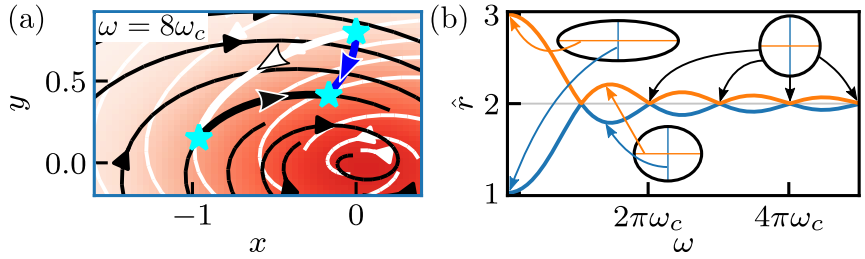

Figure 5.5.3: (a) Illustration of Eq. (5.5.10): Streamplot of the drift field $-\mathbf{A}\mathbf{x}$ (black) as in blue frame in Fig. 5.5.2c, and inverted drift field $-\mathbf{A}_{-\boldsymbol{\alpha}}\mathbf{x}$ (white). The white line depicts $\mathrm{e}^{-\mathbf{A}_{-\alpha}\tau}\mathbf{x}_0$ for $\tau \in [0,t]$, the black line is $\mathrm{e}^{-\mathbf{A}\tau}\mathrm{e}^{-\mathbf{A}_{-\alpha}t}\mathbf{x}_0$, and the blue line shows $\mathbf{X}(\tau)\mathbf{x}_0$. (b) Effective stiffness $\hat{r}_j(\omega) \equiv -\ln(x_j^t)/2t$ at $t=1$ as a function of driving ω (see Subsec. 5.5.8). For large driving the directions mix, such that the system effectively approaches a circular parabola with stiffness $(r_1+r_2)/2$, which is the real part of eigenvalues in Fig. 5.5.2d.

$$\widetilde{\mathbf{X}}(t) \equiv \boldsymbol{\beta}\mathbf{X}(t)\boldsymbol{\beta}^{-1} = \mathrm{e}^{-\widetilde{\mathbf{A}}t}\left(\mathrm{e}^{-\widetilde{\mathbf{A}}t}\right)^T, \tag{5.5.11}$$

where we used $\boldsymbol{\beta}\mathbf{A}_{-\alpha}\boldsymbol{\beta}^{-1} = \widetilde{\mathbf{A}}^T$ and $\mathrm{e}^{-\widetilde{\mathbf{A}}^T t} = (\mathrm{e}^{-\widetilde{\mathbf{A}}t})^T$. Thus, $\widetilde{\mathbf{X}}(t)$ is symmetric and hence diagonalizable with real eigenvalues. Since, $\det \mathrm{e}^{-\widetilde{\mathbf{A}}t} = \mathrm{e}^{-\mathrm{tr}\widetilde{\mathbf{A}}t}$, we have $\det \widetilde{\mathbf{X}}(t) = \mathrm{e}^{-2\mathrm{tr}\widetilde{\mathbf{A}}t} \neq 0$. Therefore, $\widetilde{\mathbf{X}}(t)$ and thus $\mathbf{X}(t)$ have positive eigenvalues[6] $x_j^t > 0$, $j = 1,\dots,d$. Although $\mathbf{A}$ may have complex eigenvalues or even be non-diagonalizable and $\exp(-\mathbf{A}t)$ may be rotational (see Figs. 5.5.2c and 5.5.3a), $\mathbf{X}(t)$ has a real eigensystem since consecutive rotations in forward and current-reversed directions effectively cancel rotations, see Eq. (5.5.10) and Fig. 5.5.3a.

Using the eigenvalues $x_j^t > 0$, we rewrite Eq. (5.5.9) as

$$\Delta\mathcal{D}_t = \sum_{j=1}^{d}\left(\frac{\delta\tilde{T}_h - \delta\tilde{T}_c}{2}x_j^t - \frac{1}{2}\ln\left[\frac{1+\delta\tilde{T}_h x_j^t}{1+\delta\tilde{T}_c x_j^t}\right]\right). \tag{5.5.12}$$

If all $x_j^t \in (0,1)$, the proof for reversible systems [73, 313] asserts that $\Delta\mathcal{D}_t > 0$. It therefore suffices to show that $x_j^t < 1$ for all j, which is equivalent to $||\mathbf{X}(t)|| < 1$, where $||\mathbf{M}|| \equiv \sup_{\mathbf{v}\in\mathbb{R}^d\setminus\mathbf{0}} ||\mathbf{M}\mathbf{v}||_2 \,/\, ||\mathbf{v}||_2$ and $||\mathbf{v}||_2 = \sqrt{\mathbf{v}^T\mathbf{v}}$ are the matrix and Euclidean norm, respectively. Eq. (5.5.10) does not help in showing this[7]; although eigenvalues of $\mathbf{A}$ have positive real parts [see Eq. (5.5.8)], it may be that $\left|\left|\mathrm{e}^{-\mathbf{A}\pm\boldsymbol{\alpha}t}\right|\right| > 1$ (e.g., the distance to $\mathbf{0}$ in Fig. 5.5.3a increases along the white line). This is possible because the eigenvectors of $\mathbf{A}$ are not orthogonal.

[6] Any matrix of the form $\mathbf{M}\mathbf{M}^T$ is symmetric, and therefore diagonalizable, with real non-negative eigenvalues, since $\mathbf{M}\mathbf{M}^T\mathbf{v} = \lambda\mathbf{v}$ implies $\lambda = \mathbf{v}^T\mathbf{M}\mathbf{M}^T\mathbf{v}/\mathbf{v}^T\mathbf{v} = (\mathbf{M}^T\mathbf{v})^T\mathbf{M}^T\mathbf{v}/\mathbf{v}^T\mathbf{v} \geq 0$.

[7] Eq. (5.5.10) suffices only at equilibrium $\mathbf{A} = \mathbf{A}_{-\boldsymbol{\alpha}} = \mathbf{D}_w\boldsymbol{\Sigma}_{\mathrm{s},w}^{-1}$ where $\mathbf{X}(t) = \exp(-2\mathbf{D}_w\boldsymbol{\Sigma}_{\mathrm{s},w}^{-1}t)$ decomposes into $x_j^t = \exp(-2\mu_j t) < 1$.

Phys. Rev. Res. **5**, L042030 (2023)

We thus change the basis as in Eq. (5.5.11) and use the log-norm inequality $||\exp(\mathbf{M}t)|| \leq \exp[\mu(\mathbf{M})t]$ [334] with log norm $\mu(\mathbf{M}) \equiv \lim_{h\to 0^+} h^{-1}\left(||\mathbb{1} + h\mathbf{M}|| - 1\right)$ yielding $\mu(-\widetilde{\mathbf{A}}) \equiv \mu(-\boldsymbol{\beta}\mathbf{A}\boldsymbol{\beta}^{-1}) = \mu(-\boldsymbol{\beta}\mathbf{D}_w\boldsymbol{\beta}) = -\mu_1$ determined by the symmetric part $(\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^T)/2 = \boldsymbol{\beta}\mathbf{D}_w\boldsymbol{\beta}$, see Subsec. 5.5.8. This basis is appropriate because $\boldsymbol{\beta}\boldsymbol{\alpha}_w\boldsymbol{\beta}$ in $\widetilde{\mathbf{A}}$ (unlike $\boldsymbol{\alpha}_w\boldsymbol{\Sigma}_{\mathrm{s},w}^{-1}$ in $\mathbf{A}$) has no symmetric part, i.e., the driving only affects the rotational part. The log-norm inequality thus implies $\left|\left|\exp(-\widetilde{\mathbf{A}}t)\right|\right| \leq \exp[\mu(-\widetilde{\mathbf{A}})t] = \exp(-\mu_1 t)$ and similarly $\left|\left|\exp(-\widetilde{\mathbf{A}}^T t)\right|\right| \leq \exp(-\mu_1 t)$, and by the submultiplicative property of the matrix norm we obtain from Eq. (5.5.11)

$$\left|\left|\widetilde{\mathbf{X}}(t)\right|\right| \leq \left|\left|\mathrm{e}^{-\widetilde{\mathbf{A}}t}\right|\right| \left|\left|\left(\mathrm{e}^{-\widetilde{\mathbf{A}}t}\right)^T\right|\right| \leq \mathrm{e}^{-2\mu_1 t} < 1. \tag{5.5.13}$$

Since $\left|\left|\widetilde{\mathbf{X}}(t)\right|\right| = ||\mathbf{X}(t)||$ this implies $x_j^t < 1$ and with Eq. (5.5.12) completes the proof of $\Delta\mathcal{D}_t > 0$ for all $t > 0$.

The proof provides important insight into the thermodynamics of the asymmetry in reversible versus driven systems. Namely, $\Delta\mathcal{D}_t$ in Eq. (5.5.12) for a driven system at any t is equal to that of any reversible system with drift matrix $\hat{\mathbf{A}}$ having eigenvalues $\hat{\mu}_i$ satisfying $\mathrm{e}^{-2\hat{\mu}_j t} = x_j^t$. Therefore, at each t the relaxation asymmetry of a driven system is isomorphic to that of an equilibrium system with different geometry (see Fig. 5.5.3b for effective stiffness axes of the 2-dimensional parabolic potential), which implies the persistence of the asymmetry. This provokes intriguing questions about the existence of the asymmetry in the presence of time-dependent driving.

5.5.6 Conclusion

We have proven that overdamped ergodic systems driven by linear drift, conservative or not, for any pair of thermodynamically equidistant temperature quenches warm up faster than they cool down. The relaxation asymmetry [73], which was recently confirmed experimentally [88], therefore persists in driven systems. As the original proof hinged on microscopic reversibility, this finding is surprising and is explained by a non-trivial isomorphism between driven and reversible processes. In the presence of driving, a striking new feature of the relaxation asymmetry appears: rotational dynamics emerge with opposite directions during heating and cooling, respectively. This further highlights that small, noisy systems do not relax by passing through local equilibria [25]. Moreover, rotations in opposing directions emphasize that heating and cooling evolve along fundamentally distinct pathways [88]. An analysis with the framework of "thermal kinematics" [88] will bring even deeper insight. Our results motivate further studies on the existence of the relaxation

Phys. Rev. Res. **5**, L042030 (2023)

asymmetry in temporally driven systems [49, 263, 264, 275–277], systems with non-linear drift [147, 251, 253, 254, 256], and in the presence of inertial effects [261].

5.5.7 Acknowledgments

Financial support from Studienstiftung des Deutschen Volkes (to C. D.) and the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2 (to A. G.) is gratefully acknowledged.

5.5.8 Supplementary Material

In this Supplementary Material we provide further details on model examples, arguments, and calculations presented above. Besides several technical details, we give the equations and parameters describing the Rouse chain in confined shear flow and derive and solve equations for the covariance. We extensively elaborate on the rotations in different directions, and address the consistent comparison of equilibrium and non-equilibrium steady states. We conclude with a discussion of the log-norm inequality.

Rouse polymer with hydrodynamic interactions and internal friction in confined shear flow

In Fig. 5.5.1 we consider the motivating example of a polymer chain with $N = 20$ beads in $3d$ space represented by the Rouse model with internal friction and hydrodynamic interactions in shear flow. That is, we assume that the beads are connected by harmonic springs with zero rest length [Eq. (5.5.14)] and additionally interact via hydrodynamic interactions [Eq. (5.5.16)] and experience internal friction [Eq. (5.5.17)]. The chain is confined in a parabolic potential and is subject to a shear flow [Eqs. (5.5.21)-(5.5.22)]. We now describe the interactions and evolution equations individually, with increasing complexity.

In the classical Rouse model (i.e., without hydrodynamic interactions, internal friction, confinement and shear), the time-dependent position of the beads $\mathbf{x}_t$ is described by the $3N$ dimensional Langevin equation (denoting the spring stiffness by κ and solvent friction by γ)

$$\gamma \mathrm{d}\mathbf{x}_t = -\kappa \mathbf{k} \mathbf{x}_t \mathrm{d}t + \sqrt{2D_i}\mathrm{d}\mathbf{W}_t, \tag{5.5.14}$$

Phys. Rev. Res. **5**, L042030 (2023)

where the connectivity matrix $\mathbf{k}$ is a $3N \times 3N$ matrix that reads ($\mathbb{1}_3$ is the 3d unit matrix and all terms not shown are 0)

$$\mathbf{k} = \begin{bmatrix} \mathbb{1}_3 & -\mathbb{1}_3 & & & & & \\ -\mathbb{1}_3 & 2\mathbb{1}_3 & -\mathbb{1}_3 & & & 0 & \\ & -\mathbb{1}_3 & 2\mathbb{1}_3 & -\mathbb{1}_3 & & & \\ & & & \dots & \dots & & \\ & & & & \dots & \dots & \\ & 0 & & & -\mathbb{1}_3 & 2\mathbb{1}_3 & -\mathbb{1}_3 \\ & & & & & -\mathbb{1}_3 & \mathbb{1}_3 \end{bmatrix}. \tag{5.5.15}$$

Note that the temperature dependence is contained in the diffusion constant $D_i \propto T_i$ as described Subsec. 5.5.2. Following Ref. [321], we introduce hydrodynamic interactions in the pre-averaging approximation via the $3N \times 3N$ matrix $\mathbf{H}$ with 3×3 entries for $m, n = 1, \dots, N$

$$H_{mn} = \gamma^{-1}\mathbb{1}_3 \times \begin{cases} 1 & \text{if } n = m \\ \frac{1}{2\pi^2\gamma|n-m|^{1/2}} & \text{if } n \neq m \end{cases}, \tag{5.5.16}$$

and also include internal friction with friction parameter ξ_{IF}, such that the equation of motion of the chain becomes

$$\mathrm{d}\mathbf{x}_t = \mathbf{H}\left(-\kappa\mathbf{k}\mathbf{x}_t\mathrm{d}t - \xi_{\mathrm{IF}}\mathbf{k}\mathrm{d}\mathbf{x}_t + \sqrt{2D_i}\mathrm{d}\mathbf{W}_t\right). \tag{5.5.17}$$

We now introduce a spatial confinement via a $3d$ harmonic potential centered at $\mathbf{x} = \mathbf{0}$ and subject the confined chain to a shear flow in the x-y-plane. To account for the confinement, we consider the three-dimensional matrix ($r_i > 0$)

$$\mathbf{C} = \mathbf{R}(\theta)\begin{bmatrix} r_x & 0 & 0 \\ 0 & r_y & 0 \\ 0 & 0 & r_z \end{bmatrix}\mathbf{R}(\theta)^T, \tag{5.5.18}$$

and for the shear flow in the x-y-plane

$$\mathbf{S} = \begin{bmatrix} 0 & \omega & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.5.19}$$

Phys. Rev. Res. **5**, L042030 (2023)

Note that by introducing the shear flow $\mathbf{S}$ to the system, the microscopic reversibility is lost, i.e., we have an example that is genuinely driven out of equilibrium.

To avoid the special case where shear and confinement are orthogonal, we introduce the rotation matrix $\mathbf{R}(\theta)$ to rotate the confinement by an angle θ within the x-$y-$plane,

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.5.20}$$

Now consider a $3N$ dimensional matrix $\mathbf{M}$ with $\mathbf{R}(\theta)\mathbf{C}\mathbf{R}^T(\theta) + \mathbf{S}$ on the N diagonal 3×3 blocks to impose the confinement and shear on each bead. Then the equation of motion for the confined polymer in shear flow reads

$$\mathrm{d}\mathbf{x}_t = \mathbf{H}\left(-(\kappa\mathbf{k} + \mathbf{M})\mathbf{x}_t\mathrm{d}t - \xi_{\mathrm{IF}}\mathbf{k}\mathrm{d}\mathbf{x}_t + \sqrt{2D_i}\mathrm{d}\mathbf{W}_t\right). \tag{5.5.21}$$

This may now be rewritten as

$$\begin{aligned} \left(\mathbf{H}^{-1} + \xi_{\mathrm{IF}}\mathbf{k}\right)\mathrm{d}\mathbf{x}_t &= -(\kappa\mathbf{k} + \mathbf{M})\mathbf{x}_t\mathrm{d}t + \sqrt{2D_i}\mathrm{d}\mathbf{W}_t, \\ \mathrm{d}\mathbf{x}_t &= -\left(\mathbf{H}^{-1} + \xi_{\mathrm{IF}}\mathbf{k}\right)^{-1}(\kappa\mathbf{k} + \mathbf{M})\mathbf{x}_t\mathrm{d}t + \sqrt{2D_i}\left(\mathbf{H}^{-1} + \xi_{\mathrm{IF}}\mathbf{k}\right)^{-1}\mathrm{d}\mathbf{W}_t. \end{aligned} \tag{5.5.22}$$

Upon identifying

$$\mathbf{A} = \left(\mathbf{H}^{-1} + \xi_{\mathrm{IF}}\mathbf{k}\right)^{-1}(\kappa\mathbf{k} + \mathbf{M}), \quad \boldsymbol{\sigma}_i = \sqrt{2D_i}\left(\mathbf{H}^{-1} + \xi_{\mathrm{IF}}\mathbf{k}\right)^{-1} \tag{5.5.23}$$

we arrive at the desired form $\mathrm{d}\mathbf{x}_t = -\mathbf{A}\mathbf{x}_t\mathrm{d}t + \boldsymbol{\sigma}_i\mathrm{d}\mathbf{W}_t$ as described in Subsec. 5.5.2.

The parameters used in Fig. 5.5.1 are $N = 20$, $r_x = 0$, $r_y = 0.1$, $r_z = 1$, $\theta = -10°$, $\gamma = 1$, $\xi_{\mathrm{IF}} = 1$, $D_w = 1$, $\kappa = 1$, $T_c = 0.05T_w$, $T_h = 4.56T_w$. In Fig. 5.5.1a we choose $\omega = 3$, while in Fig. 5.5.1b we use $\omega = 20$ to emphasize the differences between the curves. For the chosen parameters, several eigenvalues of the matrix $\mathbf{A}$ become complex, thus confirming that the Rouse model in the shear flow is an irreversible process.

Lyapunov equation and time-dependent covariance

Here we derive Eqs. (5.5.3) and (5.5.4). We consider dynamics governed by $\mathrm{d}\mathbf{x}_t = -\mathbf{A}\mathbf{x}_t\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t$ as in Eq. (5.5.2). Taking the mean value gives $\frac{\mathrm{d}}{\mathrm{d}t}\langle\mathbf{x}_t\rangle = -\mathbf{A}\langle\mathbf{x}_t\rangle$, which implies $\langle\mathbf{x}_t\rangle = \mathrm{e}^{-\mathbf{A}t}\langle\mathbf{x}_0\rangle$. For the scope of this manuscript, we only consider initial conditions with $\langle\mathbf{x}_0\rangle = \mathbf{0}$ such that for all times $\langle\mathbf{x}_t\rangle = 0$. The covariance $\boldsymbol{\Sigma}(t) \equiv \langle\mathbf{x}_t\mathbf{x}_t^T\rangle - \langle\mathbf{x}_t\rangle\langle\mathbf{x}_t^T\rangle =$

Phys. Rev. Res. **5**, L042030 (2023)

$\langle \mathbf{x}_t \mathbf{x}_t^T \rangle$ is always symmetric $\mathbf{\Sigma}(t)^T = \mathbf{\Sigma}(t)$ with strictly positive eigenvalues. Using Itô's Lemma [6, 14] we see that $\mathbf{\Sigma}(t)$ obeys the differential Lyapunov equation

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{\Sigma}(t) &= \langle \mathrm{d}\mathbf{x}_t \mathbf{x}_t^T \rangle + \langle \mathbf{x}_t \mathrm{d}\mathbf{x}_t^T \rangle + \langle \mathrm{d}\mathbf{x}_t \mathrm{d}\mathbf{x}_t^T \rangle \\ &= -\mathbf{A}\langle \mathbf{x}_t \mathbf{x}_t^T \rangle - \langle \mathbf{x}_t \mathbf{x}_t^T \rangle \mathbf{A}^T + \boldsymbol{\sigma}\boldsymbol{\sigma}^T \\ &= -\mathbf{A}\mathbf{\Sigma}(t) - \mathbf{\Sigma}(t)\mathbf{A}^T + 2\mathbf{D}. \end{aligned} \tag{5.5.24}$$

In the steady state (i.e., for $\mathbf{A}$ originating from a confining potential and $t \to \infty$) this approaches the steady-state covariance $\mathbf{\Sigma}_\mathrm{s}$ obeying the algebraic (i.e., non-differential) Lyapunov equation (Eq. (5.5.3), see also Ref. [6])

$$\mathbf{A}\mathbf{\Sigma}_\mathrm{s} + \mathbf{\Sigma}_\mathrm{s}\mathbf{A}^T = 2\mathbf{D}. \tag{5.5.25}$$

Given the solution $\mathbf{\Sigma}_\mathrm{s}$ of Eq. (5.5.25), the solution for Eq. (5.5.24) for an initial condition with covariance $\mathbf{\Sigma}(0)$ is obtained as

$$\mathbf{\Sigma}(t) = \mathbf{\Sigma}_\mathrm{s} + \mathrm{e}^{-\mathbf{A}t}\left[\mathbf{\Sigma}(0) - \mathbf{\Sigma}_\mathrm{s}\right]\mathrm{e}^{-\mathbf{A}^T t}. \tag{5.5.26}$$

This is proven by taking the derivative of the ansatz,

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Sigma}_\mathrm{s} + \mathrm{e}^{-\mathbf{A}t}\left[\mathbf{\Sigma}(0) - \mathbf{\Sigma}_\mathrm{s}\right]\mathrm{e}^{-\mathbf{A}^T t}\right) \\ &= -\mathbf{A}\left(\mathrm{e}^{-\mathbf{A}t}\left[\mathbf{\Sigma}(0) - \mathbf{\Sigma}_\mathrm{s}\right]\mathrm{e}^{-\mathbf{A}^T t}\right) - \left(\mathrm{e}^{-\mathbf{A}t}\left[\mathbf{\Sigma}(0) - \mathbf{\Sigma}_\mathrm{s}\right]\mathrm{e}^{-\mathbf{A}^T t}\right)\mathbf{A}^T \\ &= -\mathbf{A}\mathbf{\Sigma}(t) + \mathbf{A}\mathbf{\Sigma}_\mathrm{s} - \mathbf{\Sigma}(t)\mathbf{A}^T + \mathbf{\Sigma}_\mathrm{s}\mathbf{A}^T \\ &\overset{(5.5.25)}{=} -\mathbf{A}\mathbf{\Sigma}(t) - \mathbf{\Sigma}(t)\mathbf{A}^T + 2\mathbf{D}. \end{aligned} \tag{5.5.27}$$

Choosing $\mathbf{\Sigma}(0) = \mathbf{\Sigma}_{\mathrm{s},i} = T_i \mathbf{\Sigma}_{\mathrm{s},w}/T_w$ yields Eq. (5.5.4).

Generalized excess free energy during heating and cooling

The Kullback-Leibler divergence in Eq. (5.5.1) can be computed for two d-dimensional Gaussian densities $P_{1,2}$ with mean zero as $2D_\mathrm{KL}\left(P_1||P_2\right) = -\ln\left(\det\left[\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\right]\right) + \mathrm{tr}\left(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1} - \mathbb{1}\right)$ where $\mathbb{1}$ is the d-dimensional unit matrix. Using Eq. (5.5.4) (i.e., $\mathbf{\Sigma}_i^w(t) = \mathbf{\Sigma}_{\mathrm{s},w} + \mathrm{e}^{-\mathbf{A}t}\left[\mathbf{\Sigma}_{\mathrm{s},i} - \mathbf{\Sigma}_{\mathrm{s},w}\right]\mathrm{e}^{-\mathbf{A}^T t}$) with the notations $\mathbf{X}(t) \equiv \mathrm{e}^{-\mathbf{A}t}\mathbf{\Sigma}_{\mathrm{s},w}\mathrm{e}^{-\mathbf{A}^T t}\mathbf{\Sigma}_{\mathrm{s},w}^{-1}$ and $\delta\tilde{T}_i \equiv T_i/T_w - 1$ we obtain Eq. (5.5.5), i.e.,

$$\mathcal{D}_t^i = \frac{1}{2}\delta\tilde{T}_i\,\mathrm{tr}\mathbf{X}(t) - \frac{1}{2}\ln\det\left[\mathbb{1} + \delta\tilde{T}_i\,\mathbf{X}(t)\right]. \tag{5.5.28}$$

Phys. Rev. Res. **5**, L042030 (2023)

Effective rotations opposing the direction of drift

In Subsec. 5.5.3 (see Fig. 5.5.2e-f) we state that, and briefly explain why, rotations of the probability density function (quantified via covariance ellipses) during heating emerge in opposite directions. Mathematically opposing rotations can be seen from Eq. (5.5.4), $\mathbf{\Sigma}_i^w(t) = \mathbf{\Sigma}_{\mathrm{s},w} + \mathrm{e}^{-\mathbf{A}t}\left[\mathbf{\Sigma}_{\mathrm{s},i} - \mathbf{\Sigma}_{\mathrm{s},w}\right]\mathrm{e}^{-\mathbf{A}^T t}$, where $\mathbf{\Sigma}_{\mathrm{s},i} - \mathbf{\Sigma}_{\mathrm{s},w} = (T_i/T_w - 1)\mathbf{\Sigma}_{\mathrm{s},w}$ has opposing signs for $T_c < T_w$ and $T_h > T_w$. However, the physical or phenomenological understanding is more challenging. The most surprising aspect is that rotational motions occur in directions that oppose the rotational driving (e.g., during heating). Since there are no rotational motions in the absence of driving, clockwise rotational driving can in fact lead to counterclockwise (effective) rotations.

First, note that the phenomenon has to be understood on the level of probability density functions and *not* on the level of individual particles' trajectories. In Fig. 5.5.4a,b we show that the cloud of particles' positions (representing the probability density) effectively rotates in the counterclockwise direction while the individual particles on average follow the rotational drift in the clockwise direction, see Fig. 5.5.4c.

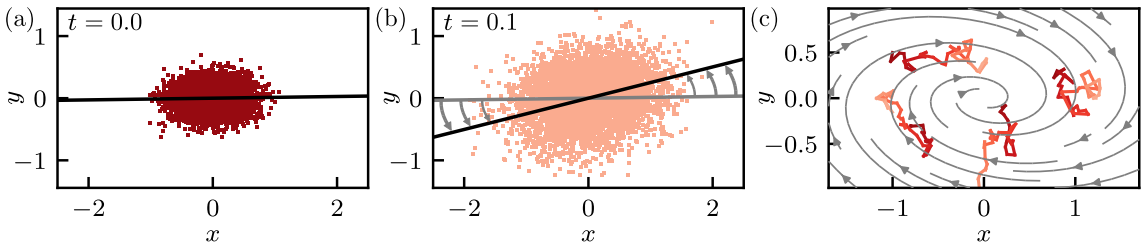

Figure 5.5.4: (a,b) Simulation of 5000 particles' trajectories evolving according to the two-dimensional overdamped Langevin equation $\mathrm{d}\mathbf{x}_t = -\mathbf{A}\mathbf{x}_t\mathrm{d}t + \sqrt{2}\mathrm{d}\mathbf{W}_t$ with $\mathbf{A} = \begin{bmatrix} r_1 & -r_2\omega \\ -r_1\omega & r_2 \end{bmatrix}$ with $r_1 = 1$, $r_2 = 3$, $\omega = 8\omega_c$ and time-step $\mathrm{d}t = 0.001$ starting from an initial condition corresponding to $T_c = 0.1$ (in units of T_w). (c) Simulated trajectories of 4 particles in time 0.0 (dark) to 0.1 (bright). Grey streamlines shows $\mathbf{A}(x, y)^T$, i.e., the direction that particles' trajectories follow on average.

The emergence of this counterintuitive opposing rotation is explained in Fig. 5.5.2f. To repeat this, during a Trotterized time-increment the diffusion propagates the initial covariance ellipse to a more circular (less eccentric) one. Next, note that the rotational drift is *not* a perfect circulation, but instead driving along elliptical contour-lines *plus* the driving into the center due to the confining (conservative) potential. This clockwise elliptical rotational driving applied to the ellipse (previously "rounded" during the diffusion Trotter-increment) leads to the counter-clockwise rotation directly by following the streamlines of the drift (see Fig. 5.5.2f).

Phys. Rev. Res. 5, L042030 (2023)

To elaborate on these rotations consider Fig. 5.5.5. Ellipses in Fig. 5.5.5 and those shown below are the covariance ellipses, while ellipses in Fig. 5.5.4 and Fig. 5.5.2f correspond to standard-deviation ellipses (i.e., square roots of covariance ellipses). In Fig. 5.5.5a we recall the opposite rotation during heating and cooling. As in Fig. 5.5.4 we then focus on the heating, where the initial rotation is in the counterintuitive direction; see Fig. 5.5.5b. To illustrate the explanation given in Fig. 5.5.2f, we show that this rotation similarly emerges if we start in a circular initial condition (see Fig. 5.5.5c).

If we instead consider a circular driving with circular steady-state density (see Fig. 5.5.5d) all rotations emerge in the (intuitive) clockwise direction, which shows that the elliptical (i.e., non-circular) component of the circular driving is a key factor in this phenomenon.

