White-noise fluctuation theorem for Langevin dynamics

M Innerbichler, A Militaru, M Frimmer, I Novotny and C Dellago

1 Faculty of Physics, University of Vienna, 1090 Vienna, Austria
2 Photonics Laboratory, ETH Zürich, CH-8093 Zurich, Switzerland
3 Quantum Center, ETH Zürich, CH-8093 Zurich, Switzerland

E-mail: christoph.dellago@univie.ac.at

Keywords: fluctuation theorems, stochastic thermodynamics, stochastic processes, parameter inference

Abstract

Fluctuation theorems (FTs) based on time-reversal have provided remarkable insight into the non-equilibrium statistics of thermodynamic quantities like heat, work, and entropy production. These types of laws impose constraints on the distributions of certain trajectory functionals that reflect underlying dynamical symmetries. In this work, we introduce a detailed FT for Langevin dynamics that follows from the statistics of Gaussian white noise rather than from time-reversal. The theorem, which originates from a point-wise symmetry in phase space, holds individually for each degree of freedom coupled to additive or multiplicative noise. The relation is independent of the phase space distribution generated by the dynamics and can be used to derive a versatile parameter inference algorithm applicable to a wide range of systems, including non-conservative and non-Markovian ones.

1. Introduction

Stochastic thermodynamics revolves around the generalization of macroscopic thermodynamics to the micro- and mesoscopic regime. At this scale, quantities like heat, work, or entropy production become functionals of stochastically evolving trajectories and typical fluctuations exhibit magnitudes comparable to the respective mean and standard deviation [1]. Hence, it is necessary to consider probability distributions rather than operating with average values only. Remarkably, the fundamental symmetries of the underlying dynamics often result in so-called fluctuation theorems (FTs), i.e. physical laws constraining these distributions [2, 3]. Time-reversal in particular has led to several important relations like the well-known Crooks FT [4]: under the assumption of microscopic reversibility (detailed balance), this theorem connects the probability distributions of work dissipated during a non-equilibrium transformation carried out forward and backward in time. Other FTs constrain only expectation values: the Jarzynski equality [5] relates the exponentiated work average to the free energy difference between equilibrium states and the Hatano–Sasa relation [6] extends this result to transitions between non-equilibrium stationary states, to name but a few [2].

Although various notions of time-reversal have been exploited to deduce FTs [3], they are not the only dynamical symmetries that permit doing so. Here, we explore this possibility for the case of general Langevin equations, which are widely used to model continuous stochastic processes and most commonly feature Gaussian white noise as a source of randomness [7, 8]. The model in this particular form and its equivalent Fokker–Planck representation [9] have been applied to a wide range of physical, chemical, and biological phenomena. Examples include micro- and macroscopic systems under the influence of thermal noise [10], active particles [11–13], reaction coordinates in transition processes [14, 15], and the concentrations of reactants in a well-stirred solution [16].
The relations introduced in this work have the form of detailed fluctuation theorems (DFTs) and emerge from the properties of Gaussian white noise without any reference to time-reversal. Seifert has previously discussed such relations for pure Langevin systems with arbitrary deterministic forces and additive noise [2, 17]. Here, we extend upon these investigations in several ways, also examining systems that are only partially described by Langevin equations with additive or multiplicative noise. We do not require the existence of a stationary state or any knowledge about the phase space distribution at any time. These white-noise DFTs offer a considerable degree of flexibility and remain valid for various systems that have previously been studied in search for reversibility-based distributional symmetries and inequalities. Examples range from magnetic forces [18–20] to active matter [21, 22] and even delayed feedback [23, 24]. The treatment of each of these cases poses unique challenges in the framework of stochastic thermodynamics due to the specific nature of time-reversal. Conversely, the noise-symmetry approach followed here circumvents these additional complications per construction, demanding no special attention for any of the issues mentioned above.

Note, however, that the generality of the white-noise DFT does come at a cost: lacking the usual time-reversal—based connection to entropy production, the theorem does not provide any immediate thermodynamic insight. Moreover, it does not even require the existence of an underlying Hamiltonian to drive the system’s dynamics, and may consequently apply to non-physical systems as well. Rather than refining our understanding of thermodynamics in particular, we use our results to devise a versatile parameter inference method building on the white-noise DFT. The general idea of this approach is to determine unknown parameters by minimizing the deviation between the statistics of the empirical data and the symmetry dictated by the DFT. We illustrate the method numerically on the example of an oscillator with delayed feedback, extracting the delay time. The method is widely applicable and only requires accurate trajectory data.

The remainder of this paper is organized as follows. In section 2, we first derive a local version of the DFT for displacements conditioned to individual points in phase space. We then extend the local version of the DFT to trajectory data, deriving equation (10) which is the main result of this work. We then demonstrate the validity of the DFT for several numerical examples in sections 3 and 4. These systems feature arbitrary initial distributions, lack of stationary densities, and non-Markovianity. In section 5 we outline the parameter inference method and discuss its application. After reviewing the impact of signal contamination and measurement noise on the white-noise DFT in section 6, we provide our concluding remarks in section 7.

2. Noise-induced FT symmetries

Let us consider a general system described by $M$ degrees of freedom $\mathbf{q} = \{q_1, q_2, \ldots, q_M\}$. These coordinates $q_i$ are in principle not restricted to positions and velocities, but may include more general quantities as well, e.g. reaction coordinates. The evolution of $\mathbf{q}$ can generally be described by an independent system of $M$ coupled equations. We assume that the first $N \leq M$ degrees of freedom evolve according to Langevin equations, that is laws of the form

$$\dot{q}_i = f_i(\mathbf{q}, t) + \sum_{j=1}^{k} g_{ij}(\mathbf{q}, t) \eta_j(t).$$

(1)

Here, $\dot{q}_i$ designates the derivative with respect to time $t$ of coordinate $q_i$, the $i$th element of state vector $\mathbf{q}$. The index $j$, on the other hand, refers to the Gaussian white noise $\eta_j$. The $k$ distinct Gaussian white noises are independent, satisfying the relation $\langle \eta_j(t) \eta_j(t') \rangle = \delta_{jj'} \delta(t - t')$. Finally, the functions $f_i$ and $g_{ij}$ prescribe the deterministic drift and noise amplitudes respectively. We deal with multiplicative noise if any $g_{ij}$ explicitly depends on $\mathbf{q}$. Note that equation (1) by itself is ill-defined in the presence of multiplicative noise—one needs to additionally prescribe a suitable integration convention [25, 26]. It is possible to map Langevin equations from one prescription to another by modifying the deterministic term $f_i$, giving statistically equivalent descriptions of the system (see appendix A). We can therefore impose the Itô convention [25, 26] on all Langevin equations appearing henceforth without loss of generality. Doing so leads to a particularly simple theoretical treatment.

It should be underlined that we do not require that each and every component of $\mathbf{q}$ follows Langevin dynamics. The upcoming considerations hold for each $q_i$ with $i \leq N$ individually, irrespective of the remaining dynamical details defining the system. Furthermore, the deterministic terms $f_i$ and noise amplitudes $g_{ij}$ may depend on time $t$ explicitly. In general, Langevin systems may contain multiple Gaussian white noise terms $\eta_i$ that are possibly shared between multiple equations. Then $g_{ij}$ is non-zero for multiple $i$ given any particular $j$. For ease of notation and clarity we focus on cases with single, independent white...
noises first, rendering the matrix of noise amplitudes \( g_{ij} \) diagonal. We examine the influence of multiple noise sources in section 2.1. Finally, our upcoming discussions presume free boundary conditions w.r.t. coordinate \( q_i \), i.e. \( q_i \in \mathbb{R} \). Otherwise it is necessary to restrain our analysis to points in phase space with a finite distance to the boundary. We will further elaborate on the importance of this premise later.

Let us consider a temporal discretization up to first order of equation (1) for \( k = 1 \) and rearrange the terms. In the Itô convention, functions \( f_i \) and \( g_i \) are evaluated at the beginning of the time step, imposing the discretization scheme

\[
\Delta q_i = f_i(q, t) \Delta t + g_i(q, t) \Delta W_i(t). \tag{2}
\]

We assume that \( g_i(q, t) \neq 0 \) holds and the time step \( \Delta t \) is small. The symbol \( \Delta q_i \) is a shorthand for the displacement \( q_i(t + \Delta t) - q_i(t) \). On the right-hand side, \( \Delta W_i(t) \) denotes the increments of a Wiener process, distributed according to a Gaussian with variance \( \Delta t \) and zero mean. Explicitly, its probability density \( P(\Delta W_i) \) reads

\[
P(\Delta W_i) = \frac{1}{\sqrt{2\pi \Delta t}} \exp \left( -\frac{\Delta W_i^2}{2\Delta t} \right). \tag{3}
\]

The displacement \( \Delta q_i \) conditioned to an initial state \( (q, t) \) where functions \( f_i \) and \( g_i \) are finite similarly follows a Gaussian distribution. This property is exclusive to the Itô convention, which induces a linear relationship between \( \Delta q_i \) and \( \Delta W_i(t) \) as seen in equation (2). One can now relate the conditional probabilities of two complementary displacements in the coordinate \( q_i \) given the same initial phase space position \( q \) at time \( t \). Selecting sufficiently small time steps also suppresses the influence of any boundaries that may constrain the possible values of \( \Delta q_i \), returning the displacement probabilities essentially to a Gaussian form. Rearranging equation (2) and applying equation (3) yields

\[
\frac{P(+\Delta q_i|q, t)}{P(-\Delta q_i|q, t)} = \exp \left[ \frac{2f_i(q, t)}{g_i^2(q, t)} \Delta q_i \right]. \tag{4}
\]

Remarkably, this expression does no longer explicitly feature the time step \( \Delta t \). To simplify our further discussion we introduce the weighted displacement \( \Delta Z_i \), obtained from \( \Delta q_i \) via a local rescaling:

\[
\Delta Z_i := \frac{2f_i(q, t)}{g_i^2(q, t)} \Delta q_i. \tag{5}
\]

One should note that the mapping between \( \Delta q_i \) and \( \Delta Z_i \) is bijective if \( q \) and \( t \) are fixed and if \( f_i(q, t) \) does not vanish. Since the weighted displacement \( \Delta Z_i \) depends linearly on \( \Delta q_i \), it is also normally distributed. The distribution’s mean \( \mu \) is closely connected to its variance \( \sigma^2 \) via \( \sigma^2 = 2\mu = 4f_i(q, t)/g_i^2(q, t) \). This condition is necessary and sufficient for a Gaussian to satisfy the symmetry relation

\[
\frac{P(+\Delta Z_i|q, t)}{P(-\Delta Z_i|q, t)} = \exp[\Delta Z_i]. \tag{6}
\]

Equation (6) exhibits the typical form of DFTs, constraining the related probability density by an exponential symmetry and holds to lowest order in \( \Delta t \). We can interpret this result as a set of point-wise DFTs, one relation for each of the \( N \) Langevin equations. Because this equality is satisfied for every point in phase space where \( g_i \) does not vanish, it seems intuitive that one would retain this symmetry for distributions of the sum of such values as well. This allows us to extend the applicability of equation (6) to arbitrary sets of initial points and times \( Q := \{ (q_j, t_j) \} \), selected independently from the observed displacements \( \Delta Z_{ij} = \Delta Z_i(q_j, t_j) \). We then define the cumulative weighted displacement \( Z_i \) associated with set \( Q \) as

\[
Z_i := \sum_j \Delta Z_{ij}(q_j, t_j). \tag{7}
\]

In contrast to most FTs known from stochastic thermodynamics, our quantity of interest \( Z_i \) is not inherently defined as a functional over continuous trajectories. Rather, it represents a random variable conditioned to a discrete set of initial states \( Q \). One can easily derive an extension of equation (5) via convolution of the local probability densities, yielding

\[
\frac{P(+Z_i|Q)}{P(-Z_i|Q)} = \exp[Z_i]. \tag{8}
\]
Inductive reasoning allows to prove this conjecture for arbitrary $P(\Delta Z|\mathbf{q})$ if they share a common exponential symmetry. Yet this particular problem can be simplified even further by noting that the involved transition probabilities are all Gaussian. Based on the well-known convolution behavior of normal distributions, we immediately find that the relation connecting mean and variance still holds

$$
\sigma^2(Q) = \sum_j \sigma^2(q_j, t_j) = \sum_j 2\mu(q_j, t_j) = 2\mu(Q).
$$

Equipped with the retained Gaussianity of $P(Z|Q)$, this directly implies equation (8) for any fixed set of initial states $Q$ as desired. However, the validity of our reasoning hinges on the statistical independence of the initial-state—displacement pairs entering $Z_i$—a premise naturally violated by successive, time-correlated $\Delta Z_i$ taken from a trajectory. Fortunately, this apparent limitation in the potential selection or generation methods for $Q$ is easily remedied by drawing random displacements from a long trajectory or large ensemble of trajectory segments, rather than correlated ones. Such an approach fetches initial-state—displacement pairs independently from the marginal probability density $P(\Delta Z_{i,k}|\mathbf{q}_j, t_j)$ of the full trajectory distribution, without reference to any $\Delta Z_i(\mathbf{q}_j, t_j)$ with $k \neq j$. The transition from the fully-detailed trajectory picture to adequate marginal distributions is discussed more thoroughly in section 2.2. For now, we proceed to further generalize the DFT to cases where the set of initial points and times $Q$ itself is generated as an independent random variable associated with an arbitrary distribution $P(Q)$. Rather than a convolution we have to deal with a mixture of Gaussians, satisfying

$$
P(Z_i) = \int dQ \ P(Z_i|Q)P(Q)
= \int dQ \ e^{Z_i P(-Z_i|Q)P(Q)}
= e^{\mu P(-\mu)}. \tag{10}
$$

Here, the symbol $dQ$ is a shorthand for $\prod_i d\mathbf{q}_i dt_i$, integrating over all possible sets of initial points and times. Equation (10) represents the main result of this work. It describes a DFT-like relation about the size of $Z_i$ of samples is needed to properly evaluate the exponential average if the expectation value of $Z_i$ itself is large. Reducing the time step or the number of addends entering $Z_i$ can help in examining this property but one should avoid $\langle Z_i \rangle$ becoming so small that the IFT is fulfilled trivially. It is therefore recommended to have a look the distribution of $Z_i$ before assessing statistical significance. This problem is well-known for IFTs in general [27].