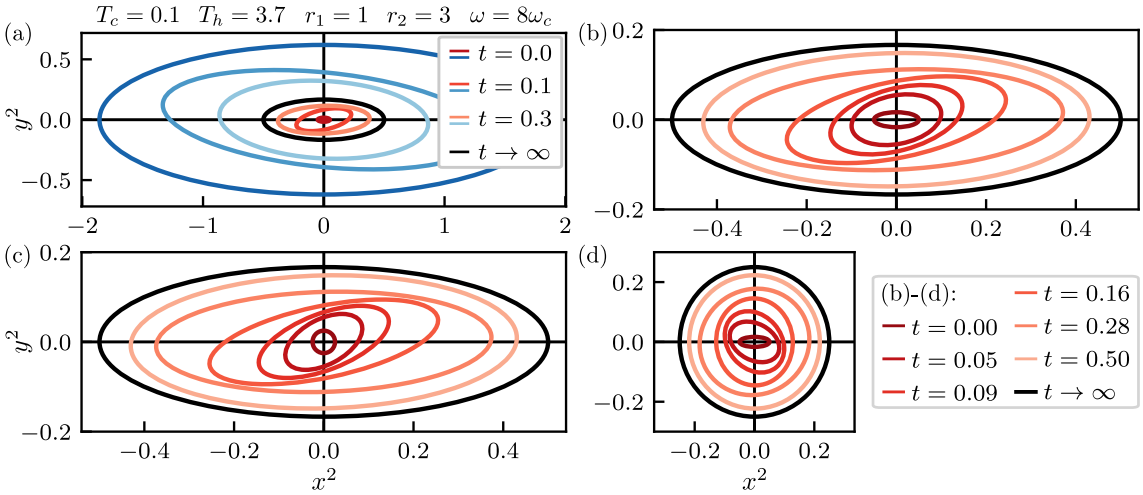

Figure 5.5.5: (a,b) Covariance ellipses from Eq. (5.5.26) for heating and cooling for the process as in Fig. 5.5.4 and Fig. 5.5.2. (c) As in (b) but with initial condition with $r_1 = r_2 = 2$. (d) As in (b) but with process (but not initial condition) defined with $r_1 = r_2 = 2$.

In Fig. 5.5.6 we further illustrate the relation between the direction of rotation and the shape of covariance ellipses. As explained above and in Fig. 5.5.2f, a more circular (less eccentric) ellipse [sign(3−ratio)=1] leads to a surprising counter-clockwise rotation [sign(angle-change)=1]. The overlap of the curves in Fig. 5.5.6c,d corroborates this explanation. Small deviations between the curves emerge, since the heuristic explanation only applies to ellipses with angle$(t) = 0$.

In Fig. 5.5.7 we repeat the presentation of Fig. 5.5.6 for a case where the eigenvalues of the drift matrix are real (see Fig. 5.5.2d for $\omega < \omega_c$). We observe that (opposite) rotational motions also occur for the case of real eigenvalues, which illustrates that (effective) rotational motions do *not only* emerge for complex eigenvalues. A difference with respect

Phys. Rev. Res. **5**, L042030 (2023)

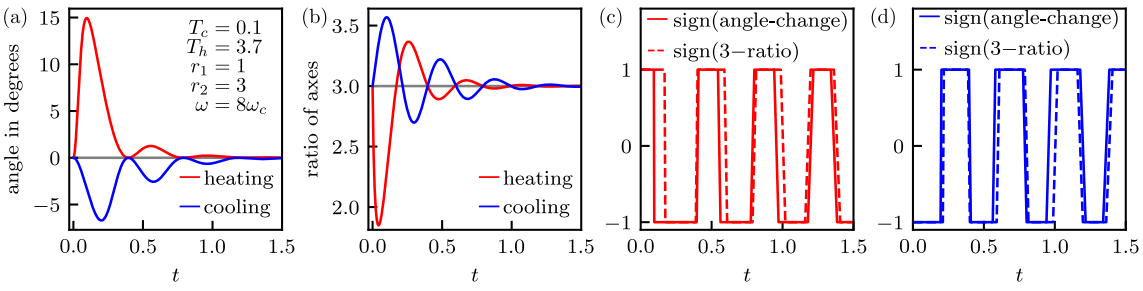

Figure 5.5.6: (a) Same as in 5.5.2e. (b) Ratio of the axes of the covariance ellipse. Values below 3 reflect more circular (less eccentric) ellipses compared to the initial condition and the steady state. (c,d) Direction of rotation (clockwise rotation is $+1$) and indicator of shape (± 1 means more/less round, i.e., less/more eccentric) for heating (c) and cooling (d).

to Fig. 5.5.6 is that the angles do not cross 0, such that the explanation for the overlaps in Fig. 5.5.7c,d only applies at $t = 0$.

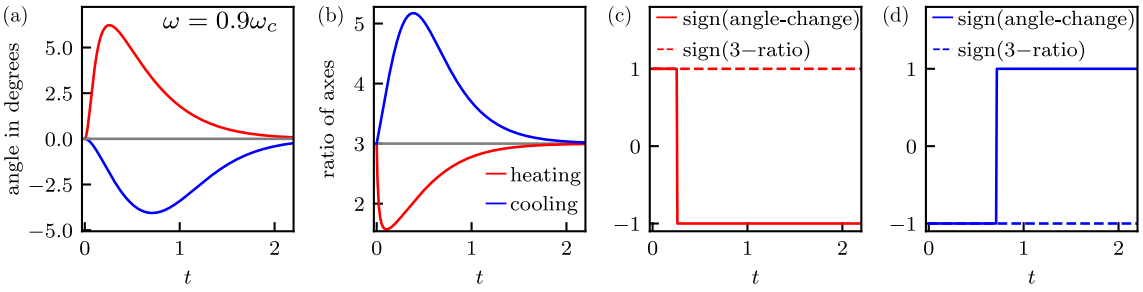

Figure 5.5.7: As in Fig. 5.5.6 but for $\omega = 0.9\omega_c$, i.e., eigenvalues of the drift matrix $\mathbf{A}$ are real (see Fig. 5.5.2d).

Relevance and generality of the observation of counterintuitive rotations:
So far, we only investigated the origin of the counterintuitive rotations in the two-dimensional example. However, since such counterintuitive rotations already occur in this linear, low-dimensional example, it is to be expected that such motions also occur for more general driven systems. In particular, if two-dimensional subspaces are described by the example above, one immediately has this rotation in the subspace of the more general dynamics. Generally, one expects opposite rotations during heating and cooling (and therefore one of the two has to rotate opposite to the driving) due to the difference in sign of $\mathbf{\Sigma}_{\mathrm{s},i} - \mathbf{\Sigma}_{\mathrm{s},w} = (T_i/T_w - 1)\mathbf{\Sigma}_{\mathrm{s},w}$ for $i = h, c$ in Eq. (5.5.26) [Eq. (5.5.4)] as pointed out above.

The relevance of this observation is twofold. On the one hand, it further emphasizes the asymmetry between heating and cooling, and that the process does *not* pass through locally equilibrated states (i.e., the system *cannot* be described by a time-dependent temperature).

Phys. Rev. Res. **5**, L042030 (2023)

On the other hand, it is also relevant for general relaxation phenomena, i.e., beyond thermal relaxation. For example, imagine one observes the part of the relaxation process in Fig. 5.5.4a,b for $t \in [0, 0.1]$. If one only observes the apparent counterclockwise rotation of the probability density, one would never guess that the underlying driving is actually in the clockwise direction. Therefore, awareness of this counterintuitive phenomenon might prove useful to avoid false conclusions; and a deep understanding of this phenomenon helps to arrive at correct conclusions.

Consistent comparison of equilibrium and non-equilibrium steady states

We here discuss under which circumstances we consider a comparison of equilibrium (EQ) and non-equilibrium steady states (NESS), or of different NESS, to be *consistent*.

In short, we consider a comparison to be consistent if tuning the driving strength does *not* change the steady-state density. Before we explain this in detail, we want to stress that a consistent comparison *is by no means required for the statement of the thermal relaxation asymmetry to be valid*, since this statement is proven for *any* NESS with linear drift in Subsec. 5.5.5. Therefore, we were able to choose the physical example of a Rouse chain in a shear flow to illustrate the relaxation asymmetry in Fig. 5.5.1 (which in fact does *not* represent a consistent comparison). The consistent comparison is, however, necessary for the statement of "accelerated relaxation" since this statement compares the relaxation speed towards an NESS with the relaxation speed in the *corresponding* passive system relaxing into an equilibrium steady state.

In Subsec. 5.5.3, we use Eq. (5.5.3), i.e., Eq. (5.5.25), to obtain the decomposition $\mathbf{A} = (\mathbf{D}_i + \boldsymbol{\alpha}_i)\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}$, see Eq. (5.5.7), with $\boldsymbol{\alpha}_i^T = -\boldsymbol{\alpha}_i$ for the linear drift matrix $\mathbf{A}$. Note that here $\mathbf{D}_i, \boldsymbol{\alpha}_i, \boldsymbol{\Sigma}_{\mathrm{s},i} \propto T_i$ all increase linearly with temperature, but the product $\mathbf{A}$ involving $\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1} \propto T_i^{-1}$ is temperature independent. Any $\mathbf{A}$ from this decomposition fulfills Eq. (5.5.3), i.e., Eq. (5.5.25), with the given $\boldsymbol{\Sigma}_{\mathrm{s},i}$, and in turn any $\mathbf{A}$ implying a steady-state covariance $\boldsymbol{\Sigma}_{\mathrm{s},i}$ via Eq. (5.5.3) can be decomposed with this $\boldsymbol{\Sigma}_{\mathrm{s},i}$ according to $\mathbf{A} = (\mathbf{D}_i + \boldsymbol{\alpha}_i)\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}$. The advantage of the latter form is that it allows to systematically compare NESS dynamics (or in the special case reversible dynamics) $\mathrm{d}\mathbf{x}_t = -\mathbf{A}\mathbf{x}_t \mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{W}_t$ with different $\mathbf{A}$ that possess different driving strengths but the same steady-state density. This comparison is performed by tuning the parameter α_i (reversible systems are obtained by setting $\alpha_i = \mathbf{0}$) for a given $\boldsymbol{\Sigma}_{\mathrm{s},i}$, which then yields $\mathbf{A}$ via $\mathbf{A} = (\mathbf{D}_i + \boldsymbol{\alpha}_i)\boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}$, see Eq. (5.5.7). We consider such a comparison to be *consistent*, in contrast to a comparison where tuning the irreversible driving alters $\boldsymbol{\Sigma}_{\mathrm{s},i}$ and thus the steady-state density.

Phys. Rev. Res. **5**, L042030 (2023)

An example of a driving that does *not* yield a consistent comparison is the shear flow in Fig. 5.5.1 and in Eqs. (5.5.14)-(5.5.22). We discuss this comparison in detail now. For simplicity, we consider a single particle $N = 1$ in the x-y plane subject to the confining potential and shear flow [see Eqs. (5.5.18)-(5.5.20)] described by the equation of motion $\mathrm{d}\mathbf{x}_t = -\mathbf{A}\mathbf{x}_t\mathrm{d}t + \sqrt{2}\mathrm{d}\mathbf{W}_t$ with drift matrix

$$\mathbf{A} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} r_x & 0 \\ 0 & r_y \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} + \begin{bmatrix} 0 & \omega \\ 0 & 0 \end{bmatrix}. \tag{5.5.29}$$

This drift originates from a (rotated) confining potential with confinement strength quantified by $r_x, r_y > 0$, plus a shear flow of strength ω (both exactly as shown in Fig. 5.5.1a). The drift without the shear flow $\omega = 0$ is symmetric and therefore gives rise to reversible dynamics with steady-state covariance [see Eqs. (5.5.3) or (5.5.7) for $\mathbf{D} = \mathbb{1}$]

$$\boldsymbol{\Sigma}_{\mathrm{s}} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1/r_x & 0 \\ 0 & 1/r_y \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{5.5.30}$$

The shear flow $\omega \neq 0$ renders the dynamics irreversible. However, since now it is not of the form $\boldsymbol{\alpha}\boldsymbol{\Sigma}_{\mathrm{s}}^{-1}$ with $\boldsymbol{\alpha}^T = -\boldsymbol{\alpha}$ as in Eq. (5.5.7), the steady-state covariance for $\omega \neq 0$ will no longer be given by Eq. (5.5.30), i.e., the steady-state Lyapunov equation [see Eq. (5.5.3) or Eq. (5.5.25)] for $\mathbf{A}$ with $\omega \neq 0$ will give rise to another steady-state different from Eq. (5.5.30) which corresponds to $\omega = 0$. Therefore, comparing systems with different ω will generally not be *consistent* [opposed a comparing systems with different $\boldsymbol{\alpha}_i$ in Eq. (5.5.7)].

We illustrate this *inconsistent* comparison by three different examples. Choosing the parameters $r_x = 1,\ r_y = 0.1,\ \omega = 3, \theta = -10°$ as in Fig. 5.5.1a, the eigenvalues of $\mathbf{A}$ are $0.55 \pm 0.51\mathrm{i}$, i.e., compared to $\omega = 0$ with eigenvalues $r_{x,y}$ the statement of faster relaxation as quantified in Eq. (5.5.8) does still hold true, even though the proof does not apply here (see also Fig. 5.5.1b where the curves with the shear flow decay faster at long times). However, if one instead takes $\omega = 0.5, \theta = 10°$ the eigenvalues of $\mathbf{A}$ are 1.08 and 0.02 i.e., the limiting relaxation is *slower* compared to the reversible system since $0.02 < r_{x,y}$. Thus, the statement of faster relaxation does not apply since the effect of the shear flow on the steady state is too large. Even more extreme is the case $\omega = 3, \theta = 10°$ where the eigenvalues are 1.365 and -0.265 where the negative eigenvalue implies that the shear flow destroyed the confining potential in the sense that the resulting drift no longer corresponds to a confined process. *This means that this process no longer relaxes into an NESS.* This can, of course, not happen for a consistent comparison, since changing only $\boldsymbol{\alpha}_i$ in Eq. (5.5.7) does not change the confinement.

Phys. Rev. Res. 5, L042030 (2023)

Adiabatic entropy production

The adiabatic entropy production is the housekeeping heat divided by the reservoir temperature and is given by [188]

$$\dot{S}_a(t) = \int \mathrm{d}\mathbf{x} P(\mathbf{x},t) \mathbf{a}_{\mathrm{irr}}(\mathbf{x})^T \mathbf{D}_i^{-1} \mathbf{a}_{\mathrm{irr}}(\mathbf{x}), \tag{5.5.31}$$

where the irreversible drift in the linear case considered reads $\mathbf{a}_{\mathrm{irr}}(\mathbf{x}) = -\mathbf{A}_{\mathrm{irr}}\mathbf{x} = -\boldsymbol{\alpha}_i \boldsymbol{\Sigma}_{\mathrm{s},i}^{-1}\mathbf{x}$. Thus, we see that the adiabatic entropy production term scales linearly with $\boldsymbol{\alpha}_i^T \mathbf{D}_i^{-1} \boldsymbol{\alpha}_i$ as mentioned in Subsec. 5.5.4, i.e., it scales quadratically in the driving strength.

Long-time scaling of the Kullback-Leibler divergence

In terms of the eigenvalue λ_1 of $\mathbf{A}$ that has the smallest real part, we know that asymptotically for large t the magnitude of $\mathrm{e}^{-\mathbf{A}t}$ is determined by $\mathrm{e}^{-\Re(\lambda_1)t}$ [there may still be oscillations (see Fig. 5.5.2h) and if $\mathbf{A}$ is not diagonalizable there may also be terms $t^k \mathrm{e}^{-\Re(\lambda_1)t}$ with $k \in \mathbb{N}$ entering, which nonetheless are dominated by $\mathrm{e}^{-\Re(\lambda_1)t}$ for sufficiently large t]. Note that $\mathrm{e}^{-\mathbf{A}t} \sim \mathrm{e}^{-\Re(\lambda_1)t}$ implies, via Eq. (5.5.6), that $\mathbf{X}(t) \sim \mathrm{e}^{-2\Re(\lambda_1)t}$ for $t \to \infty$. Recall Eq. (5.5.5), i.e., $2\mathcal{D}_t^i = \mathrm{tr}[\delta\tilde{T}_i \mathbf{X}(t)] - \ln\det[\mathbb{1} + \delta\tilde{T}_i \mathbf{X}(t)]$. Considering $\delta\tilde{T}_i \mathbf{X}(t) = \mathrm{e}^{-2\Re(\lambda_1)t}\mathbf{M}$ for some matrix $\mathbf{M}$ for large enough t and using that around $\mathrm{e}^{-2\Re(\lambda_1)t} \to 0$, we have that $\det[\mathbb{1} + \mathrm{e}^{-2\Re(\lambda_1)t}\mathbf{M}] = 1 + \mathrm{tr}[\mathrm{e}^{-2\Re(\lambda_1)t}\mathbf{M}] + \mathcal{O}[\mathrm{e}^{-4\Re(\lambda_1)t}]$, and we obtain $\mathcal{D}_t^i = \mathcal{O}[\mathrm{e}^{-4\Re(\lambda_1)t}]$ as illustrated in Fig. 5.5.2h. This confirms that the limiting relaxation speed is dictated by $\Re(\lambda_1)$, i.e., by the smallest real part of eigenvalues of $\mathbf{A}$. In the reversible case we have $\Re(\lambda_1) = \mu_1$ with the notation in Subsec. 5.5.4, and $\mu_1 = r_1$ for the example considered in Fig. 5.5.2.

Note that we did not formally exclude the case that the order $\mathrm{e}^{-4\Re(\lambda_1)t}$ also vanishes; in this situation we would need to consider even higher orders. It is likely that this case can be generally excluded, however, since no results hinge on the specific scaling, we do not go into more detail here.

Effective stiffness

The effective stiffness $\hat{r}_j(\omega) \equiv -\ln(x_j^t)/2t$ [such that $x_j^t = \mathrm{e}^{-2\hat{r}_j(\omega)t}$] is defined as the stiffness of the confining potential of a reversible system that has the same thermal relaxation properties as the considered system, where x_j^t for $j = 1, \ldots, d$ are the eigenvalues of the matrix $\mathbf{X}(t) \equiv \mathrm{e}^{-\mathbf{A}t}\boldsymbol{\Sigma}_{\mathrm{s},w}\mathrm{e}^{-\mathbf{A}^T t}\boldsymbol{\Sigma}_{\mathrm{s},w}^{-1} = \mathrm{e}^{-\mathbf{A}t}\mathrm{e}^{-\mathbf{A}_{-\boldsymbol{\alpha}}t}$ [see Eqs. (5.5.6) and (5.5.10)]. In Fig. 5.5.3b we show $\hat{r}_j(\omega)$ for $j = 1, 2$ for the two-dimensional system as shown in Fig. 5.5.2

Phys. Rev. Res. **5**, L042030 (2023)

with driving strength ω. The eigenvalues $x^t_{1,2}$ at $T_w = 1$ (i.e., T_i are measured in units of T_w) for this example are computed from

$$\begin{aligned}
\mathbf{A} &\equiv \begin{bmatrix} r_1 & -r_2\omega \\ r_1\omega & r_2 \end{bmatrix}, \qquad \boldsymbol{\sigma} = \sqrt{2}\mathbb{1}, \qquad \boldsymbol{\Sigma}_{\mathrm{s}} = \begin{bmatrix} 1/r_1 & 0 \\ 0 & 1/r_2 \end{bmatrix}, \\
M &\equiv \sqrt{(r_1 - r_2)^2 - 4r_1 r_2 \omega^2} \in \mathbb{C}, \\
\exp(-\mathbf{A}t) &= \frac{\exp[-(r_1+r_2)t/2]}{M} \times \\
&\quad \begin{bmatrix} M\cosh\left(\frac{Mt}{2}\right) - |r_1 - r_2|\sinh\left(\frac{Mt}{2}\right) & 2\omega r_2 \sinh\left(\frac{Mt}{2}\right) \\ -2\omega r_1 \sinh\left(\frac{Mt}{2}\right) & M\cosh\left(\frac{Mt}{2}\right) + |r_1 - r_2|\sinh\left(\frac{Mt}{2}\right) \end{bmatrix}, \\
x^t_1 x^t_2 &= \det[\mathbf{X}(t)] = \exp[-2(r_1+r_2)t], \\
x^t_1 + x^t_2 &= \mathrm{tr}[\mathbf{X}(t)] = 2\exp[-(r_1+r_2)t]\left[1 + 2\frac{(r_1-r_2)^2}{M^2}\sinh^2\left(\frac{Mt}{2}\right)\right] \\
x^t_{1,2} &= \frac{\mathrm{tr}(\mathbf{X}(t))}{2} \pm \sqrt{\frac{\mathrm{tr}^2(\mathbf{X}(t))}{4} - \det(\mathbf{X}(t))}.
\end{aligned} \tag{5.5.32}$$

Log-norm inequality

In Subsec. 5.5.5 we use the log-norm inequality $||\exp(\mathbf{M}t)|| \leq \exp[\mu(\mathbf{M})t]$ [334] where the log norm is defined via the matrix norm $||\mathbf{M}|| \equiv \sup_{\mathbf{v}\in\mathbb{R}^d\setminus\mathbf{0}} ||\mathbf{M}\mathbf{v}||_2 / ||\mathbf{v}||_2$ (also known as operator norm) where $||\mathbf{v}||_2 = \sqrt{\mathbf{v}^T\mathbf{v}}$ as

$$\mu(\mathbf{M}) \equiv \lim_{h\to 0^+} \frac{||\mathbb{1} + h\mathbf{M}|| - 1}{h}. \tag{5.5.33}$$

Writing the matrix norm $||\mathbf{M}||$ in the form $||\mathbf{M}\mathbf{v}||_2 / ||\mathbf{v}||_2 = \sqrt{\mathbf{v}^T\mathbf{M}^T\mathbf{M}\mathbf{v}/\mathbf{v}^T\mathbf{v}}$ one sees that $||\mathbf{M}||$ is given by the square root of the largest eigenvalue of the symmetric matrix $\mathbf{M}^T\mathbf{M}$. Splitting $\mathbf{M} = \mathbf{M}_s + \mathbf{M}_a$ with $\mathbf{M}_s \equiv (\mathbf{M}+\mathbf{M}^T)/2 = \mathbf{M}_s^T$ and $\mathbf{M}_a \equiv (\mathbf{M}-\mathbf{M}^T)/2 = -\mathbf{M}_a^T$ we find that

$$\begin{aligned}
(\mathbb{1} + h\mathbf{M})^T(\mathbb{1} + h\mathbf{M}) &= (\mathbb{1} + h\mathbf{M}_s - h\mathbf{M}_a)(\mathbb{1} + h\mathbf{M}_s + h\mathbf{M}_a) = 1 + 2h\mathbf{M}_s + \mathcal{O}(h^2) \\
&= (\mathbb{1} + h\mathbf{M}_s)^T(\mathbb{1} + h\mathbf{M}_s) + \mathcal{O}(h^2).
\end{aligned} \tag{5.5.34}$$

This implies that $||\mathbb{1} + h\mathbf{M}|| = ||\mathbb{1} + h\mathbf{M}_s||$ and via Eq. (5.5.33) that the log norm is solely determined by the symmetric part $\mu(\mathbf{M}) = \mu(\mathbf{M}_s)$. Intuitively, this states that asymmetric contributions (which account for rotations after exponentiation) do not enter the absolute value in the exponential bound $||\exp(\mathbf{M}t)|| \leq \exp[\mu(\mathbf{M})t]$, which makes the log norm very useful for our theory.

Phys. Rev. Res. 5, L042030 (2023)

From this insight, we immediately compute the result used in Subsec. 5.5.5, i.e., we use that $-\boldsymbol{\beta}\mathbf{D}_w\boldsymbol{\beta}$ with eigenvalues $-\mu_1 > -\mu_2 > \dots$ is the symmetric part of $-\boldsymbol{\beta}\mathbf{A}\boldsymbol{\beta}^{-1}$ to obtain

$$\mu(-\widetilde{\mathbf{A}}) \equiv \mu(-\boldsymbol{\beta}\mathbf{A}\boldsymbol{\beta}^{-1}) = \mu(-\boldsymbol{\beta}\mathbf{D}_w\boldsymbol{\beta}) = -\mu_1. \tag{5.5.35}$$

Phys. Rev. Res. **5**, L042030 (2023)

Chapter 6

Scattering fingerprints of two-state dynamics

We now come to the last chapter before the Conclusion. This chapter goes beyond the previous content in the sense that it no longer deals with Markov dynamics described by an overdamped Langevin equation. We introduce the considered dynamics in Sec. 6.1. In Sec. 6.2, we summarize the results of a publication that is reproduced in Sec. 6.3. Since this publication is rather self-contained and requires little previous knowledge, we only give a short introduction, focusing on putting the work into the context of the thesis.

6.1 Introduction to two-state dynamics and anomalous diffusion

In this section, we introduce *two-state dynamics* in Subsec. 6.1.1, *anomalous diffusion* in Subsec. 6.1.2, and the *intermediate scattering function* in Subsec. 6.1.3.

6.1.1 Two-state dynamics

In this chapter, we consider a new type of dynamics, namely *two-state dynamics*. To this end, we consider a process that randomly switches between two different states of motion, e.g., between diffusive motions with diffusion constants D_1 and D_2, respectively, or between two different conformations of a macromolecule. The switching is assumed to occur randomly, e.g., with waiting times drawn from an exponential distribution. Similar to the example of the different species in Subsec. 4.1.4, we assume that only the time-dependent

position is observed (or fingerprints of the position in related observables) without knowing at what time which state of motion determines the dynamics. In contrast to the example in Subsec. 4.1.4, the different motions are no longer independently coexisting, but instead stochastically convert between each other.

Often, we consider both, the dynamics in the individual states and the switching between them, to be Markovian, i.e., memoryless, see Subsec. 2.3.1. However, not resolving for the state of motion, but only for the time-dependent position, may generally give rise to memory if internal degrees of freedom evolve differently in the different state. This further distinguishes this chapter from the preceding parts of the thesis, where we almost exclusively considered Makovian dynamics.

6.1.2 Anomalous diffusion

In one part of the publication in Sec. 6.3, see Fig. 6.3.2, we go beyond the previously considered Langevin dynamics even in the individual states of motion. Instead, the individual states are assumed to follow a non-Markovian, *anomalous diffusion* [81, 82], that is often found in transport in complex, viscoelastic, or heterogeneous media [82,203,335,336]. For a one-dimensional position coordinate x_t, anomalous diffusion is characterized by a mean-squared displacement $\langle (x_t - x_0)^2 \rangle = 2Ct^\alpha$ with exponent $\alpha > 0$, where C is a generalized diffusion coefficient with units of $\mathrm{m^2 s^{-\alpha}}$. Processes with $\alpha > 1$ are called *superdiffusion*, while processes with $\alpha < 1$ are called *subdiffusion*. There are different mathematical models for anomalous diffusion. We here consider *fractional Brownian motion*, which is the Gaussian process, see Subsec. 2.3.2, fulfilling the scaling relation $\langle (x_t - x_0)^2 \rangle = 2Ct^\alpha$ for self-similar and stationary increments (for details, see [337, Proposition 3.8]).

Recall from the assumptions necessary for the emergence of Langevin dynamics in Subsec. 1.2 that we assumed that the medium relaxes much faster than the system. Though this is a reasonable assumption for Brownian particles in water, we run into problems if we consider more complex media, such as the interior of a biological cell. Though the required timescale separation may hold between a considered particle and water, the cytosol (cell fluid) contains a high concentration of particles such as proteins that are much larger than water molecules. Therefore, memory effects emerge since the medium does not relax fast. Often, this complex medium slows down the diffusion and leads to anti-persistent, subdiffusive motion.

In particular, it was recently found that tracer particle motion in living cells can be described by a two-state dynamics switching between two different subdiffusive motions [91]. We revisit this particular process in Fig. 6.3.2 in the publication reproduced in Sec. 6.3.

6.1.3 Considered observable

As in the previous chapter, the question arises how to appropriately quantify the considered process. An observable that is both, mathematically and experimentally, appropriate for the dynamics considered here is the *intermediate scattering function* $F(\mathbf{q},t)$ that is the moment generating function, see Eq. (2.2.4), of the time-dependent displacement vector

$$F(\mathbf{q},t) \equiv \left\langle \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot(\mathbf{x}_t-\mathbf{x}_0)} \right\rangle . \tag{6.1.1}$$

Mathematically, this observable has the advantage that (similar to the Feynman-Kac approach in Subsec. 3.2.2) the moment generating function is more easily accessible than $p(\mathbf{x},t)$. From an experimental point of view, note that $F(\mathbf{q},t)$ is directly probed in modern experimental techniques such as neutron spin echo spectroscopy [338–341], differential dynamic microscopy [342–344], or Fourier imaging correlation spectroscopy [345–347]. Moreover, it is directly related to the *dynamic structure factor* $S(\mathbf{q},\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} F(\mathbf{q},|t|)/2\pi$, which is measured in inelastic scattering experiments [348] (giving rise to the name "intermediate scattering function"). Alternatively, the intermediate scattering function in Eq. (6.1.1) may also be obtained from measured trajectories, as shown in Fig. 6.3.2 for particle tracking data recorded in living cells [91].

6.2 Summary of results

The publication in Sec. 6.3 investigates the intermediate scattering function and dynamic structure factor for two-state dynamics. We solve and discuss simple model systems, including a minimal model for protein (un)folding. We fit some results to experimental data, and thereby support the hypothesis from Ref. [91] of two-state dynamics switching between anomalous diffusions.

6.3 Scattering fingerprints of two-state dynamics

(New J. Phys. 2022)

This section is a slightly modified version of the publication C. Dieball, D. Krapf, M. Weiss and A. Godec, "Scattering fingerprints of two-state dynamics", *New J. Phys.* **24**, 023004 (2022).

My personal contribution as first author of the publication was in performing calculations and simulations, analyzing the results, and co-writing the manuscript.

Scattering fingerprints of two-state dynamics

Cai Dieball[1], Diego Krapf[2], Matthias Weiss[3] and Aljaž Godec[1]

[1]Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen, Germany

[2]Electrical and Computer Engineering and School of Biomedical Engineering, Colorado State University, Fort Collins, Colorado, USA

[3]Experimental Physics I, University of Bayreuth, Bayreuth, Germany

New J. Phys. **24** 023004 (2022)

Abstract

Particle transport in complex environments such as the interior of living cells is often (transiently) non-Fickian or anomalous, that is, it deviates from the laws of Brownian motion. Such anomalies may be the result of small-scale spatio-temporal heterogeneities in, or viscoelastic properties of, the medium, molecular crowding, etc. Often the observed dynamics displays multi-state characteristics, i.e., distinct modes of transport dynamically interconverting between each other in a stochastic manner. Reliably distinguishing between single- and multi-state dynamics is challenging and requires a combination of distinct approaches. To complement the existing methods relying on the analysis of the particle's mean squared displacement, position- or displacement-autocorrelation function, and propagators, we here focus on "scattering fingerprints" of multi-state dynamics. We develop a theoretical framework for two-state scattering signatures – the intermediate scattering function and dynamic structure factor – and apply it to the analysis of simple model systems as well as particle-tracking experiments in living cells. We consider inert tracer-particle motion as well as systems with an internal structure and dynamics. Our results may generally be relevant for the interpretation of state-of-the-art differential dynamic microscopy experiments on complex particulate systems, as well as inelastic or quasielastic neutron (including spin-echo) and X-ray scattering scattering probing structural and dynamical properties of macromolecules, when the underlying dynamics displays two-state transport.

6.3.1 Introduction

Transport in complex systems, such as disordered media [349], biological cells [82, 335], and cell membranes [350, 351], frequently deviates from the laws of Brownian motion that govern the dynamics of particles in simple media and at high dilution. The reason

for the deviation from the paradigmatic Brownian behavior may be sought in spatial obstruction and/or macromolecular crowding [82, 335, 351], viscoelastic properties of the medium [352–356], and spatial heterogeneities [357, 358], to name but a few. Over the years a multitude of approaches have emerged to characterize anomalous diffusion from different perspectives, and in particular to distinguish between different modes of anomalous diffusion (for excellent reviews see, e.g., [82, 203, 350, 359, 360]).

An important class of anomalous diffusion processes are multi-state dynamics – dynamics that depend on the instantaneous value of an internal state. Examples of multi-state dynamics include stochastically gated diffusion-controlled reactions [361–365], intermittent molecular search processes underlying cellular gene regulation [366–368], or the center of mass diffusion of (un)folding proteins [369] or growing polymers [370], the anomalous diffusion of membrane proteins [371], as well as tracer diffusion in heterogeneous environments [372, 373], and most recently also in mammalian cells [91, 202].

Notwithstanding all progress in the development of analytical tools [82, 203, 335, 360, 374] it often remains challenging to conclusively distinguish between different modes of motion observed in experiments [375], and in particular to conclusively identify multi-state transport [376–378]. To this end, we focus on scattering fingerprints of multi-state dynamics, that is, on the intermediate scattering function (ISF) and dynamic structure factor (DSF), observables that are typically monitored in neutron and X-ray scattering experiments (for applications in the context of biophysical systems see, e.g., [379–382]). The ISF and the DSF were originally identified as key observables in scattering experiments by van Hove [348]. While the ISF was initially introduced as an auxiliary observable in the study of the DSF, more recent methods, such as neutron spin echo (NSE) spectroscopy [338–341], as well as modern (colloidal) particle-tracking techniques, such as differential dynamic microscopy (DDM) [342–344] or Fourier imaging correlation spectroscopy (FICS) [345–347], probe the ISF directly.

Motivated by recent experimental observations (see, e.g., [91, 202, 383]) our aim is to employ the ISF and DSF in the analysis of multi-state transport, more precisely, of two-state dynamics. By combining theory and particle-tracking experiments in living cells we consider inert tracer-particle dynamics as well as systems with internal degrees of freedom. Our results may be relevant for the interpretation of DDM experiments on complex particulate systems, as well as neutron and X-ray scattering scattering, when the underlying dynamics displays two-state transport.

New J. Phys. **24** 023004 (2022)

6.3.2 Theory

Intermediate scattering function and the dynamic structure factor

The fundamental quantities, the ISF and DSF, are calculated from the so-called van Hove function $G(\vec{r},t)$, which is a generalized pair distribution function [348, 384]. In the absence of quantum effects $G(\vec{r},t)$ is the density-density time-correlation function that may be interpreted as the average number of particles (scatterers) β in a region $d\vec{r}$ around a point $\vec{r}+\vec{r}'$ at time t given that there was a particle (scatterer) α at a point $\vec{r}'$ at time $t=0$ (whereby the initial condition $\vec{r}'$ is averaged out),

$$G^{\alpha\beta}(\vec{r},t) = \langle \delta(\vec{r}_\beta(t) - \vec{r}_\alpha(0) - \vec{r}) \rangle. \tag{6.3.1}$$

The brackets denote the average over an appropriate ensemble that we will specify below. For the sake of simplicity, we here focus exclusively on systems of non-interacting molecules, which, however, may possess internal degrees of freedom. Depending on whether $\alpha=\beta$ or not (i.e., whether we consider scattering events at an individual or two distinct scatterers) the van Hove function splits into an incoherent part (self-part, $\alpha=\beta$) that probes single-scatterer motions, and a coherent part (sum of distinct-part, $\alpha\neq\beta$, and self-part) probing collective motions [384]. More precisely,

$$G_{\text{inc}}(\vec{r},t) = \frac{1}{N}\sum_{\alpha=1}^{N} G^{\alpha\alpha}(\vec{r},t), \tag{6.3.2}$$

$$G_{\text{coh}}(\vec{r},t) = \frac{1}{N}\sum_{\alpha,\beta=1}^{N} G^{\alpha\beta}(\vec{r},t), \tag{6.3.3}$$

where the sum runs over all N scattering centers within a molecule and we will in addition average over an ensemble of statistically independent trajectories of the molecule (in fact over an ensemble of many such molecules) that we denote by $\langle\cdot\rangle$. Depending on the experimental setup $G_{\text{inc}}(\vec{r},t)$ and $G_{\text{coh}}(\vec{r},t)$ may [385] or may not be monitored separately.

The intermediate scattering function (ISF; mathematically the characteristic function of displacements[1]) measured in NSE, DDM and FICS is the spatial Fourier transform of $G^{\alpha\beta}(\vec{r},t)$,

$$F^{\alpha\beta}(\vec{q},t) = \int \mathrm{d}^3r\; G^{\alpha\beta}(\vec{r},t)\mathrm{e}^{-i\vec{q}\cdot\vec{r}} = \langle \mathrm{e}^{-i\vec{q}\cdot(\vec{r}_\beta(t)-\vec{r}_\alpha(0))}\rangle, \tag{6.3.4}$$

[1] If moments of the displacement are finite, the ISF as a function of q contains the complete information about the statistics of the displacements.

New J. Phys. 24 023004 (2022)

while the dynamic structure factor (DSF) that is measured in inelastic/quasielastic scattering experiments is the space-and-time Fourier transform of $G^{\alpha\beta}(\vec{r},t)$, i.e.,

$$S^{\alpha\beta}(\vec{q},\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}t\;\mathrm{e}^{i\omega t}F^{\alpha\beta}(\vec{q},|t|)=\frac{1}{\pi}\int_{0}^{\infty}\mathrm{d}t\;\cos(\omega t)F^{\alpha\beta}(\vec{q},t), \tag{6.3.5}$$

or alternatively defined via the Laplace transform $\widehat{f}(s)\equiv\int_0^\infty \mathrm{e}^{-st}f(t)\mathrm{d}t$ as $S^{\alpha\beta}(\vec{q},\omega)=\pi^{-1}\mathrm{Re}[\widehat{F^{\alpha\beta}}(\vec{q},s=-i\omega)]$, where Re denotes the real part.

In scattering experiments the arguments $\vec{q}$ and ω correspond to the momentum and energy transfer and the DSF is proportional to the measured intensity [348]. In the next step we develop results for $F^{\alpha\beta}(\vec{q},t)$ and $S^{\alpha\beta}(\vec{q},\omega)$ for systems displaying two-state dynamics.

Two-state Dynamics

Assuming two distinct dynamic "states" interconverting in continuous time in a Markovian manner with rates $k_1,k_2>0$, that is, according to the master equation $\dot{\vec{p}}(t)=\mathrm{K}\vec{p}(t)$ with transition rate matrix

$$\mathrm{K}=\begin{bmatrix}-k_1 & k_2\\ k_1 & -k_2\end{bmatrix}. \tag{6.3.6}$$

Using

$$p_1^{\mathrm{eq}}\equiv\frac{k_2}{k_1+k_2},\qquad p_2^{\mathrm{eq}}\equiv\frac{k_1}{k_1+k_2}, \tag{6.3.7}$$

the jump-propagator is diagonalized as

$$\exp(\mathrm{K}t)=\begin{pmatrix}p_1^{\mathrm{eq}}\\ p_2^{\mathrm{eq}}\end{pmatrix}(1,1)+\mathrm{e}^{-(k_1+k_2)t}\begin{pmatrix}1\\ -1\end{pmatrix}(p_2^{\mathrm{eq}},-p_1^{\mathrm{eq}})\,. \tag{6.3.8}$$

We now assume $N=1$ and drop the indices α,β; the case $N>1$ is discussed in Sec. 6.3.5. Let $G_j(\vec{r},t)\psi_j(t)$ denote the joint density to observe a displacement $\vec{r}$ in a time t in dynamical state j without changing the state, where $G_j(\vec{r},t)$ denotes the van Hove function for the dynamics in a single state j. Its Fourier-Laplace transform will be denoted by

$$\widehat{(F\psi)}_j\equiv\widehat{(F\psi)}_j(\vec{q},s)\equiv\int\mathrm{d}^3r\int_0^\infty\mathrm{d}t\,\mathrm{e}^{-st-i\vec{q}\cdot\vec{r}}G_j(\vec{r},t)\psi_j(t). \tag{6.3.9}$$

Often the change of state erases all memory in the sense that the propagation within each sojourn can be assumed to be independent of the configuration just before the switch [386], as in the case e.g., when the dynamics in both states is translation invariant. In this case

New J. Phys. **24** 023004 (2022)

the DSF and ISF of the complete two-state dynamics can be obtained following Ref. [386] by a direct summation over all possible realizations of sojourns until time t (for the derivation, see Sec. 6.3.5). The Fourier-Laplace transform of the complete two-state propagator, $G(\vec{r},t)$, in this case reads

$$\widehat{F}(\vec{q},s)=\frac{p_1\widehat{(F\psi)}_1\left[1+k_1\widehat{(F\psi)}_2\right]+p_2\widehat{(F\psi)}_2\left[1+k_2\widehat{(F\psi)}_1\right]}{1-k_1k_2\widehat{(F\psi)}_1\widehat{(F\psi)}_2}, \tag{6.3.10}$$

where $p_{1,2}$ denote the initial occupation of states. The DSF henceforth follows immediately using

$$S(\vec{q},\omega)=\frac{1}{\pi}\Re[\widehat{F}(\vec{q},-i\omega)]. \tag{6.3.11}$$

The problem of determining the ISF and DSF for two-state dynamics with "complete memory erasure" upon each change of state thus boils down to determining $\widehat{(F\psi)}_{1,2}(\vec{q},s)$ in Eq. (6.3.9) for the specific example under consideration. In case there is no complete memory erasure (see protein (un)folding example Sec. 6.3.3) one must solve the specific model at hand using a dedicated method or resort to approximations. In the following, we address a set of insightful and experimentally relevant scenarios of two-state diffusion.

New J. Phys. 24 023004 (2022)

6.3.3 Results

Two-state diffusion

In the first example, we assume that the molecule/tracer particle undergoes free Brownian motion with a diffusion coefficient that randomly switches between values D_1 and D_2 in a Markovian fashion with rates k_1 and k_2, implying the distribution $\psi_j(t)=\mathrm{e}^{-k_jt}$ of sojourn times in the respective states. This situation arises in the context of tracer diffusion in heterogeneous media (see, e.g., [365, 372, 373]). Note that assuming sojourn times to be independent of the dynamics within the individual states is not always justified. In particular, we here assume an "annealed heterogeneity" (see, e.g., [358]). This may occur when the medium is translationally invariant, such as in the case of a state change triggered by the reversible binding to a homogeneously distributed mobile species. In general, this does not apply to quenched heterogeneous media, except in cases where a system is observed under equilibrated conditions [387].

The Fourier-Laplace transform of $G_j(\vec{r},t)\psi_j(t)$ defined in Eq. (6.3.9) reads

$$\widehat{(F\psi)}_j(\vec{q},s)=\widehat{(F\psi)}_j(q,s)=\frac{1}{s+q^2D_j+k_j}, \tag{6.3.12}$$

where $q^2 \equiv \vec{q}\cdot\vec{q}$. Plugging Eq. (6.3.12) into Eq. (6.3.10) the Fourier-Laplace image of the complete propagator for two-state diffusion can be written as

$$\widehat{F}(q,s) = \frac{s + q^2(p_1 D_2 + p_2 D_1) + k_1 + k_2}{(s+q^2 D_1 + k_1)(s + q^2 D_2 + k_2) - k_1 k_2}, \tag{6.3.13}$$

with a general initial condition $p_{1,2}$. Defining

$$\mu_\pm(q) = \frac{q^2(D_1+D_2)+k_1+k_2}{2} \pm \sqrt{\frac{[q^2(D_1-D_2)+k_1-k_2]^2}{4} + k_1 k_2}\,, \tag{6.3.14}$$

introducing the auxiliary function

$$\varphi(q) \equiv \frac{q^2(p_1 D_2 + p_2 D_1) + k_1 + k_2 - \mu_-(q)}{\mu_+(q) - \mu_-(q)}, \tag{6.3.15}$$

and inverting into the time domain we finally obtain

$$F(q,t) = [1-\varphi(q)]\mathrm{e}^{-\mu_+(q)t} + \varphi(q)\mathrm{e}^{-\mu_-(q)t}, \tag{6.3.16}$$

$$S(q,\omega) = \frac{1}{\pi}\Re[\widehat{F}(q,-i\omega)] = \frac{1}{\pi}\frac{[1-\varphi(q)]\mu_+(q)}{\omega^2+\mu_+^2(q)} + \frac{1}{\pi}\frac{\varphi(q)\mu_-(q)}{\omega^2+\mu_-^2(q)}. \tag{6.3.17}$$

In contrast, in the case when the two states co-evolve as a "frozen mixture"—an ensemble consisting of two types of molecules that do not interconvert between each other[2]—with the state occupations $p_{1,2}$ we have $F_{\mathrm{mix}}(q,t) = p_1\mathrm{e}^{-q^2 D_1 t} + p_2\mathrm{e}^{-q^2 D_2 t}$. This frozen mixture will throughout be denoted by the subscript "mix", and is the limit of Eq. (6.3.16) in the case of slow switching rates $k_1 + k_2 \ll q^2 D_{1,2}$, see Sec. 6.3.5.

It is also instructive to inspect the fast-switching limit, $k_1 + k_2 \gg q^2 D_{1,2}$, where we find (for details see Sec. 6.3.5) $\mu_+(q) \simeq q^2 D_{\mathrm{eff}} \equiv p_1^{\mathrm{eq}} D_1 + p_2^{\mathrm{eq}} D_2$ and $\mu_-(q) \simeq k_1 + k_2 \gg q^2 D_{1,2}$, which implies simple diffusion with an effective diffusion coefficient in agreement with intuition.

Our aim here is to demonstrate that the analysis allows to distinguish between single-state dynamics, two-state switching diffusion and the "frozen mixture" with the same stationary probabilities as long as the switching rates are not too fast or too slow, which would give rise to the aforementioned (fast-switching) effective diffusion and "frozen mixture", respectively.

For such moderate switching rates, the ISF shown in Fig. 6.3.1a and the DSF in Fig. 6.3.1b indeed reveal distinctive characteristics of two-state diffusion (dashed lines) when com-

[2]This "frozen mixture" is the same as the non-interacting species considered in Subsec. 4.1.4.

New J. Phys. **24** 023004 (2022)

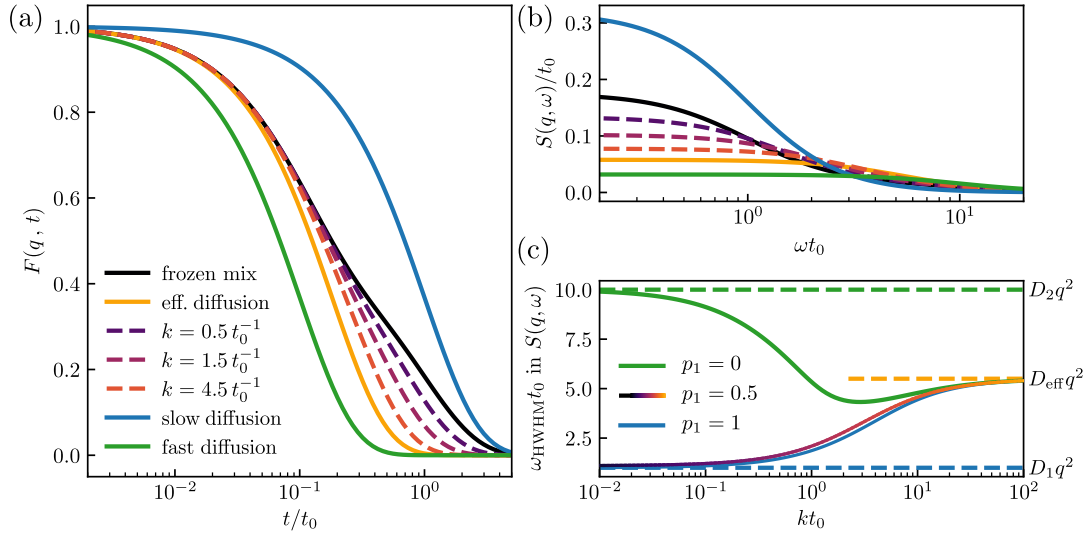

Figure 6.3.1: (a) ISF $F(q,t)$ for free diffusion switching between diffusion constants D_1 and $D_2 = 10D_1$ with switching rates $k_1 = k_2 = k$ evolving from an initial equilibrium occupation $p_1^{\mathrm{eq}} = p_2^{\mathrm{eq}} = 0.5$ and time t measured in units of $t_0 = 1/D_1q^2$. The ISF for different switching rates are shown along with the limits of the frozen mix ($k \to 0$, black curve) and the effective diffusion ($k \to \infty$, yellow curve). (b) Dynamical structure factor corresponding to the ISF in (a). (c) Half width at half maximum (HWHM) of the DSF from (b) are shown along with the HWHMs of the blue, green and yellow curves in (b) (dashed lines). To compare with a non-equilibrium initial occupation, the HWHM for initial occupation $p_1 = 0$ and $p_1 = 1$ are also shown (green and blue lines, respectively).

New J. Phys. **24** 023004 (2022)

pared to the "frozen mixture" (black full line) as well as the two individual diffusive states (blue and green full lines, respectively). To reduce these differences to a single observable we inspect the half width at half the maximum (HWHM) of the quasielastic peak as a function of the switching rate (that for simplicity we set to be symmetric, $k_1 = k_2 \equiv k$) for different initial conditions. The HWHM is determined from Eq. (6.3.17) by setting $S(q, \omega_{\mathrm{HWHM}}) = S(q, 0)/2$ and solving the corresponding bi-quadratic equation for ω_{HWHM}. The results for ω_{HWHM} are shown in Fig. 6.3.1c and together with Fig. 6.3.1a-b confirm that the DSF and ISF can reliably distinguish between a "non-communicating", frozen superposition of dynamic modes and two-state diffusion as long as the switching rates are not too fast or too slow.

Note that for an isotropic Gaussian process we have $F(q,t) = \mathrm{e}^{-q^2\langle[\vec{r}(t)-\vec{r}(0)]^2\rangle/2d}$ and hence ISF contains exactly the same information as the mean squared displacement (MSD) $\langle[\vec{r}(t) - \vec{r}(0)]^2\rangle$. For the individual states as well as for effective diffusion, the dynamics is Gaussian. However, two-state dynamics is typically non-Gaussian [91] and the ISF provides more information about the dynamics. Note, moreover, that the MSD of the two-state process reflects (normal) diffusion with effective diffusivity D_{eff} for all switching rates and thus alone cannot reveal underlying two-state dynamics. In the next example, we turn to the switching between two subdiffusive states.

Two-state Fractional Brownian Motion

Tracer transport in complex, dynamic and/or heterogeneous media such as living cells is often found to be anomalous [82, 203, 335, 336]. Particularly interesting are situations with stochastically interconverting anomalous diffusion processes such as the two-state antipersistent (i.e., subdiffusive) fractional Brownian motion observed in recent experiments on mammalian cells [91, 202]. To be more precise, a systematic analysis of the motion of quantum dots immersed in the cytoplasm of living mammalian cells revealed two-state subdiffusive fractional Brownian motion with exponent $\alpha < 1$ but with a stochastically switching anomalous diffusion coefficient $C_{1,2}$ [91]. Here we determine the DSF and ISF for the process as complementary observables and apply them to the analysis of data obtained in [91] focusing on latrunculin-treated cells. Untreated cells had been seen to feature the same dynamics [91]. Note that here a Markov switching between the two states coexists with non-Markovian subdiffusion in each of the two states, respectively, with a mean squared displacement $\langle[\vec{r}(t) - \vec{r}(0)]^2\rangle_j = 2dC_j t^{\alpha_j}$ where d (here equal two) is the dimensionality of the system, and $\langle\cdot\rangle_j$ denotes the average over an ensemble of fractional Brownian motions (FBMs) with exponent α_j and coefficient C_j. Thus, there is an inherent ambiguity in how to treat memory upon each jump. Here, in contrast to the original

New J. Phys. **24** 023004 (2022)

publication [91], we assume for convenience that the memory in $\vec{r}(t)$ is erased upon each change of state, which allows us to use the result in Eq. (6.3.10). Notably, this choice does not affect the inferred two-state FBM, which further substantiates the robustness of the proposed model.

We first treat the problem theoretically. The d-dimensional fractional Brownian motion is an isotropic and translationally invariant Gaussian process. Thus, the ISF of FBM is the characteristic function of a Gaussian process $F_j(q,t) = \exp(-q^2\langle[\vec{r}(t)-\vec{r}(0)]^2\rangle_j/2d)$. Within a sojourn in state $j = 1, 2$, the characteristic function of the displacement $\vec{r}$ in a time t without changing the state is $F_j(q,t)\psi_j(t) = \exp(-[q^2C_jt^{\alpha_j} + k_j]t)$. Its Laplace transform $\widehat{(F\psi)}_j(q,s)$ may be written in terms of Meijer G-functions in Sec. 6.3.5 (see Eq. (6.3.36)). Plugging the result for $\widehat{(F\psi)}_j(q,s)$ into Eq. (6.3.10) yields the Laplace-transformed ISF, which we invert numerically into the time domain. The DSF is in turn obtained fully analytically by setting $s = -i\omega$ and taking the real part as in Eq. (6.3.11).

The model contains six parameters $C_1, C_2, k_1, k_2, \alpha_1, \alpha_2$ to be determined. To minimize overfitting, we fix three parameters by considering the short-time behavior of the MSD, see dashed line in Fig. 6.3.2a. We see that the short time MSD (i.e., prior to any change of state) is well described by simple subdiffusion with $\alpha = 0.5$, and we thus fit at the first 10 seconds yielding $C_{\text{short}} = 0.0083\ \mu\text{m}^2/\sqrt{\text{s}}$. To reduce the number of free parameters, we constrain ourselves to exactly reproduce this short-time behavior at times sufficiently short to not be influenced by transitions, i.e., we fix $\alpha_1 = \alpha_2 = 0.5$ and $p_1^{\text{eq}}C_1 + p_2^{\text{eq}}C_2 = C_{\text{short}}$ (recall $p_1^{\text{eq}} = k_2/(k_1+k_2)$ from Eq. (6.3.7)). This leaves three free parameters, which we determine by a least-squares fit simultaneously at $q = 3\ \mu\text{m}^{-1}$ and $q = 7\ \mu\text{m}^{-1}$. The fitting yields $C_1 = 0.002\ \mu\text{m}^2/\sqrt{\text{s}}$, $C_2 = 0.052\ \mu\text{m}^2/\sqrt{\text{s}}$, $k_1 = 0.009\ \text{s}^{-1}$, $k_2 = 0.063\ \text{s}^{-1}$.

For the case $\alpha_j = 1/2$ the Fourier-Laplace transform of the joint density to observe a displacement $\vec{r}$ in a time t without changing the state reads (see Sec. 6.3.5)

$$\widehat{(F\psi)}_j(q,s) = \frac{1}{s+k_j} - \frac{\sqrt{\pi}q^2C_j}{2(s+k_j)^{3/2}}\text{erfcx}\left(\frac{q^2C_j}{2\sqrt{s+k_j}}\right), \tag{6.3.18}$$

where $\text{erfcx}(x) \equiv \exp(x^2)(1-\text{erf}(x))$ and erf denotes the error function. Via Eq. (6.3.10) and numerical Laplace inversion, this yields the two-state ISF. The theoretical ISF (blue lines) alongside the statistical uncertainty expected from a set of 400 realizations (shaded area) and the ISF determined directly from the experimental trajectories is shown in Fig. 6.3.2b for several values of q. The comparison between the corresponding mean squared displacements is shown in Fig. 6.3.2a. The theoretical fit and experimental results display a good agreement. In particular, the analysis of the ISF not only confirms the

New J. Phys. **24** 023004 (2022)

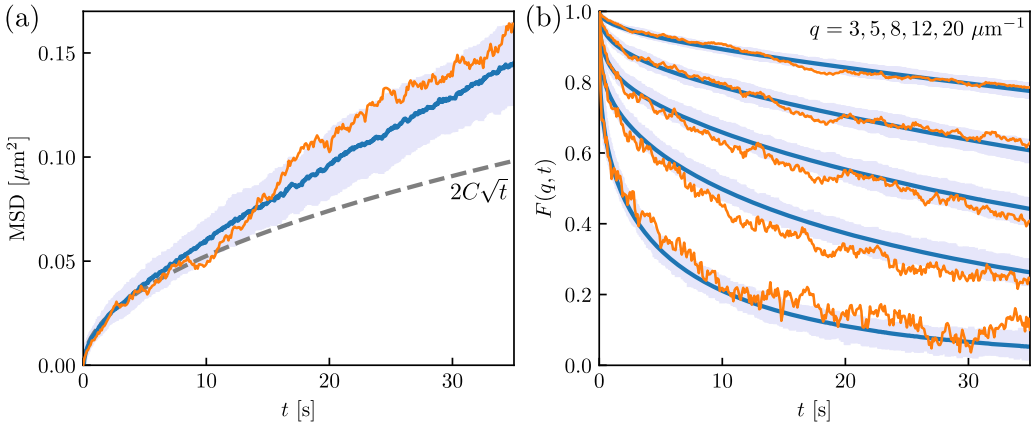

Figure 6.3.2: (a) Mean-squared displacement (MSD) derived from the experiment (orange curve) and MSD simulated with fit parameters from (b) (blue curve with blue shaded standard deviation). The gray dashed line shows a simple subdiffusive $\sqrt{t}$-fit to the first 10 seconds, yielding $C = 0.0083\ \mu\text{m}^2/\sqrt{\text{s}}$. (b) ISF $F(q,t)$ for (from top to bottom) $q = 3, 5, 8, 12, 20\ \mu\text{m}^{-1}$ from experiment (orange curve) and theory (blue curve) with standard deviation (blue shaded) obtained from simulations (experimental and simulation details are given in Sec. 6.3.5); the fitting procedure and parameters are given in the text; the range of q values was selected to capture the spatio-temporal scales in the relaxation observed experimentally, i.e., such that $F(q,t)$ spans the full range of values.

findings in Ref. [91] but also demonstrates the two-state FBM to be an appropriate model for the observed dynamics over a broad range of spatial and temporal scales. Notably, this spatio-temporally resolved information is inherently coarse-grained out in the analysis of the mean squared displacement. Recall that because the process is not Gaussian [91] the ISF and DSF, but not the MSD, contain the full information about the dynamics. The scattering fingerprints proposed here are thus well suited for a deeper analysis of particle-tracking experiments displaying multi-state anomalous diffusion.

Reversible dimerization with an internal mode

So far we have only considered two-state dynamics without any internal degrees of freedom. To go beyond this limitation, we consider diffusion in the presence of reversible dimerization (see Fig. 6.3.3a) captured in the mean field limit – two particles are assumed to associate with an effective rate that is independent of the particles' instantaneous position in the spirit of Smoluchowski [388, 389]. To be concrete, we consider non-interacting particles diffusing freely with a diffusion coefficient $D_1 = D$ and forming a dimer with a center-of-mass diffusion coefficient $D_2 = D/2$ and an internal harmonic vibrational relaxation mode with rate a (i.e., the internal coordinate – the distance between the associated particles – evolves as an Ornstein-Uhlenbeck process [102] or Rouse model with $N = 2$).

New J. Phys. **24** 023004 (2022)

Note that since $N > 1$ in contrast to the previous examples we here need to distinguish between coherent (i.e., Eq. (6.3.3)) and incoherent (i.e., Eq. (6.3.2)) contributions. We focus on the incoherent part (i.e., we set $\alpha = \beta$ in Eq. (6.3.1)) and note that the beads are equal and thus the beads $\alpha = 1$ and $\alpha = 2$ give an identical scattering contribution.

The association and dissociation rates are denoted by k_1 and k_2 respectively. Since the internal dynamics does not depend on the absolute position in space and each change of state erases all memory, we can employ the result in Sec. 6.3.2. In particular, this assumption neglects a possibly enhanced probability for re-association of the same pair of molecules immediately upon dissociation, which is expected to be a good approximation in well-mixed (i.e., "stirred") systems and/or at high monomer concentrations. The opposite case, when re-association is more likely, introduces at least one more timescale into the kinetics of association, thus rendering the two-state switching process non-Markovian. For the sake of simplicity, we stick to the Markovian scenario.