2.1. Multiple noise terms

Let us now consider a discretized Langevin equation in the Itô convention as usual with multiple white noise terms. Following the form of equation (1) we write

$$
\Delta \mathbf{q}_i = f_i(\mathbf{q}_i, t)\Delta t + \sum_j g_{ij}(\mathbf{q}_i, t)\Delta W_j. \tag{12}
$$
The total stochastic displacement results from a sum of rescaled, normally distributed random numbers \( g_{ij} \Delta W_{ij} \). The prefactors \( g_{ij}(q, t) \) set the variances of individual addends \( g_{ij} \Delta W_{ij} \) to \( g_{ij}^2 \delta t \). Their sum is distributed according to the convolution of the individual probability densities—all Gaussians with zero mean. The variance of this new random variable equals the sum of its constituents and one finds the statistically equivalent description

\[
\Delta q_i = f_i(q, t) \Delta t + \sqrt{\sum_j g_{ij}^2(q, t)} \Delta W_{ij},
\]

(13)

\[
\hat{q}_i = f_i(q, t) + \sqrt{\sum_j g_{ij}^2(q, t)} \eta_{ij},
\]

(14)

reducing the problem to a single noise term.

For completeness, one also ought to ask how the FT would be affected if the original noise histories were shared between multiple Langevin equations, a case which is frequently discussed in the literature [9]. Fortunately, the answer is quite straightforward: since the DFT emerges for each Langevin equation individually on basis of the normally distributed Wiener process, correlations between equations are not relevant to begin with. Even if we did not know how any of the other coordinates \( q_k \neq q_i \) describing the system evolved, as long as the coordinate in question \( q_i \) follows a Langevin equation in Itô form, it needs to obey the DFT as formulated. Conflating shared noise histories like in the above expression has no practical impact.

### 2.2. Data processing

In this section we discuss the preparation and treatment of (discrete) trajectory data in the context of the white-noise DFT. Let us define a discrete trajectory \( A \) as a time-ordered set of states \( q_i^A = q(t_i^A) \) that describe the system at times \( t_i^A = n \Delta t \). In this subsection we use the index \( n \) and superscript \( A \) to distinguish the time-ordered states that constitute a trajectory from the randomly drawn elements of \( Q \) entering the white-noise DFT. After generating an initial state according to an arbitrary distribution \( P(q_0^A) \), the trajectory evolves on basis of the full—and at least partially stochastic—system dynamics. If no stationary state exists it is useful to generate many realizations of this stochastic process, subsequently dubbed trajectory segments. Each segment \( A \) is independently initiated via \( P(q_0^A) \) and then propagated until a selected termination criterion is satisfied—for instance until a preset integration time has elapsed or \( q_i \) leaves the desired sampling region. This procedure randomly draws segments \( A \) from an unspecified trajectory distribution \( P(A = \{q_i^A\}) \). In principle, \( P(A) \) is fully determined by \( P(q_0^A) \), the underlying dynamics, and the chosen termination criteria.

Given a sufficiently large trajectory sample, drawing random displacements of a coordinate \( q_i \) from the realizations of \( P(A) \) is indistinguishable from sampling the corresponding marginal distribution directly

\[
P(\Delta q_{ij}, q_i, t_j) = \langle \delta(\Delta q_{ij} - \Delta q_i^A(t_j))\delta(q_i - q_i^A(t_j)) \rangle_A,
\]

(15)

where \( \langle \cdot \rangle_A \) indicates the average over \( P(A) \) and \( \Delta q_i^A(t_j) \) equals the difference \( q_i^A(t_j + \Delta t) - q_i^A(t_j) \).

Repeating this procedure generates arbitrarily large sets of single coordinate displacement \( \Delta q_{ij} \) and initial state \( (q_j, t_j) \) pairs without reference to any correlations potentially present at the trajectory level, inducing

\[
P(\{\Delta q_{ij}, q_i, t_j\}) = \prod_{ij} P(\Delta q_{ij}, q_i, t_j).
\]

(16)

This probability density instantly yields \( P(\{\Delta Z_{ij}\}, Q) \) via a simple coordinate transformation if \( q_i \) is governed by a Langevin equation in its Itô representation as usual. The associated conditional probability evidently retains the product form of equation (16) and thus implies

\[
P(\Delta Z_{ij} \mid Q, \{\Delta Z_{k \neq ij}\}) = P(\Delta Z_{ij} \mid q_i, t_j).
\]

(17)

This condition is sufficient to prove the DFT’s validity for \( Z_i = \sum_j \Delta Z_{ij} \) via convolution as mentioned in the context of equation (10). After ensuring the independence of all \( \Delta Z_{ij} \), the application of the DFT requires no further knowledge about any of the distributions discussed here.
While the delineated line of reasoning adequately demonstrates the theoretical usability of trajectory data to generate independent displacement sets, it also inherently assumes the availability of enough data to emulate the marginal distribution of equation (16). Yet, whenever a system consists of many degrees of freedom $q$, acquisition of the necessary amount of statistics to this end can become a daunting task. Fortunately, generating a representative sample of the marginal distribution is a sufficient condition, but not a necessary one: most degrees of freedom enter the white-noise DFT only through the local weight $w_i(q, t) = 2\lambda(q, t) / g_i^2(q, t)$ as associated with coordinate $q$. Therefore, obtaining a representative sample for the marginal distribution of the weight along with the displacement is sufficient, i.e. $P(\Delta Z_i, w_i)$ or $P(\Delta Z_i, w_i)$. In other words, the DFT for any $q_i$ is only concerned with the distribution of two scalars irrespective of the system dimensions, leading to much less stringent data requirements in practice.

Imposing stricter termination criteria on the trajectory segments may facilitate data collection even further if desired. Finally, we wish to point out that the idea presented in this section easily generalizes to samples obtained from single long trajectories—for instance by splitting the data into segments with a given duration. The distribution of initial states $P(q_i^0)$ for the virtual trajectory segments then approximately corresponds to the stationary phase space density derived from an infinitely long trajectory. This holds true even for non-ergodic systems. Alternatively, one could formulate effectively the same argument without reference to the trajectory ensemble if a stationary state is available to give a marginal distribution of the weight $P(\Delta Z_i, w_i)$. The numerical examples covered in the upcoming sections showcase both single- and multiple-trajectory sampling.

### 3. Overdamped systems

Let us now move on to some practical examples, starting with a setup frequently investigated in the context of stochastic thermodynamics: a Brownian particle in a potential $U(x, \lambda)$ with an external time-dependent protocol $\lambda(t)$. Applied throughout the interval $t \in [t_1, t_2]$, the protocol begins with a system in equilibrium given $\lambda_1$ and ends at $\lambda_2$. We will use this example to illustrate some prominent differences to FTs based on time-reversal, in particular the Crooks theorem [4]. For simplicity we first examine a one-dimensional system governed by the position’s equation of motion

$$\dot{x} = -\frac{1}{m\gamma} \frac{\partial U(x, \lambda)}{\partial x} + \sqrt{\frac{2k_B T}{m\gamma}} \eta(t). \tag{18}$$

The gradient of the potential $U(x, \lambda)$ provides the deterministic force, whereas the diffusivity results from the environmental temperature $T$, the particle mass $m$, and the damping coefficient $\gamma$. The weighted displacements in this setup assume the form

$$\Delta Z_x = -\frac{1}{k_B T} \frac{\partial U}{\partial x} \Delta x. \tag{19}$$

According to conventional definitions in this field [2, 5, 28], variations in the energy due to a change in the internal degree(s) of freedom are associated with the heat $Q$. Energy changes caused by the protocol are instead summarized into the work $W$. Crooks has shown that the ratio between the probabilities of performing an amount of work $W$ in the forward trajectory and an amount $-W$ in the backward trajectory is given by $\exp[(W - \Delta F)/k_B T]$ if the dynamics respect detailed balance [4]. This law holds for forward and backward processes starting in equilibrium states whose free energies are $F(\lambda_1)$ and $F(\lambda_2)$ respectively, with $\Delta F = F(\lambda_2) - F(\lambda_1)$.

However, in the context of the white-noise DFT of equation (10) the stochastic increments must be interpreted in an Itô sense. Equation (19), in general, deviates from the potential energy difference between starting and final position. Therefore $\Delta Z_x$ corresponds to neither heat nor work, but contains additional terms not covered by conventionally relevant thermodynamic measures. That being said, the conditions for the noise-based DFT of equation (10) are more easily fulfilled in practice. Besides individually holding for each and every point in phase space with finite diffusivity, the noise-symmetry DFT provides a separate relation for each spatial coordinate in higher-dimensional systems. Furthermore, Langevin dynamics abide by this symmetry even if no stationary state exists, directly rendering detailed balance irrelevant as well. The DFT can also deal with systems that feature coordinates evolving in a non-Markovian fashion. The removal of these typical requirements appearing in the derivation of time-reversal FTs equips our theorem with a wide range of applicability. For instance, non-conservative forces $\mathcal{F}$ in arbitrary dimensions modify our
overdamped equation of motion in the following way

\[
\dot{x}_i = -\frac{1}{m\gamma} \frac{\partial U(x_i, t)}{\partial x_i} + \frac{F_i(x_i, t)}{m\gamma} + \sqrt{\frac{2k_B T}{m\gamma}} \eta_i(t).
\]  

(20)

The associated quantity

\[
\Delta Z_i = \frac{1}{k_B T} \left( F_i - \frac{\partial U}{\partial x_i} \right) \Delta x
\]

(21)

then includes an additional term that resembles the infinitesimal work performed by the non-conservative force, but again cannot be interpreted as such due to the Itô convention. However, the theorem can be applied without any special treatment or additional considerations involved.

3.1. Linear non-conservative forces

As a practical example for a system to our knowledge not already covered by other DFTs, let us direct our attention to a particular two-dimensional system. Specifically, we discuss an overdamped particle subjected to a linear non-conservative force. The field lines of the non-conservative force are circular, driving the particle to revolve around the center. Due to this external force, energy is continuously fed into the system and subsequently dissipated. The particle’s dynamics is governed by the equations

\[
\begin{align*}
\dot{x} &= -\frac{ay}{m\gamma} + \sqrt{\frac{2k_B T}{m\gamma}} \eta_x, \\
\dot{y} &= \frac{ax}{m\gamma} + \sqrt{\frac{2k_B T}{m\gamma}} \eta_y,
\end{align*}
\]

(22a, 22b)

where \(x\) and \(y\) constitute the particle coordinates and \(a\) sets the external force’s strength. The deterministic forces increase in magnitude as the particle moves away from the origin. Without the presence of a confining potential the particle can drift outward indefinitely. In other words the system has no stationary state. Our example system falls into the class of systems covered by equation (20). Here, we have simulated an ensemble of trajectory segments that are terminated if the particle leaves a selected perimeter around the origin. Trajectories begin at the origin with a velocity drawn from the Maxwell–Boltzmann distribution. Irrespective of this initial condition, the termination criterion, and the non-stationarity, the noise-symmetry DFT must hold.

Each Langevin equation implies a symmetry relation for the respective weighted position displacements \(\Delta Z_x\) or \(\Delta Z_y\) analogous to equation (21), which the computed histograms shown in figure 1 indeed fulfill. Due to the problem’s symmetry, \(Z_x\) and \(Z_y\) are identically distributed. We quantify the histogram symmetry by looking at the logarithmic probability ratio, henceforth dubbed the symmetry function

\[
\Sigma(Z) := \ln \frac{P(\cdot Z)}{P(\cdot = Z)}.
\]

(23)

This measure is identical to \(Z\) itself if the noise-symmetry DFT is satisfied, becoming a linear function through the origin with unity slope. The lower panel of figure 1 compares the computed \(\Sigma(Z)\)-curve against this theoretical prediction, demonstrating excellent consistency.

4. Underdamped systems

In this section, we will apply the white-noise DFT to simulated trajectories of underdamped systems. We use our results to explore the versatility of the white-noise DFT for systems that feature multiplicative noise histories, non-Markovian dynamics, and variables not invariant under time-reversal like the velocity.