The Fourier-Laplace transforms of the joint density to observe a displacement $\vec{r}$ in a time t in the respective state 1 prior to switching, $\widehat{(F\psi)}_1(q,s)$, is given by Eq. (6.3.12) with $D_1 = D$. Conversely, introducing $s' = s + q^2D/2 + k_2$ (note that $s' = s'(q)$ depends on q) we can write $\widehat{(F\psi)}_2(q,s)$ as (for a derivation see Sec. 6.3.5)

$$\begin{aligned}\widehat{(F\psi)}_2(q,s') &= \frac{1}{s'}\mathrm{e}^{-q^2D/2a}\left(q^4D^2/4a^2\right)^{s'/a} M\left(\frac{s'}{a}, \frac{s'+1}{a}, \frac{q^2D}{2a}\right) \\ &= \mathrm{e}^{-q^2D/2a}\sum_{n=0}^{\infty}\frac{(q^2D/2a)^n}{n!}\frac{1}{s'+na},\end{aligned} \tag{6.3.19}$$

where $M(a,b,z)$ denotes Kummer's confluent hypergeometric function [390]. Using $\widehat{(F\psi)}_1(\vec{q},s)$ and $\widehat{(F\psi)}_2(\vec{q},s')$ in Eq. (6.3.10) we find the Laplace image of the incoherent ISF and, via Eq. (6.3.11), also the incoherent DSF, where we recall that here the coherent and incoherent contribution do not coincide. The ISF may in principle be obtained by analytical inversion of the Laplace transform. For the sake of simplicity, we here perform the inversion numerically with the fixed Talbot method.

The results for the incoherent ISF for the two-state dynamics with $k_1 = k_2 = k$ alongside the frozen mixture and dynamics in the individual pure states, all evolving from an equilibrium initial condition, are shown in Fig. 6.3.3b. The corresponding DSF is depicted in Fig. 6.3.3c and the comparison is further clarified by means of the HWHM shown in Fig. 6.3.3d. Altogether, one can distinguish between the different situations as long as k is not too small or too large. These results demonstrate that the scattering fingerprints proposed here may be a valuable tool for analyzing two-state dynamics also in the presence of internal motions.

New J. Phys. **24** 023004 (2022)

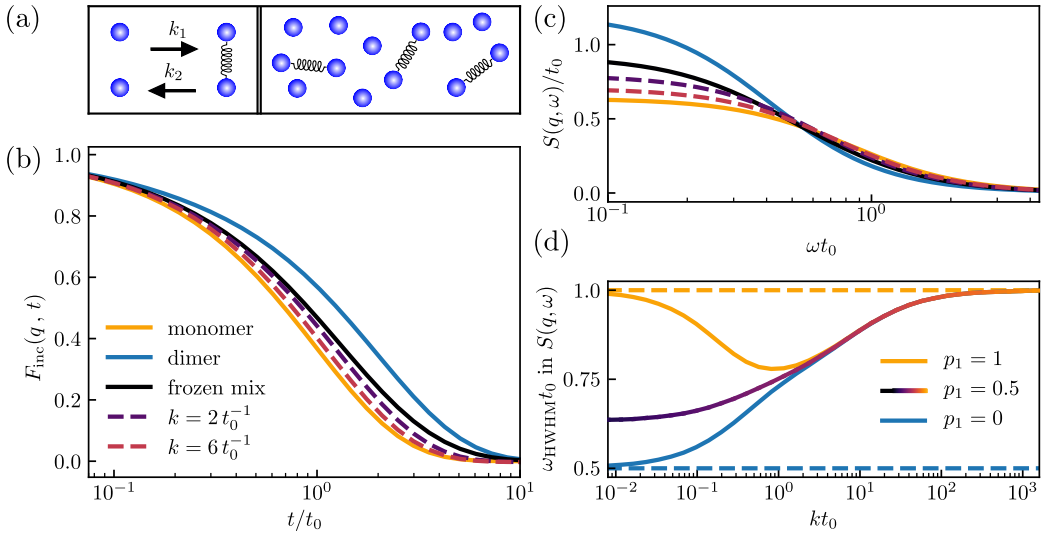

Figure 6.3.3: (a) Schematic of the dimerization process. (b) Incoherent ISF $F_{\mathrm{inc}}(q,t)$ for reversible dimerization of a two-bead (Rouse) dimer with $k = k_1 = k_2$ evolving from an equilibrium occupation $p_1 = p_2 = 0.5$ and time t measured in units of $t_0 = 1/Dq^2$. The limit for $k \to 0$ is the frozen mix (black curve) and for $k \to \infty$ the ISF approaches the ISF of a monomer without dimerization (yellow curve) since the unbinding occurs before the internal mode enters the dynamics. For the explicit derivation of this limit see Sec. 6.3.5. (c) DSFs corresponding to (a). (d) HWHMs of the DSF from (c) are shown along with the HWHMs of the blue and yellow curves in (c) (dashed lines). The HWHMs for initial occupation $p_1 = 0$ and $p_1 = 1$ are also shown (blue and yellow lines).

New J. Phys. **24** 023004 (2022)

New J. Phys. **24** 023004 (2022)

Two-state dynamics with internal structural motions – protein (un)folding

A generalization of the preceding results to two-state dynamics in the presence of some general internal structural relaxation, such as, for example, in the case of cyclization of polymers, and the reversible (un)folding or association of proteins, to name but a few, is much more difficult. In particular, when the change of state does not erase all memory such that the propagation within each sojourn depends on the internal configuration just before the switch, the results of Sec. 6.3.2 cannot be applied. Such a situation naturally arises in systems with internal (e.g., conformational) dynamics in absence of a complete timescale separation, i.e., when at the time of the jump the internal degrees of freedom have not yet completely relaxed from their initial condition or the condition immediately after the preceding change of channel, respectively. If one is to nevertheless apply a version of the theory developed here, additional simplifying assumptions are required. In particular, we must assume that the switching rate is independent of the instantaneous structural state in either dynamical state. Even in this case, the switching between states at any instance occurs from a distribution of structures that set the initial conditions for the relaxation following the jump. Further, as long as the dynamical states have substantially different equilibrium structures, the ISF is expected to depend qualitatively on q (i.e., large and intermediate scale motions are expected to differ in contrast to individual bond vibrations at large q). In full generality, this corresponds to a highly non-trivial problem. To render the general problem analytically tractable, one may, for example, inspect the limit of slow jumps that corresponds to the switching between two equilibrium populations of internal structures. However, this limit is rather uninteresting as the state switching is too slow to mix the dynamics of the individual states, i.e., deviations from a "frozen mix" are small. However, it is easy to instead look at the two-state dynamics with unidirectional/irreversible jumps evolving from a pure state (i.e., from either of the states). A neat and experimentally interesting problem is the denaturant-driven unfolding [391,392] or folding [393] of proteins. To this end, we analyze the irreversible unfolding/folding of the small 20-residue Trp-Cage protein construct TC5b into/from a Rouse chain ($N = 20$). That is, we consider a sample initially prepared in the folded state that unfolds with rate k_{u} (Fig. 6.3.4a), and an initially unfolded (Rouse chain) state folding with a rate k_{f}, respectively (Fig. 6.3.4d).

A Gaussian elastic network model [394,395] of TCPb in the folded state was constructed from the Protein Data Bank (PDB entry: 1L2Y) using the ProDy software [396], yielding a Kirchoff matrix A^{f} with dimensional non-zero entries that are integer multiples of k/ξ where k and ξ are the spring and friction constant, respectively. Conversely, Rouse chain dynamics with connectivity matrix A^{u} [397] were assumed in the unfolded state. A^{f} and A^{u} are given explicitly in Sec. 6.3.5. For simplicity, in both cases the diffusion matrix is

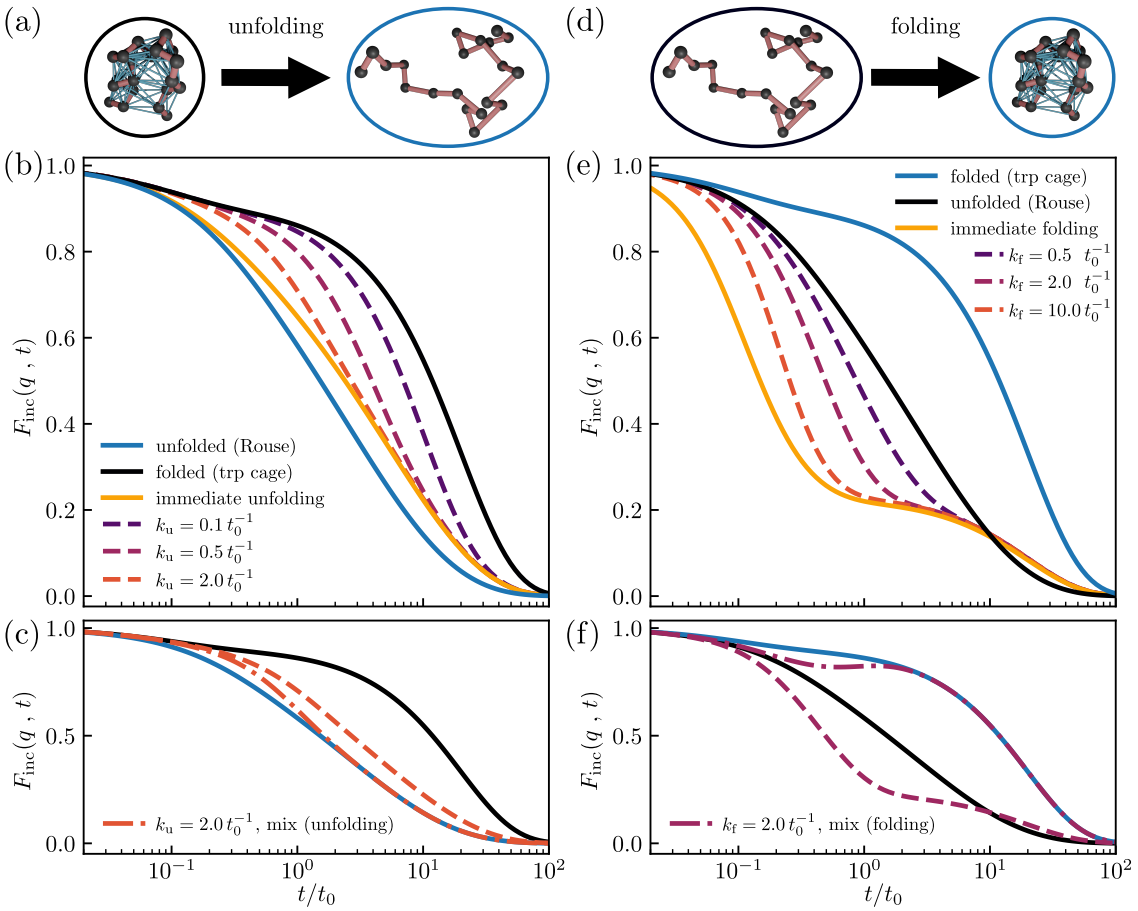

Figure 6.3.4: (a) Schematic of the Gaussian network model for the Trp-Cage Miniprotein Construct TC5b (PDB: 1L2Y) and its unfolding to a Rouse chain ($N = 20$). (b) Incoherent ISF $F_{\mathrm{inc}}(q,t)$ for the (non-reversible) unfolding process, with time measured in units of $t_0 = 1/Dq^2$ and a ratio of spring and friction constant of $k/\xi = t_0^{-1}$. The dashed lines denote various unfolding rates k_u, and the yellow curve, the "immediate unfolding" limit, is obtained by setting $\tau = 0$ in Eq. (6.3.22). Also depicted are the $F_{\mathrm{inc}}(q,t)$ of the respective relaxation processes of the folded protein (black line) and the Rouse chain (blue line). Time is expressed in standard dimensionless Rouse units (see, e.g., [394]). (c) Comparison of the unfolding process (dashed line) with the time-dependent "mix" (dash-dotted line) which corresponds to the superposition of the ISF of the Rouse chain and the protein weighted by the respective time-dependent occupation of the state. (d-f) As in (a-c) but for the converse folding process.

supposed to be equal and diagonal (isotropic) with diffusion coefficient $D = k_\mathrm{B}T/\xi$ for each individual bead. A schematic of the bead-spring model is shown in Fig. 6.3.4a and d.

Introducing the 3N-dimensional vector of beads' positions $\mathbf{R} \equiv \{\vec{R}^\alpha\}_{\alpha=1,\ldots,N}$ the equations of motion in each state follow a 3N-dimensional Ornstein-Uhlenbeck process with positive semi-definite drift matrix $\mathrm{A}^{\mathrm{f/u}}$, i.e., $\mathrm{d}\mathbf{R}_t = -\mathrm{A}^{\mathrm{f/u}}\mathbf{R}_t\mathrm{d}t + \sqrt{2D}\mathrm{d}\mathbf{W}_t$ in the respective state f and u, where $\mathbf{W}$ denotes a $3N$ dimensional vector whose entries are statistically independent Wiener processes W_t. This process describes the overdamped thermal motion of beads connected by Hookean springs with zero rest length. The analysis is most easily carried out in the respective normal coordinates $\mathbf{X}^{\mathrm{f/u}} = (\mathrm{Q}^{\mathrm{f/u}})^T\mathbf{R}$ with $(\mathrm{Q}^{\mathrm{f/u}})^T\mathrm{A}^{\mathrm{f/u}}\mathrm{Q}^{\mathrm{f/u}} = \mathrm{diag}(a_\alpha^{\mathrm{f/u}})_{\alpha=1,\ldots,N}$. For details, see Sec. 6.3.5. We choose $\alpha = 1$ to represent the centre-of-mass mode. The center-of-mass diffusion coefficient is given by $D_\mathrm{COM} = D/N$, and within this model does not depend on the conformation of the protein. For all but the center-of-mass mode, the equilibrium probability density function is Gaussian with zero mean and variance $V_{\mathrm{f/u}}^\alpha$.

To illustrate the theory, it suffices to consider the folding-process alone. Let us consider the propagation in the unfolded state u evolving from the Rouse chain equilibrium distribution. We are interested in the $\alpha\beta$-component of the ISF of the folding process, $F^{\alpha\beta}(\vec{q},t)$, assuming no or a single jump occurring at time τ with $0 < \tau < t$, where τ is exponentially distributed with rate k. Let $F_\mathrm{u}^{\alpha\beta}(\vec{q},t)$ be the ISF of the pure unfolded state (Rouse) dynamics. As we have excluded jumps back to the folded state, it is obvious that $F^{\alpha\beta}(\vec{q},t)$ corresponds to $F_\mathrm{u}^{\alpha\beta}(\vec{q},t)$ weighted by the probability e^{-kt} that no jump has occurred until time t, and a contribution of the folding. The latter corresponds to the dynamics in the folded state starting with a jump at the time τ and evolving from the equilibrium distribution of the unfolded (Rouse chain) state, $F_{\mathrm{u}\to\mathrm{f}}^{\alpha\beta}(\vec{q},t;\tau)$, averaged over the exponentially distributed jumping times[3], i.e.,

$$F^{\alpha\beta}(\vec{q},t) = \mathrm{e}^{-kt}F_\mathrm{u}^{\alpha\beta}(\vec{q},t) + \int_0^t \mathrm{d}\tau k\mathrm{e}^{-k\tau}F_{\mathrm{u}\to\mathrm{f}}^{\alpha\beta}(\vec{q},t;\tau). \tag{6.3.20}$$

Since the Ornstein-Uhlenbeck process with a Gaussian initial condition is a Gaussian process, $F_\mathrm{u}^{\alpha\beta}(\vec{q},t)$ and $F_{\mathrm{u}\to\mathrm{f}}^{\alpha\beta}(\vec{q},t;\tau)$ are characteristic functions (see, e.g., Eq. (6.3.4)) of zero mean Gaussian displacements and thus simply equal to $\exp[-q^2\langle(\vec{R}_t^\alpha - \vec{R}_0^\beta)^2\rangle_{\mathrm{u}/\mathrm{u}\to\mathrm{f}}^{\tau}/6]$, see

[3] It is straightforward to include non-exponential jumping-time distributions.

New J. Phys. **24** 023004 (2022)

Sec. 6.3.5. Via normal mode analysis (see, e.g., [320]) we show in Sec. 6.3.5 that the displacement vectors $\langle(\vec{R}_t^{\,\alpha} - \vec{R}_0^{\,\beta})^2\rangle_{\mathrm{u/u}\overset{\tau}{\to}\mathrm{f}}$ obey

$$\left\langle \left(\vec{R}_t^{\,\alpha} - \vec{R}_0^{\,\beta}\right)^2 \right\rangle_{\mathrm{u}} = 6D_{\mathrm{COM}}t + \sum_{\gamma=2}^{N} V_{\mathrm{u}}^{\gamma}\left[(Q_{\alpha\gamma}^{\mathrm{u}})^2 + (Q_{\beta\gamma}^{\mathrm{u}})^2 - 2Q_{\alpha\gamma}^{\mathrm{u}}Q_{\beta\gamma}^{\mathrm{u}}\mathrm{e}^{-a_\gamma^{\mathrm{u}}t}\right], \tag{6.3.21}$$

when the Rouse polymer does not fold until t, whereas in the case that folding occurs at a fixed time τ we find, introducing the notation $\mathrm{Q}^{\mathrm{u}\to\mathrm{f}} = (\mathrm{Q}^{\mathrm{f}})^T\mathrm{Q}^{\mathrm{u}}$ (such that analogous to $\mathbf{R}_t = \mathrm{Q}^{\mathrm{f/u}}\mathbf{X}_t^{\mathrm{f/u}}$ we have $\mathbf{X}_t^{\mathrm{f}} = \mathrm{Q}^{\mathrm{u}\to\mathrm{f}}\mathbf{X}_t^{\mathrm{u}}$),

$$\begin{aligned}\left\langle \left(\vec{R}_t^{\,\alpha} - \vec{R}_0^{\,\beta}\right)^2 \right\rangle_{\mathrm{u}\overset{\tau}{\to}\mathrm{f}} &= 6D_{\mathrm{COM}}t + \sum_{\gamma=2}^{N}(Q_{\beta\gamma}^{\mathrm{u}})^2 V_{\mathrm{u}}^{\gamma} + \sum_{\gamma=2}^{N}(Q_{\alpha\gamma}^{\mathrm{f}})^2 V_{\mathrm{f}}^{\gamma}\left[1 - \mathrm{e}^{-2a_\gamma^{\mathrm{f}}(t-\tau)}\right] \\ &+ \sum_{\omega=2}^{N} V_{\mathrm{u}}^{\omega}\left[\sum_{\gamma=2}^{N} Q_{\alpha\gamma}^{\mathrm{f}}Q_{\gamma\omega}^{\mathrm{u}\to\mathrm{f}}\mathrm{e}^{-a_\gamma^{\mathrm{f}}(t-\tau)}\right]^2 - 2\sum_{\omega,\gamma=2}^{N} Q_{\alpha\omega}^{\mathrm{f}}Q_{\omega\gamma}^{\mathrm{u}\to\mathrm{f}}Q_{\beta\gamma}^{\mathrm{u}}V_{\mathrm{u}}^{\gamma}\mathrm{e}^{-a_\omega^{\mathrm{f}}(t-\tau)-a_\gamma^{\mathrm{u}}\tau}.\end{aligned} \tag{6.3.22}$$

The converse unfolding process is obtained by interchanging $\mathrm{u} \leftrightarrow \mathrm{f}$. Plugging Eqs. (6.3.21)-(6.3.22) into Eq. (6.3.20) and performing a numerical integration over τ yields the result for $F^{\alpha\beta}(\vec{q},t)$. Summations over α,β deliver the (in)coherent ISF, see Eqs. (6.3.2)-(6.3.3).

In Figs. 6.3.4b-c and Fig. 6.3.4e-f we show, respectively, the incoherent part of the ISF for various unfolding and folding rates. The limit of large (un)folding rates is given by setting $\tau = 0$ in the Gaussian ISF for the displacement vector from Eq. (6.3.22). The discrepancy between the immediate (un)folding limit and dynamics in the target (un)folded structure (yellow and blue curves in Fig. 6.3.4b,e, respectively) is solely caused by the different initial conditions whereas the propagator is in both cases identical. Notably, this difference is small in the unfolding process but substantial during folding. Several additional remarks are in order. The ISF reflects dynamics on a given spatial (i.e., q-value) scale, and thus the projections onto different relaxation eigenmodes play an important role (see also, e.g., [73]). The aforementioned discrepancies are a result of initial conditions (and starting conditions upon a jump) that put more weight onto faster relaxation modes, rendering the decay of the ISF faster (this idea is corroborated in Fig. 6.3.4e). Note that relaxation in the folded state is faster than in the unfolded, and yet the ISF in the folded state in Fig. 6.3.4b decays the slowest. This means that said differences are not only caused by a shorter relaxation time (i.e., larger principal eigenvalue of the underlying generator).

For a comparison, in Fig. 6.3.4c,f we also show the incoherent ISF for the time-dependent mix, $F_{\mathrm{u/f}}^{\mathrm{mix}}(\vec{q},t) \equiv \mathrm{e}^{-k_{\mathrm{u/f}}t}\sum_\alpha F_{\mathrm{f/u}}^{\alpha\alpha}(\vec{q},t) + (1-\mathrm{e}^{-k_{\mathrm{u/f}}t})\sum_\alpha F_{\mathrm{u/f}}^{\alpha\alpha}(\vec{q},t)$, which represents the trivial time-dependent approximation of the ISF for the unfolding and folding processes,

New J. Phys. 24 023004 (2022)

respectively. In the case of unfolding, the (trivial) mix displays qualitative differences with respect to exact the two-state (single jump) solution, but the quantitative differences are rather small (see Fig. 6.3.4c). Conversely, in the case of folding, the mix-approximation $F_{\mathrm{f}}^{\mathrm{mix}}$ fails completely (see Fig. 6.3.4f) due to the striking dependence on the initial condition that is not included in the (trivial) mix. This clearly illustrates the importance of considering jump dynamics explicitly.

Overall, our results suggest that scattering fingerprints may be a useful observable to probe two-state dynamics even in the presence of non-trivial internal structural relaxation, such as, e.g., probed in Ref. [391–393]. A systematic study of the relaxation at various values of the dimensionless quantities $k_{u,f}t_0$ and t_0k/ξ may provide further details, which is, however, beyond the scope of the present proof-of-principle investigation. Moreover, one can easily generalize the model to include internal friction [321] or hydrodynamic interactions in the spirit of the Zimm model [319]. Finally, one may also consider springs with a non-zero rest length in the context of so-called Gaussian models (see, e.g., [394]).

New J. Phys. **24** 023004 (2022)

6.3.4 Conclusion

Scattering experiments on polymerizing chains revealed pronounced signatures of multi-state dynamics [383]. Moreover, observations of particle transport in complex environments such as the interior of living cells often reveal a non-Fickian (or non-Brownian) character that may also display multi-state characteristics [91, 361–369, 372, 373]. Reliably distinguishing between single- and multi-state dynamics remains challenging. To complement the existing approaches, we addressed scattering fingerprints of two-state dynamics – the intermediate scattering function and structure factor. A combination of theory, analysis of simple model systems and of experiments in living cells revealed the potential usefulness of these scattering fingerprints.

We addressed both, inert tracer-particle dynamics and dynamics in systems with an internal structure and dynamics. In all examples, the ISF decays faster for increasing switching rates, which is consistent with the idea of an additional relaxation channel enabling a faster decay [398, 399]. However, this view contradicts the multi-state dynamics in [383], where an empirical ansatz assumed a mode-scission that leads to a slower decay of the ISF.

The present results can be extended to incorporate non-Markovian (i.e., non-exponential) waiting time statistics in the respective states (see, e.g., [169, 400, 401]), and may be relevant and useful for digital Fourier microscopy (differential dynamic microscopy and Fourier imaging correlation spectroscopy) experiments on complex particulate systems, as

well as neutron (including spin-echo) and X-ray scattering probing structural and dynamical properties of macromolecules, as soon as the dynamics displays two-state transport.

Acknowledgments

We thank Maximilian Vossel for preparing Figs. 4a and 4d, and Adal Sabri for providing the data from Ref. [91]. The financial support from the Studienstiftung des Deutschen Volkes (to C. D.), and the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-2 (to A. G.) is gratefully acknowledged.

6.3.5 Appendix

Derivation of Eq. (6.3.10)

Here, we derive Eq. (6.3.10) for the Laplace-transformed intermediate scattering function (ISF) of the two-state dynamics following the Ref. [386]. For this approach to hold, memory in the process has to be erased at the instance of the jump and the instantaneous position at the jump must not influence the probability of subsequent displacements. These assumptions can arise as a consequence of translation invariance of the dynamics in single dynamical states, as in the examples of two-state diffusion and FBM. However, translation invariance is not a necessary condition, as e.g., in the case of the process studied in [386].

For a Markov jump process, the probability to remain in a state for a time t is given by $\psi_j(t) = \mathrm{e}^{-k_j t}$. Recall that the van Hove functions $G_j(\vec{r},t) = \langle\delta(\vec{r}(t)-\vec{r}(0)-\vec{r})\rangle_j$ in the two individual states $j = 1, 2$ give the probability of performing a displacement $\vec{r}$ in time t. We now set $N = 1$ and omit the indices α, β compared to Eq. (6.3.1). For $N > 1$, the assumption of memory erasure is rarely satisfied as a result of internal dynamics (see, e.g., Sec. 6.3.3). If it is satisfied also for $N > 1$, as e.g., in Sec. 6.3.3 where $N = 2$, then for $\alpha = \beta$ the approach does not change and Eq. (6.3.10) gives an equation for $\widehat{F}^{\alpha\alpha}(\vec{q},s)$ that can subsequently be summed over α. For $\alpha \neq \beta$ only the displacement probability prior to the first memory erasure is different, and the approach below may applied with slight modifications.

Following [386], we write down the probability $\mathcal{P}_n(\vec{r},t)$ of a displacement $\vec{r}$ in time t conditioned on a fixed number of n jumps for $n = 0, 1, 2$, given that we start in state 1 at $t = 0$, and adopting the notation $G_j(\vec{r},t)\psi_j(t) \equiv (G\psi)_j(\vec{r},t)$,

$$\mathcal{P}_0(\vec{r},t) = (G\psi)_1(\vec{r},t)$$

New J. Phys. **24** 023004 (2022)

$$\mathcal{P}_1(\vec{r},t) = \int \mathrm{d}^3 r_1 \int_0^t \mathrm{d}t_1 k_1 (G\psi)_1(\vec{r}_1,t_1)(G\psi)_2(\vec{r}-\vec{r}_1,t-t_1)$$
$$\mathcal{P}_2(\vec{r},t) = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 k_1 (G\psi)_1(\vec{r}_1,t_1) k_2 (G\psi)_2(\vec{r}_2-\vec{r}_1,t_2-t_1)\times$$
$$(G\psi)_1(\vec{r}-\vec{r}_2,t-t_2). \tag{6.3.23}$$

By the assumption of independence of the position at the time of the jump, this has a convolution structure in space and we Fourier transform $\vec{r}\to\vec{q},\ \mathcal{P}\to\widetilde{\mathcal{P}}$ to obtain, e.g., for $n=2$,

$$\widetilde{\mathcal{P}}_2(\vec{q},t) = \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 k_1 \widetilde{(G\psi)}_1(\vec{q},t_1) k_2 \widetilde{(G\psi)}_2(\vec{q},t_2-t_1) \widetilde{(G\psi)}_1(\vec{q},t-t_2). \tag{6.3.24}$$

Recalling that the intermediate scattering function (ISF) $F(\vec{q},t)$ is the Fourier transform of the van Hove function $G(\vec{r},t)$, and noting the convolution structure (defined as $[f*g](t)\equiv \int_0^t \mathrm{d}t' f(t')g(t-t')$) in time, we have

$$\begin{aligned}\widetilde{\mathcal{P}}_2(\vec{q},t) &= \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 k_1 (F\psi)_1(\vec{q},t_1) k_2 (F\psi)_2(\vec{q},t_2-t_1)(F\psi)_1(\vec{q},t-t_2)\\ &= k_1 k_2 \int_0^t \mathrm{d}t_2 (F\psi)_1(\vec{q},t-t_2)\left[(F\psi)_1(\vec{q},\cdot)*(F\psi)_2(\vec{q},\cdot)\right](t_2)\\ &= k_1 k_2 \left[(F\psi)_1 * (F\psi)_1 * (F\psi)_2\right](\vec{q},t).\end{aligned} \tag{6.3.25}$$

Taking the Laplace transform $t\to s$ gives

$$\widehat{\mathcal{P}}_2(\vec{q},s) = k_1 k_2 \left[\widehat{(F\psi)}_1(\vec{q},s)\right]^2 \widehat{(F\psi)}_2(\vec{q},s). \tag{6.3.26}$$

This generalizes to all even $n=2m$ and odd $n=2m+1$ terms as

$$\begin{aligned}\widehat{\mathcal{P}}_{2m+1}(\vec{q},s) &= k_1^{m+1} k_2^m \left[\widehat{(F\psi)}_1(\vec{q},s)\right]^{m+1} \left[\widehat{(F\psi)}_2(\vec{q},s)\right]^{m+1},\\ \widehat{\mathcal{P}}_{2m}(\vec{q},s) &= k_1^{m} k_2^m \left[\widehat{(F\psi)}_1(\vec{q},s)\right]^{m+1} \left[\widehat{(F\psi)}_2(\vec{q},s)\right]^{m}.\end{aligned} \tag{6.3.27}$$

The Fourier-Laplace transform of the probability of a displacement in the two-state dynamics that by the assumption of memory erasure and independence of the jump position is the Laplace-transformed intermediate scattering function $\widehat{F}(\vec{q},s)$, is then given by the geometric series

$$\widehat{F}(\vec{q},s) = \sum_{n=0}^{\infty} \widehat{\mathcal{P}}_n(\vec{q},s)$$

New J. Phys. **24** 023004 (2022)

$$
\begin{aligned}
&= \sum_{m=0}^{\infty} \left[\widehat{\mathcal{P}}_{2m}(\vec{q},s) + \widehat{\mathcal{P}}_{2m+1}(\vec{q},s)\right] \\
&= \widehat{(F\psi)}_1(\vec{q},s)\left[1 + k_1\widehat{(F\psi)}_2(\vec{q},s)\right] \sum_{m=0}^{\infty} \left[k_1k_2\widehat{(F\psi)}_1(\vec{q},s)\widehat{(F\psi)}_2(\vec{q},s)\right]^m \\
&= \widehat{(F\psi)}_1(\vec{q},s)\frac{1 + k_1\widehat{(F\psi)}_2(\vec{q},s)}{1 - k_1k_2\widehat{(F\psi)}_1(\vec{q},s)\widehat{(F\psi)}_2(\vec{q},s)}.
\end{aligned} \tag{6.3.28}
$$

Assuming initial occupations p_j of the two states (not necessarily $p_j = p_j^{\mathrm{eq}}$), and repeating the same treatment for starting conditions in state 2 yields the result Eq. (6.3.10),

$$
\widehat{F}(\vec{q},s) = \frac{p_1\widehat{(F\psi)}_1\left[1 + k_1\widehat{(F\psi)}_2\right] + p_2\widehat{(F\psi)}_2\left[1 + k_2\widehat{(F\psi)}_1\right]}{1 - k_1k_2\widehat{(F\psi)}_1\widehat{(F\psi)}_2}. \tag{6.3.29}
$$

Two-state diffusion in the slow- and fast-switching limit

Here we consider the slow-switching limit of two-state diffusion in Sec. 6.3.3, that is, the limit of small switching rates, $k_1 + k_2 \ll q^2D_{1,2}$. In this limit, Eq. (6.3.14) gives $\mu_+ \approx q^2D_1$ and $\mu_- \approx q^2D_2$. Since $p_2 = 1 - p_1$ we obtain $\phi(q) \approx p_2$ and thus Eq. (6.3.16) in the slow-switching limit results in the frozen mixture,

$$
F(q,t) \approx p_1\mathrm{e}^{-q^2D_1t} + p_2\mathrm{e}^{-q^2D_2t} = F_{\mathrm{mix}}(q,t). \tag{6.3.30}
$$

Now consider the fast-switching limit, $k_1 + k_2 \gg q^2D_{1,2}$ with a finite ratio $k_1/k_2 = p_2^{\mathrm{eq}}/p_1^{\mathrm{eq}}$ (when this ratio is zero or infinite, we recover single-state dynamics). We introduce the notation $k \equiv (k_1 + k_2)/2$, $\kappa \equiv (k_1 - k_2)/2$, $\bar{d} \equiv q^2(D_1 + D_2)/2$ and $\Delta \equiv q^2(D_2 - D_1)/2$. Then Eq. (6.3.14) reads

$$
\begin{aligned}
\mu_{1,2} &= \bar{d} + k \pm \sqrt{(\Delta + \kappa)^2 + k^2 - \kappa^2} \\
&= \bar{d} + k \pm \sqrt{\Delta^2 + 2\Delta\kappa + k^2}.
\end{aligned} \tag{6.3.31}
$$

In the limit of fast jumps $k \gg \bar{d} > |\Delta|$ we have

$$
\begin{aligned}
&\mu_{1,2} = \bar{d} + k \pm k\sqrt{1 + \frac{2\Delta\kappa + \Delta^2}{k^2}} \approx \bar{d} + k\left[1 \pm \left(1 + \frac{2\Delta\kappa + \Delta^2}{2k^2}\right)\right], \\
&\mu_1 \approx \bar{d} - \frac{2\Delta\kappa + \Delta^2}{2k} \approx \bar{d} - \frac{\kappa}{k}\Delta, \quad \mu_2 \approx \bar{d} + 2k + \frac{2\Delta\kappa + \Delta^2}{2k} \approx 2k \overset{k\to\infty}{\longrightarrow} \infty.
\end{aligned} \tag{6.3.32}
$$

New J. Phys. **24** 023004 (2022)

Using Eq. (6.3.7), we have $k_1 = 2p_2^{\rm eq}k$ and $k_2 = 2p_1^{\rm eq}k$ and arrive at

$$\mu_1 \approx \bar{d} - \frac{\kappa}{k}\Delta = \bar{d} + \Delta\frac{p_2^{\rm eq} - p_1^{\rm eq}}{p_2^{\rm eq} + p_1^{\rm eq}} = q^2(p_1^{\rm eq}D_1 + p_2^{\rm eq}D_2). \tag{6.3.33}$$

Thus, the exponential with rate μ_1 in Eq. (6.3.16) gives an exponential with effective diffusion constant $D_{\rm eff} \equiv p_1^{\rm eq}D_1 + p_2^{\rm eq}D_2$, while the second exponential with rate $\mu_2 \approx 2k$ immediately decays in the limit of large k. This proves that for large k, the ISF of two-state diffusion approaches effective diffusion, i.e., in Fig. 6.3.1a the dashed lines approach the yellow line.

New J. Phys. 24 023004 (2022)

Laplace transforms for fractional Brownian motion

We consider fractional Brownian motion (FBM) in d dimensions with mean squared displacement

$$\left\langle [\vec{r}(t) - \vec{r}(0)]^2 \right\rangle = 2dC_j t^{\alpha_j}. \tag{6.3.34}$$

Here, the two different states are characterized by different generalized diffusion constants C_1 and C_2 and/or different "anomalous" exponents α_1 and α_2. We still consider Markov jumps and by Gaussianity we have the Fourier-transformed displacement probabilities in a state

$$F_j(\vec{q},t)\psi_j(t) = \exp(-[q^2C_j t^{\alpha_j} + k_j]\,t). \tag{6.3.35}$$

Assuming that each jump (i.e., change of state) erases memory the structure factor for the two-state dynamics follows from Eq. (6.3.10) which requires taking the Laplace transform of Eq. (6.3.35). For our purposes we consider subdiffusive FBM with $\alpha \in (0,1)$ (for $\alpha = 1$ see Eq. (6.3.12)). The Laplace transform of Eq. (6.3.35) can of course be performed numerically [402], and for any rational $\alpha \in (0,1)$ it can be expressed in terms of Meijer G-functions, using that for two positive integers $l < m$ [403]

$$\widehat{[\mathrm{e}^{-at^{l/m}}]}(s) = \frac{\sqrt{ml}}{s(2\pi)^{(m+l)/2-1}} G_{l,m}^{m,l}\left(\frac{a^m l^l}{m^m p^l}\middle|\begin{matrix}\Delta(l,0)\\ \Delta(m,0)\end{matrix}\right). \tag{6.3.36}$$

For the special case $\alpha = 0.5$ the Laplace transform of Eq. (6.3.35) follows from [403]

$$\widehat{[\mathrm{e}^{-a\sqrt{t}}]}(s) = \frac{1}{s} - \frac{a}{2}\sqrt{\frac{\pi}{s^3}}\,\mathrm{erfcx}\left(\frac{a}{2\sqrt{s}}\right), \tag{6.3.37}$$

where $\mathrm{erfcx}(x) = \exp(x^2)(1 - \mathrm{erf}(x))$ and erf denotes the error function, and therefore

$$\widehat{(F\psi)}_j(\vec{q}, s) = \frac{1}{s + k_j} - \frac{\sqrt{\pi} q^2 C_j}{2(s + k_j)^{3/2}} \mathrm{erfcx}\left(\frac{q^2 C_j}{2\sqrt{s + k_j}} \right). \tag{6.3.38}$$

Inserting $s = -i\omega$ and taking the real part as in Eq. (6.3.17), we obtain the DSF. Numerically transforming back to the time-domain in turn yields the ISF.

For $\alpha \neq 0.5$, instead of using Meijer-G-functions or numerical solutions, we can perform a series expansion [402]. We expand e^{-ct^α} around $t = 0$

$$\begin{aligned} \int_0^\infty \mathrm{e}^{-st} \mathrm{e}^{-ct^\alpha} \mathrm{d}t &= \sum_{n=0}^\infty \frac{(-1)^n c^n}{n!} \int_0^\infty \mathrm{e}^{-st} t^{n\alpha} \mathrm{d}t \\ &= \sum_{n=0}^\infty \frac{(-1)^n \Gamma(\alpha n + 1)}{n!} c^n s^{-(\alpha n + 1)}, \end{aligned} \tag{6.3.39}$$

where the sum and integral commute since the sum is bounded by e^{-ct^α}, $c > 0$ which is integrable for $\alpha < 1$. A series for the Laplace transform of Eq. (6.3.35) is obtained by shifting $s \mapsto s + k_j$ and using $c = q^2 C_j$. This approximation works particularly well for or small q, not too small s, k and $\alpha < 1$ not too close to 1.

The complementary expansion for small s reads

$$\begin{aligned} \int_0^\infty \mathrm{e}^{-st} \mathrm{e}^{-ct^\alpha} \mathrm{d}t &\approx \sum_{n=0}^N \frac{(-1)^n s^n}{n!} \int_0^\infty t^n \mathrm{e}^{-ct^\alpha} \mathrm{d}t \\ &= \frac{1}{\alpha} \sum_{n=0}^N \frac{(-1)^n}{n!} \Gamma\left(\frac{n+1}{\alpha} \right) c^{-\frac{n+1}{\alpha}} s^n. \end{aligned} \tag{6.3.40}$$

However, note that for $\alpha < 1$ this series is only asymptotically convergent [402], and therefore has to be truncated at some finite N to be able to swap the integration and summation. We use this expansion in the very small s regime in cases where the expansion (6.3.39) does not suffice.

The two asymptotic results combined yield a very good approximation as illustrated for the DSF of a FBM in Fig. 6.3.5 (recall that the DSF follows from the Laplace transform of Eq. (6.3.35) for $k_j = 0$ by setting $s = -\mathrm{i}\omega$ and taking the real part as in Eq. (6.3.11)). Combining the two expansions allows for efficient computation of $\widehat{(F\psi)}(q, s)$ and $S(q, \omega)$ over the whole ω regime for general $\alpha \in (0, 1)$, which in turn allows via Eq. (6.3.10) to deduce the two-state ISF.

New J. Phys. **24** 023004 (2022)

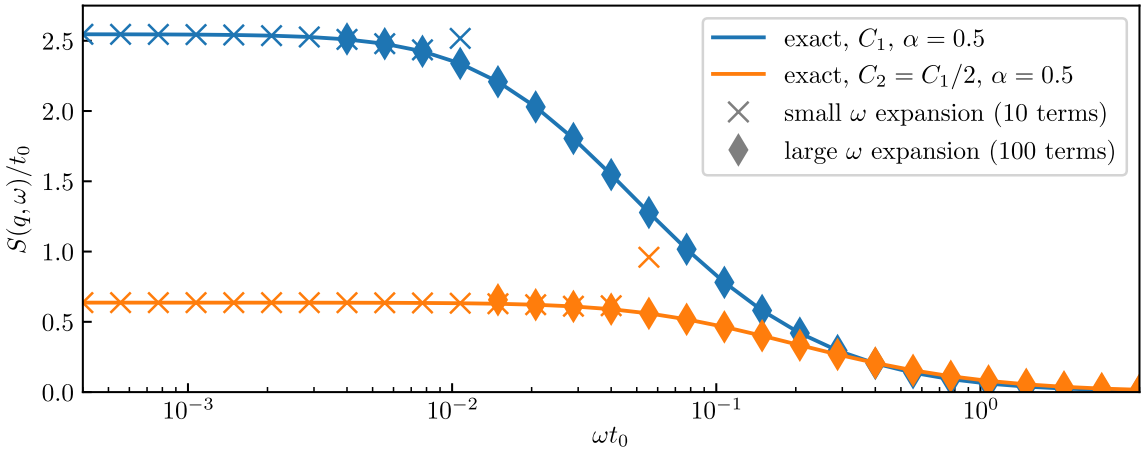

Figure 6.3.5: Illustration of the asymptotics for the DSF as introduced in Eq. (6.3.11). Time is measured in units of $t_0 = 1/(C_1^2 q^4)$. The large-ω series (6.3.39) and the small-ω expansion (6.3.40) together give a very good approximation to the exact DSF, i.e., the Fourier transform of Eq. (6.3.35) with $k = 0$.

New J. Phys. 24 023004 (2022)

Simulation details

The experimental data shown in Fig. 6.3.2 consists of 200 two-dimensional trajectories which by assuming isotropy are treated as 400 one-dimensional realizations of the two-state FBM. The experimental ISF is obtained by averaging $e^{-iq[r(t)-r(0)]}$ over the 400 realizations. Although the ISF of the model is deterministic, the experimental ISF obtained from a finite number of realizations fluctuates around the deterministic ISF. To estimate these fluctuations we performed simulations of $20,000 = 500 \times 400$ two-state FBM trajectories with the parameters obtained by the fit. The standard deviation in each point of the ISF averaged over sets of 400 trajectories is shown in Fig. 6.3.2. Moreover, the MSD and its fluctuations upon averaging over 400 realizations are obtained from the simulations and shown in Fig. 6.3.2.

Simulations were performed by drawing exponentially distributed waiting times for the state switching and simulating FBM trajectories between the switching times. The FBM simulations were performed using the Davies-Harte-algorithm [404] using the "fbm" python package available on PyPI.

Details on dimerization with an internal mode

The ISF of the dimer is the characteristic function of a Gaussian with displacement given by Eq. (6.3.21) for $N = 2$ for which

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \qquad Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \tag{6.3.41}$$

With $D_{\mathrm{COM}} = D_2 = D/2$, $V(\infty) = 3D/a$ and $\psi_2(t) = \mathrm{e}^{-k_2 t}$ we obtain

$$\begin{aligned} (F\psi)_2(q,t) &= \exp\left[-q^2 D_2 t - k_2 t - \frac{q^2 D}{2a}\left(1 - \mathrm{e}^{-at}\right)\right] \\ &= \mathrm{e}^{-\frac{q^2 D}{2a}} \mathrm{e}^{-(q^2 D_2 + k_2)t} \exp\left(\frac{q^2 D}{2a} \mathrm{e}^{-at}\right), \end{aligned} \tag{6.3.42}$$

where for convenience we have assumed equilibrium initial conditions of the dimer.

The Laplace transform can be conveniently carried out via a series expansion (the order of summation and integration can be interchanged due to dominated convergence, i.e., since the series is bounded from above by the full expression),

$$\begin{aligned} \widehat{(F\psi)}_2(q,s) &= \mathrm{e}^{-\frac{q^2 D}{2a}} \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-(s + q^2 D_2 + k_2)t} \exp\left(\frac{q^2 D}{2a} \mathrm{e}^{-at}\right) \\ &= \mathrm{e}^{-\frac{q^2 D}{2a}} \sum_{n=0}^\infty \frac{(\frac{q^2 D}{2a})^n}{n!} \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-(s + q^2 D_2 + k_2 + na)t} \\ &= \mathrm{e}^{-\frac{q^2 D}{2a}} \sum_{n=0}^\infty \frac{(\frac{q^2 D}{2a})^n}{n!} \frac{1}{s + q^2 D_2 + k_2 + na}. \end{aligned} \tag{6.3.43}$$

The series expansion converges very well, especially for small q and large t (or small ω). Note that there exists a closed form expression of $\widehat{(F\psi)}_2(q,s)$ in terms Kummer's hypergeometric function [403] (or in terms of the incomplete Gamma function [380]) as in Eq. (6.3.19).

In the limit of fast jumps $k_1, k_2 \gg q^2 D$ the time spent in a single state prior to a jump becomes very small, $at \ll 1$, for which we expand

$$\begin{aligned} (F\psi)_2(\vec{q},t) &= \mathrm{e}^{-\frac{q^2 D}{2a}} \mathrm{e}^{-(q^2 D_2 + k_2)t} \exp\left(\frac{q^2 D}{2a} \mathrm{e}^{-at}\right) \\ &\overset{t \to 0}{\approx} \mathrm{e}^{-\frac{q^2 D}{2a}} \mathrm{e}^{-(q^2 D_2 + k_2)t} \mathrm{e}^{\frac{q^2 D}{2a}(1 - at)} \\ &= \mathrm{e}^{-(q^2 D + k_2)t} \\ &= F_1(\vec{q},t)\psi_2(t). \end{aligned} \tag{6.3.44}$$

New J. Phys. **24** 023004 (2022)

Therefore, for $k_1, k_2 \gg q^2 D$ the two-state ISF converges to the simple monomer diffusion, i.e., the dashed lines approach the yellow line in Fig. 6.3.3b.

Two-state dynamics with internal structure/dynamics – irreversible protein (un)folding

Here we derive Eqs. (6.3.21)-(6.3.22) using normal mode analysis as introduced in Sec. 6.3.3, i.e., in the normal coordinates $\mathbf{X}^{\mathrm{f/u}} = (\mathrm{Q}^{\mathrm{f/u}})^T \mathbf{R}$ with $(\mathrm{Q}^{\mathrm{f/u}})^T \mathrm{A}^{\mathrm{f/u}} \mathrm{Q}^{\mathrm{f/u}} = \mathrm{diag}(a_\alpha^{\mathrm{f/u}})_{\alpha=1,\dots,N}$. The drift matrix $\mathrm{A}^{\mathrm{f/u}}$ can be considered as $N \times N$-dimensional since the model assumes isotropy in all three dimensions and thus each eigenspace of the $3N \times 3N$ drift matrix is threefold degenerate. Note that $(Q^{\mathrm{f/u}})^{-1} = (Q^{\mathrm{f/u}})^T$, and in normal coordinates (omitting index $\mathrm{f/u}$) we have $\mathrm{d}\vec{X}_t^\alpha = -a_\alpha \vec{X}_t^\alpha \mathrm{d}t + \sqrt{2D}\mathrm{d}\vec{W}_t^\alpha$.

We choose the first mode $\alpha = 1$ to correspond to the zero eigenvalue, i.e., $a_1 = 0$ and thus $\vec{X}_t^1$ evolves as a free diffusion, capturing to the center-of-mass motion with $Q_{\alpha 1}^{\mathrm{f/u}} = N^{-1/2}$ and $D_{\mathrm{COM}} = D/N$. The position at time $t = 0$ of the center of mass cancels out in all displacements and only differences between time 0 and t enter. Thus, whenever we speak of the equilibrium initial conditions we refer to the equilibrium of modes $\vec{X}^\alpha$, $\alpha > 1$, and the initial condition of the $\vec{X}^1$-mode is irrelevant.

The $\vec{X}^\alpha$ modes with $\alpha > 1$ have $a_\alpha > 0$ and are thus described by an Ornstein-Uhlenbeck process [102], for which we obtain from the stochastic differential equation that for any $\tau \in [0, t]$

$$\mathrm{d}\left\langle \vec{X}_{t'}^\alpha \cdot \vec{X}_0^\beta \right\rangle = -a_\alpha \left\langle \vec{X}_{t'}^\alpha \cdot \vec{X}_0^\beta \right\rangle \mathrm{d}t' \quad \Rightarrow \quad \left\langle \vec{X}_t^\alpha \cdot \vec{X}_0^\beta \right\rangle = \mathrm{e}^{-a_\alpha(t-\tau)} \left\langle \vec{X}_\tau^\alpha \cdot \vec{X}_0^\beta \right\rangle, \tag{6.3.45}$$

and using Itô's lemma as well as the shorthand notation $V^\alpha(t) = \frac{3D}{a_\alpha}(1 - \mathrm{e}^{-2a_\alpha t})$,

$$\begin{aligned} \mathrm{d}\left\langle \vec{X}_{t'}^\alpha \cdot \vec{X}_{t'}^\beta \right\rangle &= \left\langle \mathrm{d}\vec{X}_{t'}^\alpha \cdot \vec{X}_{t'}^\beta \right\rangle + \left\langle \vec{X}_{t'}^\alpha \cdot \mathrm{d}\vec{X}_{t'}^\beta \right\rangle + \frac{1}{2}\left\langle \mathrm{d}\vec{X}_{t'}^\alpha \cdot \mathrm{d}\vec{X}_{t'}^\beta \right\rangle \\ &= -(a_\alpha + a_\beta)\left\langle \vec{X}_{t'}^\alpha \cdot \vec{X}_{t'}^\beta \right\rangle \mathrm{d}t' + 3D\delta_{\alpha\beta}\mathrm{d}t' \\ \Rightarrow \quad \left\langle \vec{X}_t^\alpha \cdot \vec{X}_t^\beta \right\rangle &= \left\langle \vec{X}_\tau^\alpha \cdot \vec{X}_\tau^\beta \right\rangle \mathrm{e}^{-(a_\alpha+a_\beta)(t-\tau)} + \delta_{\alpha\beta} V^\alpha(t-\tau), \end{aligned} \tag{6.3.46}$$

and thus for equilibrium initial conditions $\left\langle \vec{X}_t^\alpha \cdot \vec{X}_t^\beta \right\rangle_{\mathrm{eq}} = \delta_{\alpha\beta} V^\alpha(\infty) \equiv \delta_{\alpha\beta} V^\alpha$.

Since the Ornstein-Uhlenbeck process with a Gaussian initial condition is a Gaussian process, $F_{\mathrm{u}}^{\alpha\beta}(\vec{q}, t)$ and $F_{\mathrm{u}\to\mathrm{f}}^{\alpha\beta}(\vec{q}, t; \tau)$ are characteristic functions, of zero mean Gaussian displacement vectors. For any real Gaussian stochastic process Y_t, $0 \le t \le T$, the character-

New J. Phys. 24 023004 (2022)

istic function of the joint probability distribution of the random variables $Y_{t_1} \dots Y_{t_n}$ with $0 \leq t_1 \dots t_n \leq T$ ($n < \infty$) is given by $\phi(q_1 \dots q_n) = \exp(\mathrm{i} \sum_{k=1}^{n} \langle Y_{t_k} \rangle q_k - \sum_{k,l=1}^{n} \langle [Y_{t_k} - \langle Y_{t_k} \rangle][Y_{t_l} - \langle Y_{t_l} \rangle] \rangle q_k q_l / 2)$, which naturally generalizes to d-dimensional Gaussian stochastic process with vector values $\vec{Y}_t = (Y_t^1, \dots, Y_t^d)$ ($d < \infty$) [405]. In our case we have $\vec{Y}_{t_k} = \vec{R}_{t_k}^{\alpha} - \vec{R}_{t_{k-1}}^{\beta}$, $n = 1$, and $\langle \vec{Y}_{t_k} \rangle = 0$, $\forall t_k$ because all involved Gaussian distributions are centered at zero, and the additional factor of $1/3$ in $q^2/6$ is due to isotropy. Therefore, $F^{\alpha\beta}_{\mathrm{u/u}\xrightarrow{\tau}\mathrm{f}}(\vec{q}, t) = \exp[-q^2 \langle (\vec{R}_t^{\alpha} - \vec{R}_0^{\beta})^2 \rangle_{\mathrm{u/u}\xrightarrow{\tau}\mathrm{f}} / 6]$ and we only need to compute the squared displacements $\langle (\vec{R}_t^{\alpha} - \vec{R}_0^{\beta})^2 \rangle_{\mathrm{u/u}\xrightarrow{\tau}\mathrm{f}}$ to obtain the ISFs.

We first consider the single-state process, say in the unfolded state u, with equilibrium initial conditions and now derive the result Eq. (6.3.21). Using normal coordinates $\mathbf{R} = \mathrm{Q}^{\mathrm{u}}\mathbf{X}^{\mathrm{u}}$ (from now on omit index u in $\mathbf{X}^{\mathrm{u}}$), the independence of modes, and that the mean value vanishes, we obtain

$$\begin{aligned} \left\langle \left[\vec{R}_t^{\alpha} - \vec{R}_0^{\beta} \right]^2 \right\rangle_{\mathrm{u}} &= \left\langle \left[\sum_{\gamma=1}^{N} \left(Q^{\mathrm{u}}_{\alpha\gamma} \vec{X}_t^{\gamma} - Q^{\mathrm{u}}_{\beta\gamma} \vec{X}_0^{\gamma} \right) \right]^2 \right\rangle_{\mathrm{u}} \\ &= \sum_{\gamma=1}^{N} \left\langle \left(Q^{\mathrm{u}}_{\alpha\gamma} \vec{X}_t^{\gamma} - Q^{\mathrm{u}}_{\beta\gamma} \vec{X}_0^{\gamma} \right)^2 \right\rangle_{\mathrm{u}} . \end{aligned} \tag{6.3.47}$$

The center-of-mass motion with $D_{\mathrm{COM}} = D/N$ is obtained from $Q^{\mathrm{u}}_{\alpha 1} = Q^{\mathrm{u}}_{\beta 1} = N^{-1/2}$. Using Eqs. (6.3.45)-(6.3.46) we directly obtain the result Eq. (6.3.21),

$$\left\langle \left(\vec{R}_t^{\alpha} - \vec{R}_0^{\beta} \right)^2 \right\rangle_{\mathrm{u}} = 6 D_{\mathrm{COM}} t + \sum_{\gamma=2}^{N} V_{\mathrm{u}}^{\gamma} \left[(Q^{\mathrm{u}}_{\alpha\gamma})^2 + (Q^{\mathrm{u}}_{\beta\gamma})^2 - 2 Q^{\mathrm{u}}_{\alpha\gamma} Q^{\mathrm{u}}_{\beta\gamma} \mathrm{e}^{-a^{\mathrm{u}}_{\gamma} t} \right] . \tag{6.3.48}$$

We now consider the process with a single jump at a fixed time $\tau < t$, i.e., we start in the equilibrium of the unfolded u-state and propagate in the u-state from time 0 to τ and in the folded f-state from time τ to t. Averages with respect to this process will be denoted by $\langle \cdot \rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}}$. Within the Gaussian network model the center-of-mass motion is independent of the network structure, and we obtain similar to above (but now the modes are mixed due to $\mathrm{Q}^{\mathrm{u}} \neq \mathrm{Q}^{\mathrm{f}}$ and thus no longer decouple),

$$\begin{aligned} \left\langle \left(\vec{R}_t^{\alpha} - \vec{R}_0^{\beta} \right)^2 \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} &= 6 D_{\mathrm{COM}} t + \sum_{\gamma,\nu=2}^{N} \Big[Q^{\mathrm{u}}_{\alpha\gamma} Q^{\mathrm{u}}_{\alpha\nu} \left\langle (\mathbf{X}_t^{\mathrm{f}})_{\gamma} \cdot (\mathbf{X}_t^{\mathrm{f}})_{\nu} \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} \\ &+ Q^{\mathrm{u}}_{\beta\gamma} Q^{\mathrm{u}}_{\beta\nu} \left\langle (\mathbf{X}_0^{\mathrm{f}})_{\gamma} \cdot (\mathbf{X}_0^{\mathrm{f}})_{\nu} \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} - 2 Q^{\mathrm{u}}_{\alpha\gamma} Q^{\mathrm{u}}_{\beta\nu} \left\langle (\mathbf{X}_t^{\mathrm{f}})_{\gamma} \cdot (\mathbf{X}_0^{\mathrm{f}})_{\nu} \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} \Big] . \end{aligned} \tag{6.3.49}$$

New J. Phys. **24** 023004 (2022)

Now we determine the three terms $\left\langle (\mathbf{X}_t^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_t^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}}$, $\left\langle (\mathbf{X}_0^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_0^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}}$ and $\left\langle (\mathbf{X}_t^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_0^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}}$. The first term resembles $\left\langle (\mathbf{X}_t^{\mathrm{f}})_\gamma (\mathbf{X}_t^{\mathrm{f}})_\nu \right\rangle_{\mathrm{f}} = \delta_{\gamma\nu} V_{\mathrm{f}}^\gamma$, with the difference that we do not start in the f-equilibrium but propagate the f-process from the u-equilibrium from time τ to t. In particular, modes are mixed since correlations at time τ are only diagonal in u-normal modes $\mathbf{X}^{\mathrm{u}}$ but not in $\mathbf{X}^{\mathrm{f}}$. Using Eq. (6.3.46) we have

$$\left\langle (\mathbf{X}_t^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_t^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} = \left\langle (\mathbf{X}_\tau^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_\tau^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} \mathrm{e}^{-(a_\gamma + a_\nu)(t-\tau)} + \delta_{\gamma\nu} V^\gamma (t-\tau). \tag{6.3.50}$$

Since we start in u-equilibrium and propagate in the u-process until time τ, we complete the calculation of the first term by calculating the second term, by expressing f-modes in terms of u-modes. Using the notation $\mathrm{Q}^{\mathrm{u}\to\mathrm{f}} \equiv (\mathrm{Q}^{\mathrm{f}})^T \mathrm{Q}^{\mathrm{u}}$ (such that as in $\mathbf{R} = \mathrm{Q}^{\mathrm{u}}\mathbf{X}^{\mathrm{u}} = \mathrm{Q}^{\mathrm{f}}\mathbf{X}^{\mathrm{f}}$ we have $\mathbf{X}^{\mathrm{f}} = \mathrm{Q}^{\mathrm{u}\to\mathrm{f}}\mathbf{X}^{\mathrm{u}}$),

$$\begin{aligned}\left\langle (\mathbf{X}_\tau^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_\tau^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} &= \left\langle (\mathbf{X}_0^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_0^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} = \left\langle (\mathbf{X}_0^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_0^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}} \\ &= \sum_{\omega=2}^{N} Q_{\gamma\omega}^{\mathrm{u}\to\mathrm{f}} Q_{\nu\omega}^{\mathrm{u}\to\mathrm{f}} V_{\mathrm{u}}^\omega .\end{aligned} \tag{6.3.51}$$