4.1. Energy diffusion

Particles subjected to conservative forces in a thermal bath constitute a particularly important and ubiquitous type of Langevin dynamics in physics. Subsequently, the index \(i \in \{1, \ldots, N\}\) numbers the coordinates \(x_i\) and velocities \(v_i\) of individual particles in any particular spatial dimension. The evolution of the degrees of freedom is described by
Figure 1. Non-conservative system. A non-conservative force acts on an overdamped particle in a two-dimensional space. The non-conservative driving becomes stronger as the particle moves outward. The particle’s indefinite escape is avoided by resetting the trajectory to the origin once the distance to the origin is greater than \( s_m = \sqrt{\frac{k_B T}{a}} \). Quantifying time in terms of a characteristic physical timescale \( t_0 = \sqrt{\frac{m}{a}} \), the damping is given by \( \gamma = 100 t_0^{-1} \). (a) Histograms of the cumulative weighted position increments \( Z \). Each value of \( Z \) is accumulated over 50 uncorrelated displacements with time step \( \Delta t = 0.1 t_0 \). The histograms associated with the position-displacements in the \( x \)- and \( y \)-direction coincide due to the system’s symmetry. (b) Logarithmic probability ratio \( \Sigma(Z) \). The results are in excellent agreement with the theoretical reference curve (dashed line).

Figure 2. White-noise DFT of the total energy in a conservative system. A set of non-interacting particles is trapped in a harmonic potential. (a) Histograms for three different particle numbers \( N = 1, 4, 10 \) computed by tracing the total kinetic and potential energy. Each realization of \( Z_E \) consists of 250 increments \( \Delta Z_E \) drawn from a long trajectory. The parameters of the simulation read \( \Delta t = 2 \times 10^{-3} t_0 \) and \( \gamma = t_0^{-1} \), given in units of \( t_0 = \sqrt{m/k_B} \). (b) Logarithmic probability ratio \( \Sigma(Z_E) \) as given by equation (23). The histograms respect the DFT for all particle numbers \( N \) considered.

\[
\dot{x}_i = v_i, \quad (24) \\
m\dot{v}_i = -\frac{\partial U(x)}{\partial x_i} - m\gamma v_i + \sqrt{2m\gamma k_B T}\eta_i. \quad (25)
\]

The potential \( U(x) \) only depends on the particle coordinates \( x_i \), but is otherwise arbitrary. Furthermore, energy is dissipated to the environment through a linear damping force proportional to the particle velocity \( v_i \) in each direction. Each of these relations results in a separate DFT for the weighted velocity displacements

\[
\Delta Z_i = -\frac{1}{\gamma k_B T} \left( \frac{\partial U(x)}{\partial x_i} + m\gamma v_i \right) \Delta v_i. \quad (26)
\]

However, it is also possible to consider the symmetries of composite measures, i.e. functions of the particle positions and velocities. The total energy of the system may represent the most prominent example in this context. It can be written as the sum of the potential and the total kinetic energy \( K \),

\[
E = U(x) + \sum_i \frac{mv_i^2}{2}. \quad (27)
\]

A Langevin equation for the total energy \( E \) with the appropriate integration convention is easily derived by invoking Itô’s lemma [25]:

\[
\frac{dE}{dt} = \sum_i \left( \gamma k_B T - m\gamma v_i^2 + v_i \sqrt{2m\gamma k_B T}\eta_i \right) = N\gamma k_B T - 2\gamma K + \sqrt{4\gamma K k_B T}\eta_E. \quad (28)
\]

Here, we have recycled the reasoning used in equation (14) to summarize multiple Gaussian white noises into a single one. As we can see from equation (28), the changes of the total energy only depend on the

\[
\dot{x}_i = v_i, \quad (24) \\
m\dot{v}_i = -\frac{\partial U(x)}{\partial x_i} - m\gamma v_i + \sqrt{2m\gamma k_B T}\eta_i. \quad (25)
\]

The potential \( U(x) \) only depends on the particle coordinates \( x_i \), but is otherwise arbitrary. Furthermore, energy is dissipated to the environment through a linear damping force proportional to the particle velocity \( v_i \) in each direction. Each of these relations results in a separate DFT for the weighted velocity displacements

\[
\Delta Z_i = -\frac{1}{\gamma k_B T} \left( \frac{\partial U(x)}{\partial x_i} + m\gamma v_i \right) \Delta v_i. \quad (26)
\]

However, it is also possible to consider the symmetries of composite measures, i.e. functions of the particle positions and velocities. The total energy of the system may represent the most prominent example in this context. It can be written as the sum of the potential and the total kinetic energy \( K \),

\[
E = U(x) + \sum_i \frac{mv_i^2}{2}. \quad (27)
\]

A Langevin equation for the total energy \( E \) with the appropriate integration convention is easily derived by invoking Itô’s lemma [25]:

\[
\frac{dE}{dt} = \sum_i \left( \gamma k_B T - m\gamma v_i^2 + v_i \sqrt{2m\gamma k_B T}\eta_i \right) = N\gamma k_B T - 2\gamma K + \sqrt{4\gamma K k_B T}\eta_E. \quad (28)
\]

Here, we have recycled the reasoning used in equation (14) to summarize multiple Gaussian white noises into a single one. As we can see from equation (28), the changes of the total energy only depend on the
dissipation and the fluctuations due to the environment. In the absence of environmental forces, the potential energy $U(x)$ could only be converted into kinetic energy and vice versa, leaving the total energy constant. One immediately obtains the weighted energy displacements

$$\Delta Z_E = \frac{Nk_B T - 2k}{2k_B T K} \Delta E.$$  

(29)

Irrespective of any details concerning the potentially very complex interactions in the system, energy displacements have to obey the simple law

$$\frac{P(+\Delta Z_E)}{P(-\Delta Z_E)} = \exp \left( \frac{Nk_B T - 2k}{2k_B T K} \Delta E \right).$$  

(30)

There are several points of interest here: contrary to our previous examples that tacitly investigated coordinates with free boundary conditions, the system energy often exhibits a lower bound. The presence of non-trivial boundary conditions influences the displacement probabilities in the immediate vicinity of boundaries, distorting their Gaussian statistics. Despite inducing discrepancies in the DFT, one can always find a sufficiently small time step to recover its validity if the distance to the boundary is finite. Additionally, since the DFT symmetry applies locally, one can arbitrarily prescribe the regions in parameter space from which initial points are drawn, e.g. the exclusive sampling of displacements with starting energies greater than any desired threshold. In practice, the influence of boundaries on the theorem can frequently be neglected or eliminated by choosing an adequate time discretization and sampling region. Furthermore, it is noteworthy that the energy’s evolution features multiplicative noise, i.e. a diffusivity that depends on $E$ itself via $K$.

To support the validity of the white-noise DFT for this multiplicative noise case, we consider a system of $N$ non-interacting particles in one dimension confined by a harmonic potential $U(x) = \frac{kx^2}{2}$. After generating a long trajectory as test data, we have evaluated the DFT according to equation (30), sampling 250 randomly drawn $\Delta Z_E$ for each realization of $Z_E$. We repeat our analysis for different numbers of particles $N$, achieving the distributions shown in figure 2(a). These distributions differ significantly from one another, yet they all fulfill the white-noise DFT, as we can appreciate from figure 2(b).