For the third term we employ Eq. (6.3.45) for the f-process from time τ to t and for the u-process from 0 to τ,

$$\begin{aligned}\left\langle (\mathbf{X}_t^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_0^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} &= \mathrm{e}^{-a_\gamma^{\mathrm{f}}(t-\tau)} \left\langle (\mathbf{X}_\tau^{\mathrm{f}})_\gamma \cdot (\mathbf{X}_0^{\mathrm{f}})_\nu \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} \\ &= \mathrm{e}^{-a_\gamma^{\mathrm{f}}(t-\tau)} \sum_{\omega,\omega'=2}^{N} Q_{\gamma\omega}^{\mathrm{u}\to\mathrm{f}} Q_{\nu\omega'}^{\mathrm{u}\to\mathrm{f}} \left\langle (\mathbf{X}_\tau^{\mathrm{u}})_\omega \cdot (\mathbf{X}_0^{\mathrm{u}})_{\omega'} \right\rangle_{\mathrm{u}} \\ &= \mathrm{e}^{-a_\gamma^{\mathrm{f}}(t-\tau)} \sum_{\omega=2}^{N} Q_{\gamma\omega}^{\mathrm{u}\to\mathrm{f}} Q_{\nu\omega}^{\mathrm{u}\to\mathrm{f}} V_{\mathrm{u}}^\omega \mathrm{e}^{-a_\omega^{\mathrm{u}}\tau}.\end{aligned} \tag{6.3.52}$$

Combining Eqs. (6.3.49)-(6.3.52) and using $\sum_{\gamma\nu} C_\gamma C_\nu = (\sum_\gamma C_\gamma)^2$, and for the last term $\mathrm{Q}^{\mathrm{f}}\mathrm{Q}^{\mathrm{u}\to\mathrm{f}} = \mathrm{Q}^{\mathrm{u}}$, we arrive at the result Eq. (6.3.22),

$$\begin{aligned}&\left\langle \left(\vec{R}_t^\alpha - \vec{R}_0^\beta\right)^2 \right\rangle_{\mathrm{u}\xrightarrow{\tau}\mathrm{f}} = 6D_{\mathrm{COM}} t + \sum_{\gamma=2}^{N} (Q_{\beta\gamma}^{\mathrm{u}})^2 V_{\mathrm{u}}^\gamma + \sum_{\gamma=2}^{N} (Q_{\alpha\gamma}^{\mathrm{f}})^2 V_{\mathrm{f}}^\gamma (t-\tau) \\ &+ \sum_{\omega=2}^{N} V_{\mathrm{u}}^\omega \left[\sum_{\gamma=2}^{N} Q_{\alpha\gamma}^{\mathrm{f}} Q_{\gamma\omega}^{\mathrm{u}\to\mathrm{f}} \mathrm{e}^{-a_\gamma^{\mathrm{f}}(t-\tau)}\right]^2 - 2 \sum_{\omega,\gamma=2}^{N} Q_{\alpha\omega}^{\mathrm{f}} Q_{\omega\gamma}^{\mathrm{u}\to\mathrm{f}} Q_{\beta\gamma}^{\mathrm{u}} V_{\mathrm{u}}^\gamma \mathrm{e}^{-a_\omega^{\mathrm{f}}(t-\tau) - a_\gamma^{\mathrm{u}}\tau}.\end{aligned} \tag{6.3.53}$$

New J. Phys. **24** 023004 (2022)

Finally, for completeness we here show the Kirchoff matrix A^{f} corresponding to the Trp-Cage Miniprotein Construct TC5b (PDB: 1L2Y)

$$\frac{\xi \mathrm{A}^{\mathrm{f}}}{k} = \begin{bmatrix}
6 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 9 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\
-1 & -1 & 11 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\
-1 & -1 & -1 & 10 & -1 & -1 & -1 & -1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
-1 & -1 & -1 & -1 & 11 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
-1 & -1 & -1 & -1 & -1 & 16 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & -1 & 0 & -1 & -1 & -1 & -1 & 0 \\
-1 & -1 & -1 & -1 & -1 & -1 & 14 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & -1 & -1 & -1 & -1 & -1 & 9 & -1 & -1 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & 12 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & 11 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 15 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & -1 & -1 & 11 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & -1 & -1 & 9 & -1 & -1 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 11 & -1 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & 8 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 11 & -1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & 10 & -1 & -1 & -1 \\
0 & -1 & -1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & -1 & -1 & 10 & -1 & -1 \\
0 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 9 & -1 \\
0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 5
\end{bmatrix}, \tag{6.3.54}$$

and the 20×20 tridiagonal Kirchoff matrix A^{u} for the unfolded Rouse chain

$$\mathrm{A}^{\mathrm{u}} = \frac{k}{\xi} \begin{bmatrix}
1 & -1 & & & & & & \\
-1 & 2 & -1 & & & & 0 & \\
 & -1 & 2 & -1 & & & & \\
 & & & \dots & \dots & & & \\
 & & & & \dots & \dots & & \\
 & 0 & & & -1 & 2 & -1 & \\
 & & & & & -1 & 1 &
\end{bmatrix}. \tag{6.3.55}$$

New J. Phys. **24** 023004 (2022)

Chapter 7

Conclusion

To conclude the thesis, we summarize all results and contributions to the literature presented in Chs. 3-6, and give an outlook for future research for each topic, respectively.

7.1 Time-integrated observables and necessity of coarse graining

In Ch.3, we introduced different approaches to address the random fluctuations of time-integrated observables, which are important observables for thermodynamic inference from measurements of stochastic trajectories, especially, in non-equilibrium steady states (NESS). Such trajectories are the result of, e.g., single-molecule or particle tracking experiments. Using a stochastic-calculus approach based on the fundamental methods developed in Ch. 2, we derived mathematical expressions for the (co)variances of, and correlations between, time-integrated densities and currents. In Sec. 3.5, we interpreted the obtained expressions using a generalized time-reversal symmetry, the *dual-reversal symmetry*, and in Sec. 3.6 we generalized the results beyond steady-state dynamics. To make the connection between the stochastic-calculus approach and existing methods like Feynman-Kac theory and functional calculus, we re-derived the new results also from these established methods in Sec. 3.7.

Most importantly, the results derived in Secs. 3.4-3.7 allowed to systematically investigate the role of coarse graining in space. The coarse graining may account for a finite spatial resolution (see, e.g., Fig. 3.5.2), but typically we refer to coarse graining as a means of processing measured data on chosen length scales, such as choosing the width of bins in a

histogram, see Fig. 3.5.1. In this sense, coarse graining can smoothen data by deliberately decreasing the precision. Importantly, we find that this smoothening does not necessarily result in mere loss of precision and inference quality. Instead, we found the opposite, namely that it has many positive aspects. Furthermore, some aspects make it even *necessary*, see Sec. 3.4. First, in multidimensional space, the coarse graining is strictly necessary to obtain finite variances, i.e., inference without coarse graining would not even be possible in the idealized case of unlimited statistics. Second, it is intuitively clear that for finite statistics, some coarse graining is required to avoid or mitigate undersampling. These two aspects give rise to an *experimental necessity* of coarse graining. Third, the coarse graining appears to be *mathematically necessary* to obtain well-defined observables, and to pose well-defined questions about their statistics, as stressed in Secs. 3.2-3.5. Finally, the coarse graining was also found to be *necessary for thermodynamic inference* of entropy production using the thermodynamic uncertainty relation (TUR), see Secs. 3.4 and 3.5.

We now give an outlook for future research based on the results summarized above. There are two major open questions that directly come to mind.

First, we often motivated the considered observables and insight into coarse graining in the sense that they are relevant for experimental inference. Though this seems to be well justified in theory, it has yet to be proven for actual experiments. So far, we only demonstrated the necessity of coarse graining for experimental data in Fig. 3.5.9, but have not applied any of the results to inference in practice.

Second, we discussed on many occasions the mathematical necessity of coarse graining, which means that mathematical problems occur if the empirical density and current are defined with a delta function, see Eq. (3.2.18). In Appendix II in Sec. 3.4 and in Subsec. 3.5.9, we stressed that the divergent variance in $d \geq 2$-dimensional space in the absence of coarse graining prevents any large deviation principle that includes a central limit regime with a finite variance. This raises one important question: How can our results be reconciled with the celebrated Donsker-Varadhan rate functional, see Eq. (3.2.20), and with level 2.5 large deviation theory, see Eq. (3.2.21)? Or, to put it plainly: Do our results imply that the Donsker-Varadhan large deviation principle is wrong in $d \geq 2$-dimensional space?

The simple answer is that our results suggest that the empirical density and current do not exist as well-defined random variables if $d \geq 2$ when defined with a delta function as in Eq. (3.2.18). Therefore, one cannot make any mathematical statement about them, which renders the Donsker-Varadhan result as presented in Refs. [45, 204] an empty statement. From a physical perspective, however, this simple answer is unsatisfying. The relevant ques-

tion becomes, whether the rate functional correctly describes the dynamical fluctuations when interpreted in some way that avoids the aforementioned mathematical problems. To address this question, we now sketch some unpublished results that make a quantitative connection between our results and the Donsker-Varadhan rate functional in Eq. (3.2.20).

Consider d-dimensional Brownian motion with diffusion constant $D = 1$ confined in a box $[0,1]^d$ such that $p_\mathrm{s}(\mathbf{x}) = 1$ for all $\mathbf{x} \in [0,1]^d$. The simplest observable to describe expected fluctuations of the density field $\rho_t(\cdot)$, where $\rho_t(\mathbf{x}) \equiv \int_0^t \delta(\mathbf{x}_\tau - \mathbf{x})\mathrm{d}\tau/t$, is the integrated squared deviation, or L_2-variance, of the field $\rho_t(\cdot)$,

$$V_\rho(t) \equiv \left\langle \int \mathrm{d}\mathbf{x} \left[\rho_t(\mathbf{x}) - p_\mathrm{s}(\mathbf{x})\right]^2 \right\rangle . \tag{7.1.1}$$

By interchanging the order of integration and taking the average (for justification see Subsec. 2.1.4), we see that $V_\rho(t)$ is the integral over individual-point variances, which diverge for $d \geq 2$, see first line of Eq. (3.4.6) for $h = 0$. Thus, our results predict $V_\rho(t) = \infty$ for $d \geq 2$. Before we discuss this result further, we also compute $V_\rho(t)$ from the Donsker-Varadhan rate functional. For Brownian motion in a box, the rate functional (3.2.20) becomes $I[\rho(\cdot)] \propto \int_{[0,1]^d} \mathrm{d}\mathbf{x}[\nabla\rho(\mathbf{x})]^2/\rho(\mathbf{x})$. To obtain the second cumulant in Eq. (7.1.1) it suffices to consider a functional Taylor expansion up to second order around $\rho = p_\mathrm{s} = 1$, which gives approximately $I[\rho(\cdot)] \propto \int_{[0,1]^d} \mathrm{d}\mathbf{x}[\nabla\rho(\mathbf{x})]^2$. By writing the average in Eq. (7.1.1) as a functional integral weighted by this rate functional, and transforming to Fourier space, one can show that $V_\rho(t) \propto \sum_{k_1,\ldots,k_d=1}^{\infty} (\sum_{m=1}^{d} k_m^2)^{-1} = \infty$ for[1] $d \geq 2$. Thus, our approach and the Donsker-Varadhan rate functional consistently give the same result for $V_\rho(t)$. Based on this preliminary result, there is no contradiction between our result and the Donsker-Varadhan large deviation principle, i.e., possibly both theories can be reconciled by interpreting the latter in some well-defined sense (e.g., by coarse graining or by *contracting* (see Subsec. 3.2.3) to one-dimensional variables).

However, we can already deduce from the preliminary result $V_\rho(t) = \infty$ that the standard interpretation of the large deviation principle does *not* apply to the empirical density defined with a delta function. This interpretation, in particular asserted to hold for time-integrated observables [43], states that there is a single zero of the rate function(al), which reflects the mean and the most likely point, and that the central-limit theorem is included as the regime of small deviations around the mean, see Subsec. 3.2.3 and Refs. [43, 110]. In the case $V_\rho(t) = \infty$, however, already these small deviations diverge, and the central limit theorem does *not* apply. Although the rate functional is non-negative with a single

[1] For $d = 1$, the sum converges, $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$, and we obtain $V_\rho(t) = 1/3t$ for long times from both, the Taylor expansion of the Donsker-Varadhan rate functional and integration over single-point variances.

zero for $\rho = p_{\mathrm{s}}$, this does not reflect the most likely point for $d \geq 2$ since for any $t < \infty$, the trajectory $(\mathbf{x}_\tau)_{0\leq\tau\leq t}$ cannot hit all points in the d-dimensional space, which would be required for $\rho_t = p_{\mathrm{s}}$.

Future work will have to clarify if and how the, in principle, very powerful results of the Donsker-Varadhan and level 2.5 large deviation principles, see Eqs. (3.2.20) and (3.2.21), can be applied to multidimensional systems.

7.2 Thermodynamic uncertainty relations

While TURs are inequalities involving the current variance, see Ch. 4, we derived an explicit equality for the current variance in Ch. 3. Since the approach in Ch. 3 applies to any time-integrated current for overdamped Langevin dynamics, we wondered how to approach the TUR from the developed methods in a more general sense, i.e., beyond the notion of coarse graining. To this end, we rederived the TUR from the stochastic-calculus approach developed in Chs. 2 and 3. The new derivation was very direct, and established the TUR as an inherent property of Langevin dynamics, following immediately from an average over fluctuating trajectories, see Eqs. (4.5.10) and (4.5.11). Importantly, it allowed us to find sharper versions of TURs, and in particular, to theoretically saturate a TUR for time-homogeneous systems with arbitrary initial conditions, see Sec. 4.5 and Fig. 4.5.1. This shows that, for fully observed dynamics starting from some initial condition, the generalized correlation TUR in Eq. (4.5.22) is a very suitable tool for inference of entropy production.

As for the results for the time-integrated density and current, a direct objective of future research is the application to experimental data to infer entropy production from observed dynamics. There, it will be particularly interesting to which extent the time derivative $t\partial_t\langle J_t\rangle$, e.g., in Eq. (4.5.22), becomes problematic in regard to sufficient sampling, since terms involving time derivatives are expected to require significantly more statistics than, e.g., $\langle J_t\rangle$.

Apart from applications, also several possible generalizations come to mind. We mentioned how to generalize the approach to Markov jump dynamics in Subsec. 4.5.12. Future work could complete this approach to derive the results, such as the generalized transient TUR (4.5.22), also for Markov-jump dynamics. Further generalizations could follow in a similar direction as the TUR for indistinguishable species in Eq. (4.1.4). A promising strategy for such generalizations is to consider the equality $\langle A_t(J_t - \langle J_t\rangle)\rangle = \langle J_t\rangle + \langle A_t J_t^{\mathrm{II}}\rangle$ in Eq. (4.5.10), and to take another average (such as the one over the distinct species) over this

equality before applying the Cauchy-Schwarz inequality. This not only gives an alternative proof for the TUR for indistinguishable species in Eq. (4.1.4), but more importantly presents a promising strategy to derive TURs for complex dynamics, such as multi-state dynamics (see Ch. 6), or *diffusing diffusivity* systems [357,358], where the diffusion coefficient itself is a stochastic quantity.

7.3 Thermal relaxation far from equilibrium

In Ch. 5, we shifted from the topic of time-integrated currents to thermal relaxation. As described in detail in Secs. 5.1 and 5.2, it was known from the conventional framework of linear irreversible thermodynamics, that thermal relaxation is symmetric in the sign of the temperature change up to linear order, and recently, it was predicted that for paradigmatic examples of overdamped motion beyond this linear regime, heating is always faster than cooling [73]. Building on the framework and results of Ref. [73], we investigated the asymmetry between heating and cooling in more detail. We experimentally confirmed the predicted asymmetry using an optically trapped particle, see Sec. 5.4. Although the quantification of thermal relaxation using the *excess free energy* has several advantages, in particular its thermodynamic interpretation, it does not allow comparing the thermal relaxation between a pair of arbitrary temperatures, see Sec. 5.2. As a remedy to this shortcoming, we developed a framework that we called *thermal kinematics*. This framework, unlike the approach using the excess free energy, allows to define *distances* and *velocities* for relaxation, and in particular, it allows to consistently compare heating and cooling between any pair of temperatures. We predicted, and experimentally confirmed, that also in this framework, heating is indeed faster than cooling, see Sec. 5.4.

To provide some intuition about the faster heating, we noted that heating starts from a more localized initial condition than cooling, which in a spectral representation of the process implies more weight on the higher, faster-decaying modes. For heating and cooling between two arbitrary temperatures, we furthermore noted that the heating process occurs at a higher temperature than the cooling process.

Another generalization of the thermal relaxation asymmetry was established in Sec. 5.5, where we proved that, in terms of the (generalized) excess free energy, heating is also faster than cooling for driven systems that relax into an NESS. The general insight into linearly driven diffusion gained in Sec. 5.5 also lead to observations relevant beyond thermal relaxation, such as the role and geometrical intuition of complex eigenvectors and eigenvalues (see, e.g., Fig. 5.5.2), and in particular the phenomenon of effective rotations that surprisingly oppose the direction of driving, see Figs. 5.5.2 and 5.5.4-5.5.7.

We now come to an outlook for further research. While heating is always faster than cooling in terms of the excess free energy in harmonic potentials, counterexamples to this asymmetry are known, e.g., for multiwell potentials [73]. It would be interesting to study these counterexamples through the lens of the thermal-kinematics framework developed in Sec. 5.4. Moreover, connecting the relaxation asymmetry to time-dependent processes, in the spirit of the Brownian Carnot cycle [275] and its optimization, represents an intriguing question.

Very concrete objectives for future research based on Secs. 5.4 and 5.5 are the generalization of the relaxation asymmetry to underdamped dynamics (i.e., in the presence of inertial effects), to generalize the thermal-kinematics approach to relaxation towards NESS as considered in Sec. 5.5, and to experimentally confirm the relaxation asymmetry for relaxation towards NESS as predicted in Sec. 5.5. These experiments are indeed planned.

Overall, we are really still at the beginning of understanding thermal relaxation, and relaxation in general, in systems far from equilibrium. The complexity of relaxation in Langevin systems is highlighted by the opposing rotations observed in Sec. 5.5, that illustrate how complicated and counterintuitive even the supposedly simple model of Langevin dynamics with a linear drift can behave.

7.4 Two-state dynamics

In addition to the analysis of time-averaged observables and thermal relaxation of time-homogeneous dynamics, we addressed time-heterogeneous[2] *two-state dynamics* in Ch. 6. For several physically relevant examples of two-state dynamics, we computed the *intermediate scattering function*. Since it is accessible from different microscopy and scattering experiments, or from measured trajectories, this observable is suitable to resolve for two-state dynamics in experiments. To give a concrete example for the application of our results, we fitted the results for two-state fractional Brownian motion to particle tracking data that was recorded in living cells in Ref. [91]. We thereby provided further support to the interpretation of complex intracellular motion as a nominally time-heterogeneous two-state process [91].

In all examples, we found that the intermediate scattering function decays faster in time for increasing switching rates, which is consistent with the idea of an additional relaxation

[2] *Time-heterogeneous* here refers to the fact that the equations of motion of the observable may change in time because of an explicit dependence on the instantaneous state of a hidden degree of freedom reflecting the dynamical state. The motion of all degrees of freedom, observed and hidden, is still time-homogeneous.

channel enabling a faster decay [398, 399]. It would be interesting to see whether this is a general feature of the intermediate scattering function for multi-state dynamics, which appears to be likely, even though it directly contradicts the empirical ansatz in Ref. [383].

The theory developed in Sec. 6.3 can be extended to multi-state dynamics and to non-Markovian waiting-time statistics in the respective states. Moreover, an objective for future research is to apply the results to experimental data, either from microscopy or scattering experiments, or, as in Fig. 6.3.2, to Fourier transformed data from particle tracking experiments.

7.5 Final remarks

Altogether, we addressed several aspects of non-equilibrium statistical mechanics, with particular emphasis on overdamped stochastic dynamics studied through the lens of stochastic calculus. Our investigations encompassed analyses of time-integrated, coarse-grained densities and currents, as well as the study of thermal relaxation in Langevin systems and of optically trapped particles. Additionally, we delved into two-state dynamics, with an application to random motion in living cells. While we obtained interesting results and new insights, what remains striking is the vast number of remaining challenges and unresolved questions in non-equilibrium physics, including very fundamental ones, that will undoubtedly captivate our attention for many more years to come.

Bibliography

[1] R. Brown, "XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies," *Philos. Mag.* **4** (1828) 161.

[2] A. Einstein, "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen," *Ann. Phys. (Leipzig)* **322** (1905) 549.

[3] M. von Smoluchowski, "Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen," *Ann. Phys. (Leipzig)* **326** (1906) 756.

[4] J. Perrin, *Atoms*. Ox Bow, Woodbridge, Conn., 1990 [originally 1913].

[5] J. Perrin, "Discontinuous structure of matter, Nobel lecture," 1926.

[6] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer-Verlag, Berlin New York, second ed., 1985.

[7] N. van Kampen, *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library. Elsevier, Amsterdam, third ed., 2007.

[8] P. Langevin, "Sur la théorie du mouvement brownien," *C. R. Acad. Sci. (Paris)* **146** (1908) 530.

[9] R. Zwanzig, *Nonequilibrium statistical mechanics*. Oxford University Press (OUP), 2001.

[10] H. Mori, "Transport, collective motion, and Brownian motion," *Prog. Theor. Phys.* **33** (1965) 423.

[11] G. Wilemski, "On the derivation of Smoluchowski equations with corrections in the classical theory of Brownian motion," *J. Stat. Phys.* **14** (1976) 153.

[12] N. Wiener, "The average of an analytic functional and the Brownian movement," *Proc. Natl. Acad. Sci. U.S.A.* **7** (1921) 294.

[13] A. Fick, "Über Diffusion," *Ann. Phys. (Leipzig)* **170** (1855) 59.

[14] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*. North Holland, 1st edition ed., 1981.

[15] B. Øksendal, *Stochastic Differential Equations*. Springer Berlin Heidelberg, 2003.

[16] G. A. Pavliotis, *Stochastic Processes and Applications*. Springer New York, 2014.

[17] H. Risken, *The Fokker-Planck Equation*. Springer Berlin Heidelberg, Sept., 1989.

[18] D. T. Gillespie, "The chemical Langevin equation," *J. Chem. Phys.* **113** (2000) 297.

[19] J.-P. Bouchaud and R. Cont, "A Langevin approach to stock market fluctuations and crashes," *Eur. Phys. J. B.* **6** (1998) 543.

[20] R. van Zon, S. Ciliberto, and E. G. D. Cohen, "Power and heat fluctuation theorems for electric circuits," *Phys. Rev. Lett.* **92** (2004) 130601.

[21] R. Kubo, "The fluctuation-dissipation theorem," *Rep. Prog. Phys.* **29** (1966) 255.

[22] R. Kubo, "Brownian motion and nonequilibrium statistical mechanics," *Science* **233** (1986) 330.

[23] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II : Nonequilibrium Statistical Mechanics*. Springer Berlin Heidelberg, 1991.

[24] R. Kubo, "Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems," *J. Phys. Soc. Jpn.* **12** (1957) 570.

[25] S. R. de Groot and P. Mazur, *Non-equilibrium Thermodynamics*. North-Holland, Amsterdam, 2nd ed. ed., 1962.

[26] L. Onsager, "Reciprocal relations in irreversible processes. I.," *Phys. Rev.* **37** (1931) 405.

[27] L. Onsager, "Reciprocal Relations in Irreversible Processes. II.," *Phys. Rev.* **38** (1931) 2265.

[28] U. Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," *Rep. Prog. Phys.* **75** (2012) 126001.

[29] F. Jülicher, A. Ajdari, and J. Prost, "Modeling molecular motors," *Rev. Mod. Phys.* **69** (1997) 1269.

[30] J. Parrondo and B. de Cisneros, "Energetics of Brownian motors: a review," *Appl. Phys. A* **75** (2002) 179.

[31] A. B. Kolomeisky and M. E. Fisher, "Molecular motors: A theorist's perspective," *Annu. Rev. Phys. Chem.* **58** (2007) 675.

[32] S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, "Nonequilibrium energetics of a single $\mathbf{f}_1$-ATPase molecule," *Phys. Rev. Lett.* **104** (2010) 198103.

[33] S. N. Majumdar, "Brownian functionals in physics and computer science," *Curr. Sci.* **89** (2005) 2075.

[34] C. Maes, K. Netočný, and B. Wynants, "Steady state statistics of driven diffusions," *Physica (Amsterdam)* **387A** (2008) 2675.

[35] S. N. Majumdar and A. J. Bray, "Large-deviation functions for nonlinear functionals of a Gaussian stationary Markov process," *Phys. Rev. E* **65** (2002) 051112.

[36] S. N. Majumdar and D. S. Dean, "Exact occupation time distribution in a non-Markovian sequence and its relation to spin glass models," *Phys. Rev. E* **66** (2002) 041102.

[37] A. Rebenshtok and E. Barkai, "Weakly non-ergodic statistical physics," *J. Stat. Phys.* **133** (2008) 565.

[38] G. Bel and E. Barkai, "Occupation times and ergodicity breaking in biased continuous time random walks," *J. Phys. Condens. Matter* **17** (2005) S4287.

[39] S. Burov and E. Barkai, "Occupation time statistics in the quenched trap model," *Phys. Rev. Lett.* **98** (2007) 250601.

[40] S. Carmi and E. Barkai, "Fractional Feynman-Kac equation for weak ergodicity breaking," *Phys. Rev. E* **84** (2011) 061104.

[41] A. Lapolla, D. Hartich, and A. Godec, "Spectral theory of fluctuations in time-average statistical mechanics of reversible and driven systems," *Phys. Rev. Res.* **2** (2020) 043084.

[42] M. Kac, "On distributions of certain Wiener functionals," *Trans. Am. Math. Soc.* **65** (1949) 1.

[43] H. Touchette, "Introduction to dynamical large deviations of Markov processes," *Physica (Amsterdam)* **504A** (2018) 5.

[44] A. C. Barato and R. Chetrite, "A formal view on level 2.5 large deviations and fluctuation relations," *J. Stat. Phys.* **160** (2015) 1154.

[45] J. Hoppenau, D. Nickelsen, and A. Engel, "Level 2 and level 2.5 large deviation functionals for systems with and without detailed balance," *New J. Phys.* **18** (2016) 083010.

[46] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, I," *Commun. Pure. Appl. Math.* **28** (1975) 1.

[47] A. C. Barato and U. Seifert, "Thermodynamic uncertainty relation for biomolecular processes," *Phys. Rev. Lett.* **114** (2015) 158101.

[48] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, "Dissipation bounds all steady-state current fluctuations," *Phys. Rev. Lett.* **116** (2016) 120601.

[49] T. Koyuk and U. Seifert, "Thermodynamic uncertainty relation for time-dependent driving," *Phys. Rev. Lett.* **125** (2020) 260604.

[50] L. Peliti and S. Pigolotti, *Stochastic thermodynamics*. Princeton University Press, Princeton, NJ, 2021.

[51] K. Sekimoto, "Kinetic characterization of heat bath and the energetics of thermal ratchet models," *J. Phys. Soc. Japan* **66** (1997) 1234.

[52] K. Sekimoto, "Langevin equation and thermodynamics," *Prog. Theor. Phys. Supp.* **130** (1998) 17.

[53] K. Sekimoto, *Stochastic Energetics*. Springer Berlin Heidelberg, 2010.

[54] U. Seifert, "Entropy production along a stochastic trajectory and an integral fluctuation theorem," *Phys. Rev. Lett.* **95** (2005) 040602.

[55] C. Maes, "Local detailed balance," *SciPost Phys. Lect. Notes* (2021) 32.

[56] C. Jarzynski, "Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale," *Annu. Rev. Condens. Matter Phys.* **2** (2011) 329.

[57] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, "Probability of second law violations in shearing steady states," *Phys. Rev. Lett.* **71** (1993) 2401.

[58] G. Gallavotti and E. G. D. Cohen, "Dynamical ensembles in nonequilibrium statistical mechanics," *Phys. Rev. Lett.* **74** (1995) 2694.

[59] J. Kurchan, "Fluctuation theorem for stochastic dynamics," *J. Phys. A: Math. Gen.* **31** (1998) 3719.

[60] J. L. Lebowitz and H. Spohn, "A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics," *J. Stat. Phys.* **95** (1999) 333.

[61] C. Jarzynski, "Nonequilibrium equality for free energy differences," *Phys. Rev. Lett.* **78** (1997) 2690.

[62] C. Jarzynski, "Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach," *Phys. Rev. E* **56** (1997) 5018.

[63] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, "Equilibrium information from nonequilibrium measurements in an experimental test of jarzynski's equality," *Science* **296** (2002) 1832.

[64] G. E. Crooks, "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences," *Phys. Rev. E* **60** (1999) 2721.

[65] G. E. Crooks, "Path-ensemble averages in systems driven far from equilibrium," *Phys. Rev. E* **61** (2000) 2361.

[66] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante, "Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies," *Nature* **437** (2005) 231.

[67] J. M. Horowitz and T. R. Gingrich, "Thermodynamic uncertainty relations constrain non-equilibrium fluctuations," *Nat. Phys.* **16** (2019) 15.

[68] J. Li, J. M. Horowitz, T. R. Gingrich, and N. Fakhri, "Quantifying dissipation using fluctuating currents," *Nat. Commun.* **10** (2019) 1666.

[69] U. Seifert, "From stochastic thermodynamics to thermodynamic inference," *Annu. Rev. Condens. Matter Phys.* **10** (2019) 171.

[70] A. Dechant, "Multidimensional thermodynamic uncertainty relations," *J. Phys. A: Math. Theor.* **52** (2018) 035001.

[71] K. Liu, Z. Gong, and M. Ueda, "Thermodynamic uncertainty relation for arbitrary initial states," *Phys. Rev. Lett.* **125** (2020) 140602.

[72] A. Dechant and S.-i. Sasa, "Fluctuation–response inequality out of equilibrium," *Proc. Natl. Acad. Sci. U.S.A.* **117** (2020) 6430.

[73] A. Lapolla and A. Godec, "Faster uphill relaxation in thermodynamically equidistant temperature quenches," *Phys. Rev. Lett.* **125** (2020) 110602.

[74] J. Meibohm, D. Forastiere, T. Adeleke-Larodo, and K. Proesmans, "Relaxation-speed crossover in anharmonic potentials," *Phys. Rev. E* **104** (2021) L032105.