### 4.2. Delayed feedback

In this subsection we investigate an example of non-Markovian dynamics to which the white-noise DFT can be applied. We consider a one-dimensional harmonic oscillator in a thermal bath that is additionally subjected to an external feedback force $z(t)$. The feedback is proportional to the particle’s previous position and applied with a delay of constant size $\tau$, rendering the system non-Markovian as intended. For simplicity, we assume that the feedback force is linear like its conservative counterpart, leading us to the set of equations shown below

$$\dot{x} = v,$$  

(31a)

$$z = x(t - \tau),$$  

(31b)

$$m \dot{v} = -kv + lx + \sqrt{2m\gamma k_B T} \eta.$$  

(31c)

The spring constant $k$ is assumed to be positive, whereas the proportionality constant $l$ of the feedback force may be selected arbitrarily. This system constitutes the first example that is not solely represented by Langevin equations. The evolution of the associated density in probability space cannot be described by a single Fokker–Planck equation either. Nonetheless, our symmetry relation remains applicable for the velocity Langevin equation (31c), as confirmed by the simulation results depicted in figure 3. We show histograms for three different delay times $\tau$, ranging from values much shorter than the potential’s characteristic oscillation time $2\pi \sqrt{m/k}$ to delays that are in the same order of magnitude. Despite leading to very distinct distributions of the displacement variable $\Delta Z_E$, all of the computed distributions clearly share the predicted symmetry as highlighted in figure 3(b). The full evolution of this system depends on a prior point in time and is thus non-Markovian. Furthermore, the delay time may assume arbitrarily large values. For systems such as these no other detailed or even just integral FTs exist to our knowledge.

### 5. Parameter inference

Contrary to FTs based on time-reversal, the quantities $Z$ appearing in the white-noise DFT are not directly tied to entropy, heat or work. In fact, $Z$ only describes coordinate displacements that are reweighted by the
local competition of deterministic and stochastic forces. The weight takes the value $2f_i/g_i^2$ as shown in equation (5), equal to the ratio of the local deterministic force and diffusivity. Assigning physical meaning to $Z$ beyond this is difficult. Yet it is precisely this difference from conventional FTs that opens up the large variety of systems it can be applied to. In this section, we will demonstrate how this versatility leads to an efficient and flexible method to infer unknown parameters of a system, and how parameters of an effective/simplified model can be optimized to be as consistent as possible with some reference data. For notational simplicity we exclusively discuss one-dimensional systems with time-independent forces in the following. Note however that every aspect of our subsequent considerations is easily generalized to time-dependent multi-dimensional systems.

5.1. Uniqueness

Assume we are given a trajectory of a coordinate $q_i$, which evolves according to a Langevin equation $\dot{q} = f(q) + g(q)\eta$, yet is modelled on the basis of a different one $q = f_m(q) + g_m(q)\eta$. Discrepancies of this kind are very common in practice and may stem from experimental uncertainties in one or more parameters or the application of simplified models like the overdamped approximation. The question we want to answer in the following is: how do differences between model and true dynamics quantitatively affect the white-noise DFT?

Let us first define the model displacement $\Delta Q_m := 2f_m(q)\Delta q / g_m^2(q)$ and deviation factor $r(q) := \Delta Z / \Delta Z_m$. Note that the latter does not depend on $\Delta q$. We restrict our analysis to positions $q$ where neither the model’s deterministic term nor the diffusivity vanish, guaranteeing that $r$ exists and is finite. With these definitions and the Gaussianity of $\Delta Z_m$ inherited from $\Delta q$, it is easy to see that the distribution $P(\Delta Z_m)$ satisfies the relation

$$P(\Delta Z_m) = e^{r(q)\Delta Z_m}.$$  \hspace{1cm} (32)

We can consequently extract the correction factor $r(q)$ from the slope of our local symmetry function $\Sigma(\Delta Z_m)$. Alternatively, since the underlying distribution is still Gaussian one can take advantage of the numerically more stable expression $r(q) = 2\mu_m(q) / \sigma_m^2(q)$ instead, where $\mu_m$ and $\sigma_m^2$ designate the mean value and variance of $\Delta Z_m$ respectively. In contrast, the true $\Delta Z$ derived from $f(q)$ and $g(q)$ would fulfill the relation $2\mu(q) / \sigma^2(q) = 1$ for all $q$ per construction. Any discrepancy between the true dynamics and the model at positions $q$ is thus uniquely inferable via up to a common prefactor in $f(q)$ and $g^2(q)$. To put it another way: by invoking the equality $r(m) / g_m^2 = f / g^2$ we can determine $h(q)f(q)$ and $h(q)g^2(q)$, where $h(q)$ is an arbitrary real number. This multiplicity is a consequence of the DFTs time step independence. Although the above equation can be employed to extract $f / g^2$ for any initial position, this approach is impractical as one would need to gather sufficient statistics for each $q$. Even if enough data were available, it would be more practical to calculate the average displacement and its variance to obtain $f(q)$ and $g^2(q)$ without the delineated multiplicity. Nonetheless, the DFT can be used in other ways, which we will discuss shortly. Before doing so, however, it is instructive to investigate the distribution of a sum of model displacements $Z_m = \sum_j \Delta Z_m(q_j)$ taken from different positions in phase space $Q = \{q_j\}$. As a convolution, the associated distribution is another Gaussian with mean $\mu_m(Q) = \sum_j \mu_m(q_j)$ and variance $\sigma_m^2(Q) = \sum_j \sigma_m^2(q_j)$. Therefore we obtain
We find that the correction factor of the sum equals an average which is weighted by its constituents’ variances. The exponential symmetry is retained even for arbitrary sums of \( \Delta Z_m \), provided that one fixes the initial conditions of the displacements. We discuss the implications of dropping this constraint in the next section.

**5.2. Convergence**

It is far more convenient to consider the statistics of \( \Delta Z_m \) or \( Z_m \) without sorting the displacements in a discrete/discretized trajectory w.r.t. their initial positions. The corresponding distribution is no longer a Gaussian, but rather a mixture of Gaussians that can be written in the decomposed form

\[
P(\Delta Z_m) = \int dq \ P(\Delta Z_m|q)P(q),
\]

where \( P(q) \) is the probability of observing initial condition \( q \) (or a set of initial points) throughout the dataset. As justified in section 2.2, this probability can be defined even if no stationary state exists. The conditional probability \( P(\Delta Z|q) \) satisfies equation (32). We proceed to have a look at the mixture distribution’s symmetry

\[
\frac{P(+\Delta Z_m)}{P(-\Delta Z_m)} = \int \frac{d\Delta Z_m e^{\Delta Z_m}P(-\Delta Z_m|q)P(q)}{\int dq P(-\Delta Z_m|q)P(q)}.
\]

The above expression takes the form of a conditional expectation value for \( e^{\Delta Z_m} \), constraining the density to \( -\Delta Z_m \). No simple exponential symmetry exists if \( r(q) \) is not constant throughout the phase space portion considered. However, it is easy to show that for positive \( \Delta Z_m \) the symmetry function \( \Sigma(\Delta Z_m) \) is bounded by

\[
\min_q \ r(q) \Delta Z_m \leq \Sigma(\Delta Z_m) \leq \max_q \ r(q) \Delta Z_m.
\]

For \( \Delta Z_m < 0 \) one obtains a similar relation, but the inequalities are reversed, implying that \( \Sigma \) is encased by two linear functions intersecting at the origin. Fortunately, we can drastically improve upon providing simple bounds by considering sums of displacements drawn randomly from the trajectory or dataset. Again, drawing them randomly serves the purpose of removing correlations between subsequent \( r(q) \) if applicable, sampling the marginal distribution instead. As the number of terms in the sum is increased, we approach the weighted average of the dataset

\[
R = \int \frac{dq \ r(q)\sigma_m^2(q)P(q)}{\int dq \ \sigma_m^2(q)P(q)},
\]

equivalent to the limit of infinite samples in equation (33). While the potential minimum and maximum values of \( r(Q = \{q_j \}) \) coincide with the respective single point values, the distribution becomes sharply peaked around \( R \) for a large number of addends. The symmetry function \( \Sigma \) then quickly approaches the value \( RZ_m \) for any finite \( Z_m \), converging toward a linear expression again. The slope of a linear fit to \( \Sigma \) can therefore be seen as an adequate quantifier for the average consistency of a dynamical model with the data. This statement can also be adapted to selected portions of phase space, if desired.

On basis of the delineated properties we propose a simple, yet versatile optimization algorithm consisting of three steps: (1) postulate a model with a set of free parameters \( k = \{k_i \} \), (2) compute the distribution of randomly added displacements of the trace to be analyzed and extract \( \Sigma(Z(k)) \), (3) vary the parameters \( k_i \) to minimize the deviation between \( \Sigma(Z(k)) \) and the theoretical reference \( Z(k) \), repeating step (2) for each test value of \( k \). One can quantify this deviation via the (initial) slope of our symmetry function \( R(k) \) or other simple measures.

Our method is closely related to Bayesian parameter inference, as both approaches aim to replicate the expected statistics of Gaussian white noise. Indeed, by extracting the noise history or increments of the Wiener process from the stochastic displacements, one can also optimize the parameters by maximizing the associated likelihood function \( L(k) \) \cite{29}. While the computational effort necessary for both approaches scales linearly with the number of displacements available, they offer different advantages. On a general basis, however, the performances of these two methods are quite similar and they are almost equally versatile.
5.3. Comparison with Bayesian inference
Let us proceed to elaborate upon several important differences between direct Bayesian inference and the DFT-method. Bayesian inference is able to identify \( f \) and \( g \) directly as indicated, avoiding the multiplicity of shared prefactors. Yet it does feature some arbitrariness in the form of the selected prior, influencing the final result. The DFT analysis in turn drops any explicit dependence on the time step, a very sensitive parameter in Bayesian estimates that may itself be subject to uncertainties. Also, whether the incurred multiplicity is actually relevant for the optimization depends on the set of parameters investigated. Only if one of the parameters leads to a modification of \( f \) and \( g \) that leaves \( f' / g^2 \) invariant is there additional information to be extracted with the Bayesian approach. The DFT-method additionally offers a direct quantification of how accurate a given set of parameters is via \( R \)—more so than the value of the likelihood function \( L \) in Bayesian estimation. Except for statistical errors, \( R \) becomes unity at the position of the optimal parameter set. The procedure can be terminated once the difference to the reference value is smaller than a threshold value \( \epsilon \).

Next, we underline that whenever \( f_m \) and \( g_m \) are continuous in \( k_i \), both \( R(k) \) and \( L(k) \) depend continuously on \( k_i \) as well, covering many practically relevant cases. In such scenarios, precise knowledge about our target value \( R = 1 \) is valuable in reducing the computational cost expended to reach the optimal parameter set. For instance, after finding two parameter sets \( k_1 \) and \( k_2 \) such that \( R(k_1) < 1 < R(k_2) \) holds, one can close in on the optimum with a bisection algorithm. This procedure reaches a minimum far faster than the Markov chain Monte Carlo approaches typically used with Bayesian inference. The possibility of using bisection with similar Bayes methods is not precluded of course, but necessitates the identification of a concave (boundary) region of \( L(k) \). Locating such regions is more computationally expensive as it inherently requires a greater set of likelihood evaluations, especially in higher-dimensional parameter spaces. Even if bisection was used in the Bayesian inference after this preparation step, knowledge about \( R \) can be utilized to accelerate convergence more effectively: the slope provides additional information to gauge the distance to the optimum and potentially lower the number of iterations necessary to reach it. Nevertheless, although the total computational cost scales identically, individual iterations are slightly more expensive in the DFT-method due to the extraction of the symmetry function \( \Sigma(Z(k)) \) from the histograms and its linear fit. One can circumvent this step only under the condition that the set of initial states \( Q \) is fixed, at which point \( R \) is efficiently inferred from the usual expression \( 2\mu(Q)/\sigma^2(Q) \).

Lastly, the methods significantly differ in their treatment of false positives—that is local maxima of \( L(k) \) or \( R = 1 \) parameter sets that are incompatible with the system’s true dynamics, even as a simplification. Broadly speaking, one does not necessarily reach \( R = 1 \) for any set \( k \) in cases of model incompatibility, but can generally still locate a maximum in the likelihood function. A less efficient optimization in the sense that one searches a minimum of \( |R - 1| \) remains possible, but the DFT-based algorithm communicates such discrepancies more directly. In the presence of multiple Langevin equations the parameter set must even satisfy several independent DFT relations, making it the much more stringent requirement. Moreover, dynamics not governed by Gaussian white noise generally do not lead to linear \( \Sigma \)-curves, providing an additional tool for error assessment. Nevertheless, false positives can in principle occur in either algorithm. One strategy for their identification consists in examining one or more phase space subregions of initial states. Checking for continued consistency is once more computationally cheaper for the DFT-method, as one only needs to find the \( R \) value for the test parameter set and see whether the selected termination criterion \( |R - 1| < \epsilon \) is still satisfied. The Bayes approach necessitates a number of evaluations in the immediate vicinity again to confirm that a local minimum remains approximately at the tested position.