[75] T. Van Vu and Y. Hasegawa, "Toward relaxation asymmetry: Heating is faster than cooling," *Phys. Rev. Res.* **3** (2021) 043160.

[76] S. K. Manikandan, "Equidistant quenches in few-level quantum systems," *Phys. Rev. Res.* **3** (2021) 043108.

[77] S. Ramaswamy, "The mechanics and statistics of active matter," *Annu. Rev. Condens. Matter Phys.* **1** (2010) 323.

[78] J. Elgeti, R. G. Winkler, and G. Gompper, "Physics of microswimmers—single particle motion and collective behavior: a review," *Rep. Prog. Phys.* **78** (2015) 056601.

[79] G. Gompper, R. G. Winkler, T. Speck, A. Solon, C. Nardini, F. Peruani, H. Löwen, R. Golestanian, U. B. Kaupp, L. Alvarez, T. Kiørboe, E. Lauga, W. C. K. Poon, A. DeSimone, S. Muiños-Landin, A. Fischer, N. A. Söker, F. Cichos, R. Kapral, P. Gaspard, M. Ripoll, F. Sagues, A. Doostmohammadi, J. M. Yeomans, I. S. Aranson, C. Bechinger, H. Stark, C. K. Hemelrijk, F. J. Nedelec, T. Sarkar, T. Aryaksama, M. Lacroix, G. Duclos, V. Yashunsky, P. Silberzan, M. Arroyo, and S. Kale, "The 2020 motile active matter roadmap," *J. Condens. Matter Phys.* **32** (2020) 193001.

[80] M. J. Bowick, N. Fakhri, M. C. Marchetti, and S. Ramaswamy, "Symmetry, thermodynamics, and topology in active matter," *Phys. Rev. X* **12** (2022) 010501.

[81] R. Metzler and J. Klafter, "The random walk's guide to anomalous diffusion: a fractional dynamics approach," *Phys. Rep.* **339** (2000) 1.

[82] F. Höfling and T. Franosch, "Anomalous transport in the crowded world of biological cells," *Rep. Prog. Phys.* **76** (2013) 046602.

[83] C. Dieball and A. Godec, "Mathematical, thermodynamical, and experimental necessity for coarse graining empirical densities and currents in continuous space," *Phys. Rev. Lett.* **129** (2022) 140601.

[84] C. Dieball and A. Godec, "Coarse graining empirical densities and currents in continuous-space steady states," *Phys. Rev. Res.* **4** (2022) 033243.

[85] C. Dieball and A. Godec, "On correlations and fluctuations of time-averaged densities and currents with general time-dependence," *J. Phys. A: Math. Theor.* **55** (2022) 475001.

[86] C. Dieball and A. Godec, "Feynman-Kac theory of time-integrated functionals: Itô versus functional calculus," *J. Phys. A: Math. Theor.* **56** (2023) 155002.

[87] C. Dieball and A. Godec, "Direct route to thermodynamic uncertainty relations and their saturation," *Phys. Rev. Lett.* **130** (2023) 087101.

[88] M. Ibáñez, C. Dieball, A. Lasanta, A. Godec, and R. A. Rica, "Heating and cooling are fundamentally asymmetric and evolve along distinct pathways," *Nat. Phys.* (2024) .

[89] C. Dieball, G. Wellecke, and A. Godec, "Asymmetric thermal relaxation in driven systems: Rotations go opposite ways," *Phys. Rev. Res.* **5** (2023) L042030.

[90] C. Dieball, D. Krapf, M. Weiss, and A. Godec, "Scattering fingerprints of two-state dynamics," *New J. Phys.* **24** (2022) 023004.

[91] A. Sabri, X. Xu, D. Krapf, and M. Weiss, "Elucidating the origin of heterogeneous anomalous diffusion in the cytoplasm of mammalian cells," *Phys. Rev. Lett.* **125** (2020) 058101.

[92] P. Billingsley, *Probability and Measure*. Wiley Series in Probability & Mathematical Statistics: Probability & Mathematical Statistics. John Wiley & Sons, Nashville, TN, 3 ed., May, 1995.

[93] I. A. Ibragimov, I. V. Linnik, and I. V. Linnik, *Independent and stationary sequences of random variables*. Groningen, Wolters-Noordhoff, 1971.

[94] J. Dedecker, P. Doukhan, G. Lang, L. R. J. Rafael, S. Louhichi, and C. Prieur, *Weak Dependence: With Examples and Applications*. Springer New York, 2007.

[95] W. Feller, *An Introduction to Probability Theory and its Applications. Vol. I and II*. Third edition. John Wiley & Sons Inc., New York, 1968.

[96] A. Lapolla and A. Godec, "Toolbox for quantifying memory in dynamics along reaction coordinates," *Phys. Rev. Res.* **3** (2021) L022018.

[97] W. Feller, "Non-Markovian processes with the semigroup property," *Ann. Math. Stat.* **30** (1959) 1252.

[98] J. A. R. Cruz and I. C. Diniz, "On weak and strong ergodicity," *J. Stat. Theory Pract.* **13** (2019) 28.

[99] J.-F. L. Gall, *Brownian Motion, Martingales, and Stochastic Calculus*. Springer International Publishing, 2016.

[100] U. G. Haussmann and E. Pardoux, "Time reversal of diffusions," *Ann. Probab.* **14** (1986) 1188.

[101] H. De and A. Godec, "Vorticity and level-set variations of invariant current bound steady-state dissipation," (2023) arXiv:2306.13647.

[102] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion," *Phys. Rev.* **36** (1930) 823.

[103] A. Berezhkovskii and A. Szabo, "Time scale separation leads to position-dependent diffusion along a slow coordinate," *J. Chem. Phys.* **135** (2011) 074108.

[104] S. Pigolotti, I. Neri, É. Roldán, and F. Jülicher, "Generic properties of stochastic entropy production," *Phys. Rev. Lett.* **119** (2017) 140604.

[105] R. E. Spinney and I. J. Ford, "Entropy production in full phase space for continuous stochastic dynamics," *Phys. Rev. E* **85** (2012) 051113.

[106] C. Maes and K. Netočný, "Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states," *Europhys. Lett.* **82** (2008) 30003.

[107] R. Chetrite and H. Touchette, "Nonequilibrium Markov processes conditioned on large deviations," *Ann. Henri Poincaré* **16** (2014) 2005.

[108] F. den Hollander, *Large Deviations*, vol. 14 of *Fields Institute Monograph*. AMS, Providence, RI, 2000.

[109] B. Derrida, "Non-equilibrium steady states: fluctuations and large deviations of the density and of the current," *J. Stat. Mech.* **2007** (2007) P07023.

[110] H. Touchette, "The large deviation approach to statistical mechanics," *Phys. Rep.* **478** (2009) 1.

[111] W. Bryc, "A remark on the connection between the large deviation principle and the central limit theorem," *Stat. Probab. Lett.* **18** (1993) 253.

[112] J. du Buisson and H. Touchette, "Dynamical large deviations of linear diffusions," *Phys. Rev. E* **107** (2023) 054111.

[113] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, II," *Commun. Pure. Appl. Math.* **28** (1975) 279.

[114] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, III," *Commun. Pure. Appl. Math.* **29** (1976) 389.

[115] J. Gladrow, N. Fakhri, F. C. MacKintosh, C. F. Schmidt, and C. P. Broedersz, "Broken detailed balance of filament dynamics in active networks," *Phys. Rev. Lett.* **116** (2016) 248301.

[116] F. S. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, "Broken detailed balance and non-equilibrium dynamics in living systems: a review," *Rep. Prog. Phys.* **81** (2018) 066601.

[117] F. Ritort, "Single-molecule experiments in biological physics: methods and applications," *J. Phys.: Cond. Matt.* **18** (2006) R531.

[118] W. J. Greenleaf, M. T. Woodside, and S. M. Block, "High-resolution, single-molecule measurements of biomolecular motion," *Annu. Rev. Biophys. Biomol. Struct.* **36** (2007) 171–190.

[119] J. R. Moffitt, Y. R. Chemla, S. B. Smith, and C. Bustamante, "Recent advances in optical tweezers," *Annu. Rev. Biochem.* **77** (2008) 205–228.

[120] S. Vaikuntanathan and C. Jarzynski, "Dissipation and lag in irreversible processes," *Europhys. Lett.* **87** (2009) 60005.

[121] C. Maes, K. Netočný, and B. Wynants, "Monotonic return to steady nonequilibrium," *Phys. Rev. Lett.* **107** (2011) 010601.

[122] H. Qian, "A decomposition of irreversible diffusion processes without detailed balance," *J. Math. Phys. (N.Y.)* **54** (2013) 053302.

[123] C. Maes, "Frenetic bounds on the entropy production," *Phys. Rev. Lett.* **119** (2017) 160601.

[124] N. Shiraishi and K. Saito, "Information-theoretical bound of the irreversibility in thermal relaxation processes," *Phys. Rev. Lett.* **123** (2019) 110603.

[125] D.-Q. Jiang, M. Qian, and M.-P. Qian, *Mathematical Theory of Nonequilibrium Steady States*. Springer Berlin Heidelberg, 2004.

[126] A. Dechant and S.-i. Sasa, "Current fluctuations and transport efficiency for general Langevin systems," *J. Stat. Mech.* (2018) 063209.

[127] U. Seifert and T. Speck, "Fluctuation-dissipation theorem in nonequilibrium steady states," *Europhys. Lett.* **89** (2010) 10007.

[128] E. N. M. Cirillo, M. Colangeli, O. Richardson, and L. Rondoni, "Deterministic model of battery, uphill currents, and nonequilibrium phase transitions," *Phys. Rev. E* **103** (2021) 032119.

[129] D. J. Evans, S. R. Williams, D. J. Searles, and L. Rondoni, "On typicality in nonequilibrium steady states," *J. Stat. Phys.* **164** (2016) 842.

[130] S. Burov, J.-H. Jeon, R. Metzler, and E. Barkai, "Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking," *Phys. Chem. Chem. Phys.* **13** (2011) 1800.

[131] H. Qian, "Nonequilibrium steady-state circulation and heat dissipation functional," *Phys. Rev. E* **64** (2001) 022101.

[132] T. R. Gingrich, G. M. Rotskoff, and J. M. Horowitz, "Inferring dissipation from current fluctuations," *J. Phys. A: Math. Theor.* **50** (2017) 184004.

[133] A. Dechant and S.-i. Sasa, "Continuous time reversal and equality in the thermodynamic uncertainty relation," *Phys. Rev. Res.* **3** (2021) L042012.

[134] A. Dechant and S.-i. Sasa, "Improving thermodynamic bounds using correlations," *Phys. Rev. X* **11** (2021) 041061.

[135] D. A. Darling and M. Kac, "On occupation times for Markoff processes," *Trans. Am. Math. Soc.* **84** (1957) 444.

[136] E. Aghion, D. A. Kessler, and E. Barkai, "From non-normalizable Boltzmann-Gibbs statistics to infinite-ergodic theory," *Phys. Rev. Lett.* **122** (2019) 010601.

[137] S. N. Majumdar and A. Comtet, "Local and occupation time of a particle diffusing in a random medium," *Phys. Rev. Lett.* **89** (2002) 060601.

[138] A. J. Bray, S. N. Majumdar, and G. Schehr, "Persistence and first-passage properties in nonequilibrium systems," *Adv. Phys.* **62** (2013) 225.

[139] G. Bel and E. Barkai, "Weak ergodicity breaking in the continuous-time random walk," *Phys. Rev. Lett.* **94** (2005) 240602.

[140] C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, and F. C. MacKintosh, "Broken detailed balance at mesoscopic scales in active biological systems," *Science* **352** (2016) 604.

[141] S. Kusuoka, K. Kuwada, and Y. Tamura, "Large deviation for stochastic line integrals as L^p -currents," *Probab. Theory Relat. Fields* **147** (2009) 649.

[142] R. Chetrite and H. Touchette, "Nonequilibrium microcanonical and canonical ensembles and their equivalence," *Phys. Rev. Lett.* **111** (2013) 120601.

[143] E. Mallmin, J. du Buisson, and H. Touchette, "Large deviations of currents in diffusions with reflective boundaries," *J. Phys. A: Math. Theor.* **54** (2021) 295001.

[144] C. Monthus, "Inference of Markov models from trajectories via large deviations at level 2.5 with applications to random walks in disordered media," *J. Stat. Mech.* **2021** (2021) 063211.

[145] É. Roldán and J. M. R. Parrondo, "Estimating dissipation from single stationary trajectories," *Phys. Rev. Lett.* **105** (2010) 150607.

[146] É. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, "How far from equilibrium is active matter?," *Phys. Rev. Lett.* **117** (2016) 038103.

[147] J. Gladrow, M. Ribezzi-Crivellari, F. Ritort, and U. F. Keyser, "Experimental evidence of symmetry breaking of transition-path times," *Nat. Commun.* **10** (2019) 55.

[148] A. M. Berezhkovskii and D. E. Makarov, "From nonequilibrium single-molecule trajectories to underlying dynamics," *J. Phys. Chem. Lett.* **11** (2020) 1682.

[149] D. S. Seara, B. B. Machta, and M. P. Murrell, "Irreversibility in dynamical phases and transitions," *Nat. Commun.* **12** (2021) 392.

[150] L. F. Cugliandolo, D. S. Dean, and J. Kurchan, "Fluctuation-dissipation theorems and entropy production in relaxational systems," *Phys. Rev. Lett.* **79** (1997) 2168.

[151] D. Mizuno, C. Tardin, C. F. Schmidt, and F. C. MacKintosh, "Nonequilibrium mechanics of active cytoskeletal networks," *Science* **315** (2007) 370.

[152] U. Seifert, "Stochastic thermodynamics: From principles to the cost of precision," *Physica (Amsterdam)* **504A** (2018) 176.

[153] S. Otsubo, S. Ito, A. Dechant, and T. Sagawa, "Estimating entropy production by machine learning of short-time fluctuating currents," *Phys. Rev. E* **101** (2020) 062106.

[154] S. K. Manikandan, D. Gupta, and S. Krishnamurthy, "Inferring entropy production from short experiments," *Phys. Rev. Lett.* **124** (2020) 120603.

[155] H. Qian, S. Saffarian, and E. L. Elson, "Concentration fluctuations in a mesoscopic oscillating chemical reaction system," *Proc. Natl. Acad. Sci. U.S.A.* **99** (2002) 10376.

[156] S. Pilgram, A. N. Jordan, E. V. Sukhorukov, and M. Büttiker, "Stochastic path integral formulation of full counting statistics," *Phys. Rev. Lett.* **90** (2003) 206801.

[157] T. Bodineau and B. Derrida, "Current fluctuations in nonequilibrium diffusive systems: An additivity principle," *Phys. Rev. Lett.* **92** (2004) 180601.

[158] D. Andrieux and P. Gaspard, "Fluctuation theorem for currents and Schnakenberg network theory," *J. Stat. Phys.* **127** (2007) 107.

[159] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, "Current fluctuations in stochastic lattice gases," *Phys. Rev. Lett.* **94** (2005) 030601.

[160] R. K. P. Zia and B. Schmittmann, "Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states," *J. Stat. Mech.* (2007) 07012.

[161] M. Baiesi, C. Maes, and K. Netočný, "Computation of current cumulants for small nonequilibrium systems," *J. Stat. Phys.* **135** (2009) 57.

[162] P. Pietzonka, A. C. Barato, and U. Seifert, "Universal bounds on current fluctuations," *Phys. Rev. E* **93** (2016) 052145.

[163] T. R. Gingrich and J. M. Horowitz, "Fundamental bounds on first passage time fluctuations for currents," *Phys. Rev. Lett.* **119** (2017) 170601.

[164] S. C. Kapfer and W. Krauth, "Irreversible local Markov chains with rapid convergence towards equilibrium," *Phys. Rev. Lett.* **119** (2017) 240603.

[165] K. Macieszczak, K. Brandner, and J. P. Garrahan, "Unified thermodynamic uncertainty relations in linear response," *Phys. Rev. Lett.* **121** (2018) 130601.

[166] M. Kaiser, R. L. Jack, and J. Zimmer, "Canonical structure and orthogonality of forces and currents in irreversible Markov chains," *J. Stat. Phys.* **170** (2018) 1019.

[167] A. C. Barato, R. Chetrite, A. Faggionato, and D. Gabrielli, "Bounds on current fluctuations in periodically driven systems," *New J. Phys.* **20** (2018) 103023.

[168] S. Marcantoni, C. Pérez-Espigares, and J. P. Garrahan, "Symmetry-induced fluctuation relations for dynamical observables irrespective of their behavior under time reversal," *Phys. Rev. E* **101** (2020) 062142.

[169] D. Hartich and A. Godec, "Emergent memory and kinetic hysteresis in strongly driven networks," *Phys. Rev. X* **11** (2021) 041047.

[170] E. Suárez, R. P. Wiewiora, C. Wehmeyer, F. Noé, J. D. Chodera, and D. M. Zuckerman, "What Markov state models can and cannot do: Correlation versus path-based observables in protein-folding models," *J. Chem. Theory Comput.* **17** (2021) 3119.

[171] F. Flandoli, M. Gubinelli, M. Giaquinta, and V. M. Tortorelli, "Stochastic currents," *Stochastic Process. Appl.* **115** (2005) 1583.

[172] V. Y. Chernyak, M. Chertkov, S. V. Malinin, and R. Teodorescu, "Non-equilibrium thermodynamics and topology of currents," *J. Stat. Phys.* **137** (2009) 109.

[173] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, "Macroscopic fluctuation theory," *Rev. Mod. Phys.* **87** (2015) 593.

[174] A. Lapolla and A. Godec, "Unfolding tagged particle histories in single-file diffusion: exact single- and two-tag local times beyond large deviation theory," *New J. Phys.* **20** (2018) 113021.

[175] S.-i. Sasa, "Possible extended forms of thermodynamic entropy," *J. Stat. Mech.* **2014** (2014) P01004.

[176] T. Hatano and S.-i. Sasa, "Steady-state thermodynamics of Langevin systems," *Phys. Rev. Lett.* **86** (2001) 3463.

[177] M. S. Green, "Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids," *J. Chem. Phys.* **22** (1954) 398.

[178] R. Durrett, *Stochastic Calculus: A Practical Introduction*. CRC Press, Boca Raton, 1st ed., 1996.

[179] A. Frishman and P. Ronceray, "Learning force fields from stochastic trajectories," *Phys. Rev. X* **10** (2020) 021009.

[180] D. B. Brückner, P. Ronceray, and C. P. Broedersz, "Inferring the dynamics of underdamped stochastic systems," *Phys. Rev. Lett.* **125** (2020) 058103.

[181] F. Ferretti, V. Chardès, T. Mora, A. M. Walczak, and I. Giardina, "Building general Langevin models from discrete datasets," *Phys. Rev. X* **10** (2020) 031018.

[182] D. Hartich and U. Seifert, "Optimal inference strategies and their implications for the linear noise approximation," *Phys. Rev. E* **94** (2016) 042416.

[183] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer Berlin Heidelberg, 2010.

[184] V. Holubec, K. Kroy, and S. Steffenoni, "Physically consistent numerical solver for time-dependent Fokker-Planck equations," *Phys. Rev. E* **99** (2019) 032117.

[185] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, v. Roučka, A. Saboo, I. Fernando, S. Kulal, R. Cimrman, and A. Scopatz, "SymPy: symbolic computing in Python," *PeerJ Comput. Sci.* **3** (2017) e103.

[186] L. Dabelow, S. Bo, and R. Eichhorn, "Irreversibility in active matter systems: Fluctuation theorem and mutual information," *Phys. Rev. X* **9** (2019) 021009.

[187] M. Esposito and C. Van den Broeck, "Three faces of the second law. I. Master equation formulation," *Phys. Rev. E* **82** (2010) 011143.

[188] C. Van den Broeck and M. Esposito, "Three faces of the second law. II. Fokker-Planck formulation," *Phys. Rev. E* **82** (2010) 011144.

[189] C. Schroeder, R. Teixeira, E. Shaqfeh, and S. Chu, "Characteristic periodic motion of polymers in shear flow," *Phys. Rev. Lett.* **95** (2005) 018301.

[190] M. Harasim, B. Wunderlich, O. Peleg, M. Kröger, and A. R. Bausch, "Direct observation of the dynamics of semiflexible polymers in shear flow," *Phys. Rev. Lett.* **110** (2013) 108302.

[191] S. Gerashchenko and V. Steinberg, "Statistics of tumbling of a single polymer molecule in shear flow," *Phys. Rev. Lett.* **96** (2006) 038304.

[192] A. Alexander-Katz, M. F. Schneider, S. W. Schneider, A. Wixforth, and R. R. Netz, "Shear-flow-induced unfolding of polymeric globules," *Phys. Rev. Lett.* **97** (2006) 138101.

[193] H. Qian and M. Qian, "Pumped biochemical reactions, nonequilibrium circulation, and stochastic resonance," *Phys. Rev. Lett.* **84** (2000) 2271.

[194] H. Qian, "Phosphorylation energy hypothesis: Open chemical systems and their biological functions," *Annu. Rev. Phys. Chem.* **58** (2007) 113.

[195] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, "Hydrodynamics of soft active matter," *Rev. Mod. Phys.* **85** (2013) 1143.

[196] N. Fakhri, A. D. Wessel, C. Willms, M. Pasquali, D. R. Klopfenstein, F. C. MacKintosh, and C. F. Schmidt, "High-resolution mapping of intracellular fluctuations using carbon nanotubes," *Science* **344** (2014) 1031.

[197] P. Hänggi and H. Thomas, "Stochastic processes: Time evolution, symmetries and linear response," *Phys. Rep.* **88** (1982) 207.

[198] Y. Klimontovich, "Itô, Stratonovich and kinetic forms of stochastic equations," *Physica (Amsterdam)* **163A** (1990) 515.

[199] D. Carbone and L. Rondoni, "Necessary and sufficient conditions for time reversal symmetry in presence of magnetic fields," *Symmetry* **12** (2020) 1336.

[200] J. L. Doob, "Conditional Brownian motion and the boundary limits of harmonic functions," *Bull. Soc. Math. Fra.* **85** (1957) 431.

[201] T. Van Vu, V. T. Vo, and Y. Hasegawa, "Entropy production estimation with optimal current," *Phys. Rev. E* **101** (2020) 042138.

[202] J. Janczura, M. Balcerek, K. Burnecki, A. Sabri, M. Weiss, and D. Krapf, "Identifying heterogeneous diffusion states in the cytoplasm by a hidden Markov model," *New J. Phys.* **23** (2021) 053018.

[203] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, "Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking," *Phys. Chem. Chem. Phys.* **16** (2014) 24128.

[204] F. Coghi, R. Chetrite, and H. Touchette, "Role of current fluctuations in nonreversible samplers," *Phys. Rev. E* **103** (2021) 062142.

[205] D. S. Grebenkov, "Probability distribution of the time-averaged mean-square displacement of a Gaussian process," *Phys. Rev. E* **84** (2011) 031124.

[206] D. S. Grebenkov, "Time-averaged quadratic functionals of a Gaussian process," *Phys. Rev. E* **83** (2011) 061117.

[207] D. Boyer, D. S. Dean, C. Mejía-Monasterio, and G. Oshanin, "Optimal fits of diffusion constants from single-time data points of Brownian trajectories," *Phys. Rev. E* **86** (2012) 060101.

[208] D. Boyer, D. S. Dean, C. Mejía-Monasterio, and G. Oshanin, "Optimal estimates of the diffusion coefficient of a single Brownian trajectory," *Phys. Rev. E* **85** (2012) 031136.

[209] D. Boyer, D. S. Dean, C. Mejía-Monasterio, and G. Oshanin, "Distribution of the least-squares estimators of a single Brownian trajectory diffusion coefficient," *J. Stat. Mech.: Theory Exp.* **2013** (2013) P04017.

[210] P. Pietzonka, F. Ritort, and U. Seifert, "Finite-time generalization of the thermodynamic uncertainty relation," *Phys. Rev. E* **96** (2017) 012101.

[211] T. Koyuk and U. Seifert, "Operationally accessible bounds on fluctuations and entropy production in periodically driven systems," *Phys. Rev. Lett.* **122** (2019) 230601.

[212] T. Koyuk, U. Seifert, and P. Pietzonka, "A generalization of the thermodynamic uncertainty relation to periodically driven systems," *J. Phys. A: Math. Theor.* **52** (2018) 02LT02.

[213] A. C. Barato and U. Seifert, "Cost and precision of Brownian clocks," *Phys. Rev. X* **6** (2016) 041053.

[214] D. Hartich and A. Godec, "Thermodynamic uncertainty relation bounds the extent of anomalous diffusion," *Phys. Rev. Lett.* **127** (2021) 080601.

[215] B. Derrida and J. L. Lebowitz, "Exact large deviation function in the asymmetric exclusion process," *Phys. Rev. Lett.* **80** (1998) 209.

[216] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, "First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories," *J. Phys. A: Math. Theor.* **42** (2009) 075007.

[217] T. Bodineau and C. Toninelli, "Activity phase transition for constrained dynamics," *Commun. Math. Phys.* **311** (2012) 357.

[218] C. Maes, "Frenesy: Time-symmetric dynamical activity in nonequilibria," *Phys. Rep.* **850** (2020) 1.

[219] R. L. Jack, I. R. Thompson, and P. Sollich, "Hyperuniformity and phase separation in biased ensembles of trajectories for diffusive systems," *Phys. Rev. Lett.* **114** (2015) 060601.

[220] M. Vanicat, E. Bertin, V. Lecomte, and E. Ragoucy, "Mapping current and activity fluctuations in exclusion processes: consequences and open questions," *SciPost Phys.* **10** (2021) 028.

[221] G. Dell'Antonio, "Lecture 6: Lie-Trotter formula, Wiener process, Feynman-Kac formula," *Lectures on the Mathematics of Quantum Mechanics II: Selected Topics* (2016) 133.

[222] G. C. M. A. Ehrhardt, S. N. Majumdar, and A. J. Bray, "Persistence exponents and the statistics of crossings and occupation times for Gaussian stationary processes," *Phys. Rev. E* **69** (2004) 016106.

[223] S. Sabhapandit, S. N. Majumdar, and A. Comtet, "Statistical properties of functionals of the paths of a particle diffusing in a one-dimensional random potential," *Phys. Rev. E* **73** (2006) 051102.

[224] R. H. Cameron and W. T. Martin, "Transformations of weiner integrals under translations," *Ann. Math.* **45** (1944) 386.

[225] I. V. Girsanov, "On transforming a certain class of stochastic processes by absolutely continuous substitution of measures," *Theory Probab. Appl.* **5** (1960) 285.

[226] R. F. Fox, "Functional-calculus approach to stochastic differential equations," *Phys. Rev. A* **33** (1986) 467.

[227] R. F. Fox, "Stochastic calculus in physics," *J. Stat. Phys.* **46** (1987) 1145.

[228] R. L. Stratonovich, "A new representation for stochastic integrals and equations," *SIAM J. Control* **4** (1966) 362.

[229] N. Tizón-Escamilla, V. Lecomte, and E. Bertin, "Effective driven dynamics for one-dimensional conditioned langevin processes in the weak-noise limit," *J. Stat. Mech.* **2019** (2019) 013201.

[230] G. Falasco, M. Esposito, and J.-C. Delvenne, "Unifying thermodynamic uncertainty relations," *New J. Phys.* **22** (2020) 053046.

[231] C. Radhakrishna Rao, "Information and the accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.* **37** (1945) 81.

[232] H. Cramér, *Mathematical Methods of Statistics*. Princeton Univ. Press, Princeton, NJ, 1946.

[233] N. Shiraishi, "Optimal thermodynamic uncertainty relation in Markov jump processes," *J. Stat. Phys.* **185** (2021) 19.

[234] P. Pietzonka, "Classical pendulum clocks break the thermodynamic uncertainty relation," *Phys. Rev. Lett.* **128** (2022) 130606.

[235] K. Proesmans and C. Van den Broeck, "Discrete-time thermodynamic uncertainty relation," *Europhys. Lett.* **119** (2017) 20001.

[236] A. C. Barato, R. Chetrite, A. Faggionato, and D. Gabrielli, "A unifying picture of generalized thermodynamic uncertainty relations," *J. Stat. Mech.* **2019** (2019) 084017.

[237] T. Speck, "Stochastic thermodynamics for active matter," *Europhys. Lett.* **114** (2016) 30006.

[238] M. J. Bowick, N. Fakhri, M. C. Marchetti, and S. Ramaswamy, "Symmetry, thermodynamics, and topology in active matter," *Phys. Rev. X* **12** (2022) 010501.

[239] F. Jülicher, S. W. Grill, and G. Salbreux, "Hydrodynamic theory of active matter," *Rep. Prog. Phys.* **81** (2018) 076601.

[240] E. Fodor, R. L. Jack, and M. E. Cates, "Irreversibility and biased ensembles in active matter: Insights from stochastic thermodynamics," *Annu. Rev. Condens. Mat. Phys.* **13** (2022) 215.

[241] T. Koyuk and U. Seifert, "Quality of the thermodynamic uncertainty relation for fast and slow driving," *J. Phys. A: Math. Theor.* **54** (2021) 414005.

[242] R.-S. Fu and T. R. Gingrich, "Thermodynamic uncertainty relation for Langevin dynamics by scaling time," *Phys. Rev. E* **106** (2022) 024128.

[243] J. M. Horowitz and T. R. Gingrich, "Proof of the finite-time thermodynamic uncertainty relation for steady-state currents," *Phys. Rev. E* **96** (2017) 020103(R).

[244] J. C. Crocker and D. G. Grier, "Microscopic measurement of the pair interaction potential of charge-stabilized colloid," *Phys. Rev. Lett.* **73** (1994) 352.

[245] J. E. Curtis, B. A. Koss, and D. G. Grier, "Dynamic holographic optical tweezers," *Opt. Commun.* **207** (2002) 169.

[246] R. Dasgupta, R. S. Verma, and P. K. Gupta, "Microfluidic sorting with blinking optical traps," *Opt. Lett.* **37** (2012) 1739.

[247] L. Onsager, "The motion of ions: Principles and concepts, Nobel lecture," 1968.

[248] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Stat.* **22** (1951) 79.

[249] R. Jordan, D. Kinderlehrer, and F. Otto, "The variational formulation of the Fokker-Planck equation," *SIAM J. Math. Anal.* **29** (1998) 1.

[250] R. Kubo, M. Yokota, and S. Nakajima, "Statistical-mechanical theory of irreversible processes. II. response to thermal disturbance," *J. Phys. Soc. Jpn.* **12** (1957) 1203.

[251] Z. Lu and O. Raz, "Nonequilibrium thermodynamics of the Markovian Mpemba effect and its inverse," *Proc. Natl. Acad. Sci. U.S.A.* **114** (2017) 5083.

[252] A. Lasanta, F. Vega Reyes, A. Prados, and A. Santos, "When the hotter cools more quickly: Mpemba effect in granular fluids," *Phys. Rev. Lett.* **119** (2017) 148001.

[253] M. Baity-Jesi, E. Calore, A. Cruz, L. A. Fernandez, J. M. Gil-Narvion, A. Gordillo-Guerrero, D. Iñiguez, A. Lasanta, A. Maiorano, E. Marinari, V. Martin-Mayor, J. Moreno-Gordo, A. Muñoz Sudupe, D. Navarro, G. Parisi, S. Perez-Gaviro, F. Ricci-Tersenghi, J. J. Ruiz-Lorenzo, S. Fabio Schifano, B. Seoane, A. Tarancon, R. Tripiccione, and D. Yllanes, "The Mpemba effect in spin glasses is a persistent memory effect," *Proc. Natl. Acad. Sci. U.S.A.* **116** (2019) 15350.

[254] A. Kumar and J. Bechhoefer, "Exponentially faster cooling in a colloidal system," *Nature* **584** (2020) 64.

[255] F. Carollo, A. Lasanta, and I. Lesanovsky, "Exponentially accelerated approach to stationarity in Markovian open quantum systems through the Mpemba effect," *Phys. Rev. Lett.* **127** (2021) 060401.

[256] A. Kumar, R. Chétrite, and J. Bechhoefer, "Anomalous heating in a colloidal system," *Proc. Natl. Acad. Sci. U.S.A.* **119** (2022) e2118484119.

[257] I. Klich, O. Raz, O. Hirschberg, and M. Vucelja, "Mpemba index and anomalous relaxation," *Phys. Rev. X* **8** (2019) 021060.

[258] C. Josserand, A. V. Tkachenko, D. M. Mueth, and H. M. Jaeger, "Memory effects in granular materials," *Phys. Rev. Lett.* **85** (2000) 3632.

[259] Y. Lahini, O. Gottesman, A. Amir, and S. M. Rubinstein, "Nonmonotonic aging and memory retention in disordered mechanical systems," *Phys. Rev. Lett.* **118** (2017) 085501.

[260] I. L. Morgan, R. Avinery, G. Rahamim, R. Beck, and O. A. Saleh, "Glassy dynamics and memory effects in an intrinsically disordered protein construct," *Phys. Rev. Lett.* **125** (2020) 058001.

[261] A. Militaru, A. Lasanta, M. Frimmer, L. L. Bonilla, L. Novotny, and R. A. Rica, "Kovacs memory effect with an optically levitated nanoparticle," *Phys. Rev. Lett.* **127** (2021) 130603.

[262] B. Riechers, L. A. Roed, S. Mehri, T. S. Ingebrigtsen, T. Hecksher, J. C. Dyre, and K. Niss, "Predicting nonlinear physical aging of glasses from equilibrium relaxation via the material time," *Sci. Adv.* **8** (2022) eabl9809.

[263] V. Blickle and C. Bechinger, "Realization of a micrometre-sized stochastic heat engine," *Nat. Phys.* **8** (2011) 143.

[264] I. A. Martínez, E. Roldán, L. Dinis, D. Petrov, and R. A. Rica, "Adiabatic processes realized with a trapped Brownian particle," *Phys. Rev. Lett.* **114** (2015) 120601.

[265] J. L. Lebowitz and P. G. Bergmann, "Irreversible Gibbsian ensembles," *Ann. Phys.* **1** (1957) 1.

[266] S. Ito and A. Dechant, "Stochastic time evolution, information geometry, and the Cramér-Rao bound," *Phys. Rev. X* **10** (2020) 021056.

[267] M. Okuyama and M. Ohzeki, "Quantum speed limit is not quantum," *Phys. Rev. Lett.* **120** (2018) 070402.

[268] N. Shiraishi, K. Funo, and K. Saito, "Speed limit for classical stochastic processes," *Phys. Rev. Lett.* **121** (2018) 070601.

[269] G. E. Crooks, "Measuring thermodynamic length," *Phys. Rev. Lett.* **99** (2007) 100602.

[270] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, Hoboken, New Jersey, 2nd ed., 2006.

[271] A. Patrón, A. Prados, and C. A. Plata, "Thermal brachistochrone for harmonically confined Brownian particles," *Eur. Phys. J. Plus* **137** (2022) 1011.

[272] A. Zygmund, *Trigonometric series*, vol. 1. Cambridge university press, 2002.

[273] L. Brandolini and L. Colzani, "Localization and convergence of eigenfunction expansions," *J. Fourier Anal. Appl.* **5** (1999) 431.

[274] P. Resibois and M. F. D. Leener, *Classical Kinetic Theory of Fluids*. Wiley, Hoboken, New Jersey, 1977.

[275] I. A. Martínez, É. Roldán, L. Dinis, D. Petrov, J. M. R. Parrondo, and R. A. Rica, "Brownian Carnot engine," *Nat. Phys.* **12** (2015) 67.

[276] S. Krishnamurthy, S. Ghosh, D. Chatterji, R. Ganapathy, and A. K. Sood, "A micrometre-sized heat engine operating between bacterial reservoirs," *Nat. Phys.* **12** (2016) 1134.

[277] M. Rademacher, M. Konopik, M. Debiossac, D. Grass, E. Lutz, and N. Kiesel, "Nonequilibrium control of thermal and mechanical changes in a levitated system," *Phys. Rev. Lett.* **128** (2022) 070601.

[278] I. A. Martínez, A. Petrosyan, D. Guéry-Odelin, E. Trizac, and S. Ciliberto, "Engineered swift equilibration of a Brownian particle," *Nat. Phys.* **12** (2016) 843.

[279] D. Guéry-Odelin, C. Jarzynski, C. A. Plata, A. Prados, and E. Trizac, "Driving rapidly while remaining in control: classical shortcuts from Hamiltonian to stochastic dynamics," *Rep. Prog. Phys.* **86** (2023) 035902.

[280] M. Polettini and M. Esposito, "Nonconvexity of the relative entropy for Markov dynamics: A Fisher information approach," *Phys. Rev. E* **88** (2013) 012112.

[281] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Adv. Comput. Math.* **5** (1996) 329.

[282] S. M. Stewart, "On certain inequalities involving the Lambert W function," *J. Inequal. Pure Appl. Math.* **10** (2009) 96.

[283] I. A. Martínez, É. Roldán, J. M. R. Parrondo, and D. Petrov, "Effective heating to several thousand kelvins of an optically trapped sphere in a liquid," *Phys. Rev. E* **87** (2013) 032159.

[284] M. Baiesi and C. Maes, "An update on the nonequilibrium linear response," *New J. Phys.* **15** (2013) 013004.

[285] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, "Minimum dissipation principle in stationary non-equilibrium states," *J. Stat. Phys.* **116** (2004) 831.

[286] R. S. Maier and D. L. Stein, "Escape problem for irreversible systems," *Phys. Rev. E* **48** (1993) 931.

[287] F. Bouchet and J. Reygner, "Generalisation of the Eyring-Kramers transition rate formula to irreversible diffusion processes," *Ann. Henri Poincare* **17** (2016) 3499.

[288] J. Casademunt, R. Mannella, P. V. E. McClintock, F. E. Moss, and J. M. Sancho, "Relaxation times of non-Markovian processes," *Phys. Rev. A* **35** (1987) 5183.

[289] I. M. Sokolov, "Solutions of a class of non-Markovian Fokker-Planck equations," *Phys. Rev. E* **66** (2002) 041101.

[290] I. M. Sokolov, "Cyclization of a polymer: First-passage problem for a Non-Markovian process," *Phys. Rev. Lett.* **90** (2003) 080601.

[291] B. Dybiec, I. M. Sokolov, and A. V. Chechkin, "Relaxation to stationary states for anomalous diffusion," *Commun. Nonlinear Sci. Numer. Simul.* **16** (2011) 4549.

[292] D. E. Makarov, "Interplay of non-Markov and internal friction effects in the barrier crossing kinetics of biopolymers: Insights from an analytically solvable model," *J. Chem. Phys.* **138** (2013) 014102.

[293] R. Satija and D. E. Makarov, "Generalized langevin equation as a model for barrier crossing dynamics in biomolecular folding," *J. Phys. Chem. B* **123** (2019) 802.

[294] A. Lapolla and A. Godec, "Manifestations of projection-induced memory: General theory and the tilted single file," *Front. Phys.* **7** (2019) 182.

[295] A. Lapolla and A. Godec, "Single-file diffusion in a bi-stable potential: Signatures of memory in the barrier-crossing of a tagged-particle," *J. Chem. Phys.* **153** (2020) 194104.

[296] P. Talkner and P. Hänggi, "Colloquium: Statistical mechanics and thermodynamics at strong coupling: Quantum and classical," *Rev. Mod. Phys.* **92** (2020) 041002.

[297] C. Ayaz, L. Scalfi, B. A. Dalton, and R. R. Netz, "Generalized langevin equation with a nonlinear potential of mean force and nonlinear memory friction from a hybrid projection scheme," *Phys. Rev. E* **105** (2022) 054138.

[298] E. Lippiello, M. Baiesi, and A. Sarracino, "Nonequilibrium fluctuation-dissipation theorem and heat production," *Phys. Rev. Lett.* **112** (2014) 140602.

[299] A. Gal and O. Raz, "Precooling strategy allows exponentially faster heating," *Phys. Rev. Lett.* **124** (2020) 060602.

[300] R. Holtzman and O. Raz, "Landau theory for the Mpemba effect through phase transitions," *Commun. Phys.* **5** (2022) 280.

[301] J. Degünther and U. Seifert, "Anomalous relaxation from a non-equilibrium steady state: An isothermal analog of the Mpemba effect," *Europhys. Lett.* **139** (2022) 41002.

[302] F. Bouchet, K. Gawedzki, and C. Nardini, "Perturbative calculation of quasi-potential in non-equilibrium diffusions: A mean-field example," *J. Stat. Phys.* **163** (2016) 1157.

[303] Y. Baek and Y. Kafri, "Singularities in large deviation functions," *J. Stat. Mech.* **2015** (2015) P08026.

[304] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, "Dynamical first-order phase transition in kinetically constrained models of glasses," *Phys. Rev. Lett.* **98** (2007) 195702.

[305] T. Speck, A. Engel, and U. Seifert, "The large deviation function for entropy production: the optimal trajectory and the role of fluctuations," *J. Stat. Mech.* **2012** (2012) P12001.

[306] P. Tsobgni Nyawo and H. Touchette, "A minimal model of dynamical phase transition," *Europhys. Lett.* **116** (2016) 50009.

[307] V. Ermolaev and C. Külske, "Low-temperature dynamics of the curie-weiss model: Periodic orbits, multiple histories, and loss of Gibbsianness," *J. Stat. Phys.* **141** (2010) 727.

[308] J. Meibohm and M. Esposito, "Finite-time dynamical phase transition in nonequilibrium relaxation," *Phys. Rev. Lett.* **128** (2022) 110603.

[309] J. Meibohm and M. Esposito, "Landau theory for finite-time dynamical phase transitions," *New J. Phys.* **25** (2023) 023034.

[310] K. Blom and A. Godec, "Global speed limit for finite-time dynamical phase transition in nonequilibrium relaxation," (2022) arXiv:2209.14287.

[311] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, "Optimal protocols and optimal transport in stochastic thermodynamics," *Phys. Rev. Lett.* **106** (2011) 250601.

[312] K. Yoshimura and S. Ito, "Thermodynamic uncertainty relation and thermodynamic speed limit in deterministic chemical reaction networks," *Phys. Rev. Lett.* **127** (2021) 160601.

[313] A. Lapolla and A. Godec, "Erratum: Faster uphill relaxation in thermodynamically equidistant temperature quenches [Phys. Rev. Lett. **125**, 110602 (2020)]," *Phys. Rev. Lett.* **128** (2022) 229901.

[314] C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu, "Accelerating diffusions," *Ann. Appl. Probab.* **15** (2005) 1433.

[315] L. Rey-Bellet and K. Spiliopoulos, "Irreversible Langevin samplers and variance reduction: a large deviations approach," *Nonlinearity* **28** (2015) 2081.

[316] C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu, "Accelerating Gaussian diffusions," *Ann. Appl. Probab.* **3** (1993) 897.

[317] M. C. Mackey, "The dynamic origin of increasing entropy," *Rev. Mod. Phys.* **61** (1989) 981.

[318] M. Esposito and C. Van den Broeck, "Three detailed fluctuation theorems," *Phys. Rev. Lett.* **104** (2010) 090601.

[319] B. H. Zimm, "Dynamics of polymer molecules in dilute solution: Viscoelasticity, flow birefringence and dielectric loss," *J. Chem. Phys.* **24** (1956) 269.

[320] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*. International Series of Monographs on Physics. Clarendon Press, Oxford, England, Nov., 1988.

[321] R. R. Cheng, A. T. Hawk, and D. E. Makarov, "Exploring the role of internal friction in the dynamics of unfolded proteins using simple polymer models," *J. Chem. Phys.* **138** (2013) 074112.

[322] A. Soranno, B. Buchli, D. Nettels, R. R. Cheng, S. Müller-Späth, S. H. Pfeil, A. Hoffmann, E. A. Lipman, D. E. Makarov, and B. Schuler, "Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy," *Proc. Natl. Acad. Sci. U.S.A.* **109** (2012) 17800.

[323] B. Schuler, A. Soranno, H. Hofmann, and D. Nettels, "Single-molecule FRET spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins," *Annu. Rev. Biophys.* **45** (2016) 207. PMID: 27145874.

[324] J. O. Daldrop, J. Kappler, F. N. Brünig, and R. R. Netz, "Butane dihedral angle dynamics in water is dominated by internal friction," *Proc. Natl. Acad. Sci. U.S.A.* **115** (2018) 5169.

[325] D. E. Smith, H. P. Babcock, and S. Chu, "Single-polymer dynamics in steady shear flow," *Science* **283** (1999) 1724.

[326] P. S. Doyle, B. Ladoux, and J.-L. Viovy, "Dynamics of a tethered polymer in shear flow," *Phys. Rev. Lett.* **84** (2000) 4769.

[327] T. T. Perkins, D. E. Smith, R. G. Larson, and S. Chu, "Stretching of a single tethered polymer in a uniform flow," *Science* **268** (1995) 83.

[328] R. E. Teixeira, H. P. Babcock, E. S. G. Shaqfeh, and S. Chu, "Shear thinning and tumbling dynamics of single polymers in the flow-gradient plane," *Macromolecules* **38** (2005) 581.

[329] T. T. Perkins, D. E. Smith, and S. Chu, "Single polymer dynamics in an elongational flow," *Science* **276** (1997) 2016.

[330] Y. Roichman, B. Sun, A. Stolarski, and D. G. Grier, "Influence of nonconservative optical forces on the dynamics of optically trapped colloidal spheres: The fountain of probability," *Phys. Rev. Lett.* **101** (2008) 128301.

[331] G. Metafune, "l^p-spectrum of Ornstein-Uhlenbeck operators," *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **30** (2001) 97.

[332] G. Metafune, D. Pallara, and E. Priola, "Spectrum of Ornstein-Uhlenbeck operators in l^p spaces with respect to invariant measures," *J. Funct. Anal.* **196** (2002) 40–60.

[333] H. F. Trotter, "On the product of semi-groups of operators," *Proc. Amer. Math. Soc.* **10** (1959) 545.

[334] G. Dahlquist, *Stability and Error Bounds in the Numerical Integration of Ordinary Differential Equations*. Almqvist & Wiksell, Uppsala, Sweden, 1958.

[335] E. Barkai, Y. Garini, and R. Metzler, "Strange kinetics of single molecules in living cells," *Phys. Today* **65** (2012) 29.

[336] M. Weiss, "Crowding, diffusion, and biochemical reactions," *Int. Rev. Cell Mol. Biol* (2014) 383.

[337] B. B. Mandelbrot and J. W. V. Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.* **10** (1968) 422.

[338] F. Mezei, C. Pappas, and T. Gutberlet, eds., *Neutron Spin Echo Spectroscopy: Basics, Trends, and Applications*, vol. 601 of *Lecture Notes in Physics*. Springer, 2002.

[339] D. Richter, M. Monkenbusch, A. Arbe, and J. Colmenero, *Neutron Spin Echo in Polymer Systems*. Springer Berlin Heidelberg, 2005.

[340] D. J. Callaway and Z. Bu, "Visualizing the nanoscale: protein internal dynamics and neutron spin echo spectroscopy," *Curr. Opin. Struct. Biol.* **42** (2017) 1.

[341] Y. Liu, "Intermediate scattering function for macromolecules in solutions probed by neutron spin echo," *Phys. Rev. E* **95** (2017) 020501.

[342] C. Kurzthaler, C. Devailly, J. Arlt, T. Franosch, W. C. K. Poon, V. A. Martinez, and A. T. Brown, "Probing the spatiotemporal dynamics of catalytic Janus particles with single-particle tracking and differential dynamic microscopy," *Phys. Rev. Lett.* **121** (2018) 078001.

[343] R. Cerbino and V. Trappe, "Differential dynamic microscopy: Probing wave vector dependent dynamics with a microscope," *Phys. Rev. Lett.* **100** (2008) 188102.

[344] L. G. Wilson, V. A. Martinez, J. Schwarz-Linek, J. Tailleur, G. Bryant, P. N. Pusey, and W. C. K. Poon, "Differential dynamic microscopy of bacterial motility," *Phys. Rev. Lett.* **106** (2011) 018101.

[345] E. N. Senning and A. H. Marcus, "Subcellular dynamics and protein conformation fluctuations measured by Fourier imaging correlation spectroscopy," *Annu. Rev. Phys. Chem.* **61** (2010) 111.

[346] M. Knowles, T. Grassman, and A. Marcus, "Measurement of the dynamic structure function of fluorescently labeled complex fluids by Fourier imaging correlation spectroscopy," *Phys. Rev. Lett.* **85** (2000) 2837.

[347] D. L. Kolin, D. Ronis, and P. W. Wiseman, "k-space image correlation spectroscopy: A method for accurate transport measurements independent of fluorophore photophysics," *Biophys. J.* **91** (2006) 3061.

[348] L. Van Hove, "Correlations in space and time and Born approximation scattering in systems of interacting particles," *Phys. Rev.* **95** (1954) 249.

[349] J.-P. Bouchaud and A. Georges, "Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications," *Phys. Rep.* **195** (1990) 127.

[350] D. Krapf, "Mechanisms underlying anomalous diffusion in the plasma membrane," *Curr. Top. Membr.* (2015) 167.

[351] J.-H. Jeon, M. Javanainen, H. Martinez-Seara, R. Metzler, and I. Vattulainen, "Protein crowding in lipid bilayers gives rise to non-Gaussian anomalous lateral diffusion of phospholipids and proteins," *Phys. Rev. X* **6** (2016) 021006.

[352] G. Guigas and M. Weiss, "Sampling the cell with anomalous diffusion—the discovery of slowness," *Biophys. J.* **94** (2008) 90.

[353] S. A. McKinley, L. Yao, and M. G. Forest, "Transient anomalous diffusion of tracer particles in soft matter," *J. Rheol.* **53** (2009) 1487.

[354] I. M. Tolić-Nørrelykke, E.-L. Munteanu, G. Thon, L. Oddershede, and K. Berg-Sørensen, "Anomalous diffusion in living yeast cells," *Phys. Rev. Lett.* **93** (2004) 078102.

[355] I. Y. Wong, M. L. Gardel, D. R. Reichman, E. R. Weeks, M. T. Valentine, A. R. Bausch, and D. A. Weitz, "Anomalous diffusion probes microstructure dynamics of entangled f-actin networks," *Phys. Rev. Lett.* **92** (2004) 178101.

[356] A. Godec, M. Bauer, and R. Metzler, "Collective dynamics effect transient subdiffusion of inert tracers in flexible gel networks," *New J. Phys.* **16** (2014) 092002.

[357] M. V. Chubynsky and G. W. Slater, "Diffusing diffusivity: A model for anomalous, yet Brownian, diffusion," *Phys. Rev. Lett.* **113** (2014) 098302.

[358] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, "Brownian yet non-Gaussian diffusion: From superstatistics to subordination of diffusing diffusivities," *Phys. Rev. X* **7** (2017) 021002.

[359] C. Manzo and M. F. Garcia-Parajo, "A review of progress in single particle tracking: from methods to biophysical insights," *Rep. Prog. Phys.* **78** (2015) 124601.

[360] Y. Meroz and I. M. Sokolov, "A toolbox for determining subdiffusive mechanisms," *Phys. Rep.* **573** (2015) 1.

[361] A. Szabo, D. Shoup, S. H. Northrup, and J. A. McCammon, "Stochastically gated diffusion-influenced reactions," *J. Chem. Phys.* **77** (1982) 4484.

[362] R. Zwanzig, "Dynamical disorder: Passage through a fluctuating bottleneck," *J. Chem. Phys.* **97** (1992) 3587.

[363] M. Bier and R. D. Astumian, "Matching a diffusive and a kinetic approach for escape over a fluctuating barrier," *Phys. Rev. Lett.* **71** (1993) 1649.

[364] C. R. Doering and J. C. Gadoua, "Resonant activation over a fluctuating barrier," *Phys. Rev. Lett.* **69** (1992) 2318.

[365] A. Godec and R. Metzler, "First passage time statistics for two-channel diffusion," *J. Phys. A: Math. Theor.* **50** (2017) 084001.

[366] Z. Wunderlich and L. A. Mirny, "Spatial effects on the speed and reliability of protein-DNA search," *Nucl. Acids Res.* **36** (2008) 3570.

[367] E. F. Koslover, M. A. Díaz de la Rosa, and A. J. Spakowitz, "Theoretical and computational modeling of target-site search kinetics in vitro and in vivo," *Biophys. J.* **101** (2011) 856.

[368] M. Sheinman, O. Bénichou, Y. Kafri, and R. Voituriez, "Classes of fast and specific search mechanisms for proteins on dna," *Rep. Prog. Phys.* **75** (2012) 026601.

[369] E. Yamamoto, T. Akimoto, A. Mitsutake, and R. Metzler, "Universal relation between instantaneous diffusivity and radius of gyration of proteins in aqueous solution," *Phys. Rev. Lett.* **126** (2021) 128101.

[370] F. Baldovin, E. Orlandini, and F. Seno, "Polymerization induces non-Gaussian diffusion," *Front. Phys.* **7** (2019) 124.

[371] H. B. Maizón and F. J. Barrantes, "A deep learning-based approach to model anomalous diffusion of membrane proteins: the case of the nicotinic acetylcholine receptor," *Brief. Bioinformatics* **23** (2021) bbab435.

[372] J. Kärger, "NMR self-diffusion studies in heterogeneous systems," *Adv. Coll. Interf. Sci.* **23** (1985) 129.

[373] D. S. Grebenkov, "Time-averaged mean square displacement for switching diffusion," *Phys. Rev. E* **99** (2019) 032133.

[374] D. Ernst, J. Köhler, and M. Weiss, "Probing the type of anomalous diffusion with single-particle tracking," *Phys. Chem. Chem. Phys.* **16** (2014) 7686.

[375] M. Woringer, I. Izeddin, C. Favard, and H. Berry, "Anomalous subdiffusion in living cells: Bridging the gap between experiments and realistic models through collaborative challenges," *Front. Phys.* **8** (2020) 134.

[376] G. Sikora, A. Wyłomańska, J. Gajda, L. Solé, E. J. Akin, M. M. Tamkun, and D. Krapf, "Elucidating distinct ion channel populations on the surface of hippocampal neurons via single-particle tracking recurrence analysis," *Phys. Rev. E* **96** (2017) 062404.

[377] T. Akimoto and E. Yamamoto, "Detection of transition times from single-particle-tracking trajectories," *Phys. Rev. E* **96** (2017) 052138.

[378] Y. Lanoiselée and D. S. Grebenkov, "Unraveling intermittent features in single-particle trajectories by a local convex hull method," *Phys. Rev. E* **96** (2017) 022144.

[379] G. Zaccai, "How soft is a protein? A protein dynamics force constant measured by neutron scattering," *Science* **288** (2000) 1604.

[380] G. R. Kneller, "Quasielastic neutron scattering and relaxation processes in proteins: analytical and simulation-based models," *Phys. Chem. Chem. Phys.* **7** (2005) 2641.

[381] R. Biehl, M. Monkenbusch, and D. Richter, "Exploring internal protein dynamics by neutron spin echo spectroscopy," *Soft Matter* **7** (2011) 1299.

[382] P. Martel, "Biophysical aspects of neutron scattering from vibrational modes of proteins," *Prog. Biophys. molec. Biol.* **57** (1992) 129.

[383] M. Monkenbusch, M. Krutyeva, W. Pyckhout-Hintzen, W. Antonius, C. H. Hövelmann, J. Allgaier, A. Brás, B. Farago, A. Wischnewski, and D. Richter, "Molecular view on supramolecular chain and association dynamics," *Phys. Rev. Lett.* **117** (2016) 147802.

[384] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*. Academic Press, Oxford, fourth ed., 2013.

[385] V. Réat, H. Patzelt, M. Ferrand, C. Pfister, D. Oesterhelt, and G. Zaccai, "Dynamics of different functional parts of bacteriorhodopsin: H-2H labeling and neutron scattering," *Proc. Natl. Acad. Sci. U.S.A.* **95** (1998) 4970.

[386] K. S. Singwi and A. Sjölander, "Diffusive motions in water and cold neutron scattering," *Phys. Rev.* **119** (1960) 863.

[387] E. B. Postnikov, A. Chechkin, and I. M. Sokolov, "Brownian yet non-Gaussian diffusion in heterogeneous media: from superstatistics to homogenization," *New J. Phys.* **22** (2020) 063046.

[388] M. V. Smoluchowski, "Drei Vorträge uber Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen," *Phys. Z.* **17** (1916) 557.

[389] F. C. Collins and G. E. Kimball, "Diffusion-controlled reaction rates," *J. Colloid Sci.* **4** (1949) 425.

[390] H. Bateman and A. Erdélyi, *Higher transcendental functions, vol. 1*. California Institute of technology. Bateman Manuscript project. McGraw-Hill, New York, NY, 1955.

[391] O. Matsarskaia, L. Bühl, C. Beck, M. Grimaldo, R. Schweins, F. Zhang, T. Seydel, F. Schreiber, and F. Roosen-Runge, "Evolution of the structure and dynamics of bovine serum albumin induced by thermal denaturation," *Phys. Chem. Chem. Phys.* **22** (2020) 18507.

[392] L. Henry, M. R. Panman, L. Isaksson, E. Claesson, I. Kosheleva, R. Henning, S. Westenhoff, and O. Berntsson, "Real-time tracking of protein unfolding with time-resolved X-ray solution scattering," *Struct. Dyn.* **7** (2020) 054702.

[393] T. W. Kim, S. J. Lee, J. Jo, J. G. Kim, H. Ki, C. W. Kim, K. H. Cho, J. Choi, J. H. Lee, M. Wulff, Y. M. Rhee, and H. Ihee, "Protein folding from heterogeneous unfolded state revealed by time-resolved X-ray solution scattering," *Proc. Natl. Acad. Sci. U.S.A.* **117** (2020) 14996.

[394] A. Lapolla, M. Vossel, and A. Godec, "Time- and ensemble-average statistical mechanics of the Gaussian network model," *J. Phys. A: Math. Theor.* **54** (2021) 355601.

[395] T. Haliloglu, I. Bahar, and B. Erman, "Gaussian dynamics of folded proteins," *Phys. Rev. Lett.* **79** (1997) 3090.

[396] A. Bakan, L. M. Meireles, and I. Bahar, "ProDy: Protein Dynamics Inferred from Theory and Experiments," *Bioinformatics* **27** (2011) 1575.

[397] P. E. Rouse, "A theory of the linear viscoelastic properties of dilute solutions of coiling polymers," *J. Chem. Phys.* **21** (1953) 1272.

[398] R. Granek and M. E. Cates, "Stress relaxation in living polymers: Results from a Poisson renewal model," *J. Chem. Phys.* **96** (1992) 4758.

[399] E. B. Stukalin and K. F. Freed, "Minimal model of relaxation in an associating fluid: Viscoelastic and dielectric relaxations in equilibrium polymer solutions," *J. Chem. Phys.* **125** (2006) 184905.

[400] D. Krapf and R. Metzler, "Strange interfacial molecular dynamics," *Phys. Today* **72** (2019) 48.

[401] A. Weron, K. Burnecki, E. J. Akin, L. Solé, M. Balcerek, M. M. Tamkun, and D. Krapf, "Ergodicity breaking on the neuronal surface emerges from random switching between diffusive states," *Sci. Rep.* **7** (2017) 5404.

[402] J. Wuttke, "Laplace–Fourier transform of the stretched exponential function: Analytic error bounds, double exponential transform, and open-source implementation "libkww"," *Algorithms* **5** (2012) 604.

[403] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series - Direct Laplace Transforms*, vol. 4. Gordon and Breach Science Publishers, 1992.

[404] R. B. Davies and D. S. Harte, "Tests for Hurst effect," *Biometrika* **74** (1987) 95.

[405] J. L. Doob, *Stochastic Processes*. John Wiley & Sons, New York, 1953.