5.4. Example: delayed system
Let us revisit one of our previous examples—the harmonic oscillator with delayed feedback. The FT offers an approach to infer most of the system’s parameters, for instance \( k, l \), and \( \tau \) from equation (31). Here, we check the accuracy of our optimization algorithm by extracting the delay time \( \tau \) that prescribes the degree of non-Markovianity of the system, assuming that the remaining parameters are known beforehand. The displacement weights \( f(x, v, z) / D \) explicitly depend on the delay time, implying that remnant discrepancies to the theoretical reference will only disappear at \( \tau = \tau_0 \) for arbitrary phase space regions and infinitesimal time steps. We generate trajectories with a given delay \( \tau_0 \) and subsequently compute the \( \tau \)-dependent weighted displacement distributions \( P(Z, \tau) \). In this artificial example we have computed the displacement statistics of \( \tau \)-values on a uniformly-spaced grid ranging from zero delay to \( 2\tau_0 \).

Figure 4(a) demonstrates how the symmetry of the extracted distributions can deviate from the theoretical prediction if the test value for \( \tau \) strays significantly from \( \tau_0 \). Nonetheless, the logarithmic probability ratio \( \Sigma(Z) \) remains well-approximated by linear fitting functions: as predicted in section 5.2, increasing the number of adds induces convergence to linear \( \Sigma(Z) \)-functions even for deviating models. Here we use the same number of displacements per \( Z_\tau \)-value as in figure 3.
Figure 4. Inference of delay time $\tau$. (a) Logarithmic probability ratio $\Sigma$ of the delayed system described by equation (31) for three different test delay times, equal to $\tau/\tau_0 = 0, 1, 2$. Only the true delay time $\tau = \tau_0$, used for the generation of the trajectory data, results in a graph consistent with the theoretical reference $R(\tau) = 1$ (shown by the black dashed line). Irrespective of the test value $\tau$, the initial behavior of $\Sigma$ is well-described by a linear approximation. Here, we have analyzed the $\tau_0/t_0 = 0.16$ data of figure 3. (b) Initial slope $R$ of $\Sigma$ as a function of the delay time $\tau$. The theoretically predicted slope of unity is found precisely at the true delay time $\tau_0$.

Figure 5. Slope of $\Sigma(Z_x)$ as a function of the signal noise parameters. The heatmaps showcase the deviation w.r.t. the expected slope $R = 1$ as a function of the noise’s magnitude $\sigma_{\epsilon,x}$ and correlation time $\tau_{\epsilon,x}$ as it contributes to the velocity signal. The respective values are given in terms of the unbiased velocity’s standard deviation $\sigma_v$ and decorrelation time $\tau_v = 1/\gamma$. The two panels only differ in the magnitude of the position error, equaling $\sigma_{\epsilon,x} = 0$ in (a) and $\sigma_{\epsilon,x} = \sigma_x/2$ in (b).

Figure 4(b) showcases the extracted initial slopes $R$ as a function of the delay time $\tau$. Corresponding to the average deviation factor $r(x, v, z)$ of the sample, we find that the theoretically predicted slope of $R = 1$ is only obtained for the true delay time $\tau_0$. Any remnant discrepancies at this point are negligible and caused primarily by statistical uncertainties.

6. Signal contamination

The practical relevance of the white-noise DFT lies in the ability to infer parameters from simulated and experimental trajectories. Especially in the latter case it may help in extracting otherwise elusive information from the short-time evolution of the system. Yet, this application inherently demands the availability of sufficiently accurate displacement data. While our inference method can aid in identifying accurate dynamical descriptions and extracting adequate parameter sets, measurement noise remains a fundamentally unavoidable source of errors. It therefore merits a more thorough qualitative discussion.

We subsequently investigate the sensitivity of the DFT with respect to noise as a function of its standard deviation $\sigma_\epsilon$ and correlation time $\tau_\epsilon$. Specifically, we examine an underdamped harmonic oscillator in a thermal bath described by

$$\dot{x} = v,$$  \hspace{1cm} (38a)

$$\dot{v} = -\gamma v - \omega_0^2 x + \sqrt{2D}\eta.$$  \hspace{1cm} (38b)
Its position and velocity are subject to Gaussian measurement uncertainties $\epsilon_x$ and $\epsilon_v$, generating signals $x' = x + \epsilon_x$ and $v' = v + \epsilon_v$. We base the evolution of the signal errors on Ornstein–Uhlenbeck processes

$$\dot{\epsilon}_x = -\frac{\epsilon_x}{\tau_{\epsilon,x}} + \sqrt{\frac{2\sigma^2_{\epsilon}}{\tau_{\epsilon,x}}} \eta_{\epsilon,x},$$

$$\dot{\epsilon}_v = -\frac{\epsilon_v}{\tau_{\epsilon,v}} + \sqrt{\frac{2\sigma^2_{\epsilon}}{\tau_{\epsilon,v}}} \eta_{\epsilon,v},$$

allowing us to freely prescribe the noise magnitudes $\sigma_{\epsilon,x}$, $\sigma_{\epsilon,v}$, and correlation times $\tau_{\epsilon,x}$, $\tau_{\epsilon,v}$. The Ornstein–Uhlenbeck toy model further offers the unique advantage of ensuring that even the corrupted signal exhibits Gaussian transition probabilities, retaining the linear DFT curves as discussed in the main text.

In a real experiment the velocity signal would be typically extracted from the position via finite difference methods or fitting procedures. For greater simplicity and generality we make no particular assumption on the signal retrieval method. Here we consider the errors $\epsilon_x$ and $\epsilon_v$ to originate from uncorrelated red noises. Without any information about the error’s value, the signals $x'$ and $v'$ become substitutes for the real position $x$ and velocity $v$ and we evaluate the DFT of the corresponding velocity Langevin equation, i.e. for the quantity

$$\Delta Z' := \frac{-\gamma v' - \omega^2_x x'}{D} \Delta v'. \tag{40}$$

Figure 5 outlines the impact of the signal velocity's decorrelation time and noise magnitude. As expected, the slope of the logarithmic probability ratio curves returns to unity if the signal errors become negligibly small. Deviations appear and become evermore prominent as we increase in noise intensity. However, longer noise correlation results in a slower onset of this effect, partially mitigating the impact of the signal error. This observation is likely owed to the fact that errors in the DFT originate from two sources: the phase space point before and after the displacement. The ensemble of measured initial points remains invariant under changes in $\tau_x$, however the displacement itself is deeply sensitive w.r.t. this parameter. Longer correlation times diminish the signal error’s change between time steps and therefore allow for the extraction of a more accurate estimate of the velocity difference—but not of the velocity itself. Although white noise finds more frequent application as an error model in practice, it is precisely this correlation that leads to more faithful $Z'_v$-histograms. Nonetheless, as one steps further into the regime of strong corruption, we seemingly encounter a threshold beyond which the DFT breaks down irrespective of $\tau_x$. This occurs at a signal-to-noise ratio of approximately one, i.e. when the signal corruption starts to overshadow the physical movement completely.

Although qualitatively unchanged, this effect is also strongly sensitive to changes in the time step used throughout the analysis. Small time steps lead to reduced displacement magnitudes as well. Therefore fast fluctuations caused by signal errors may bury the physical information more easily. This circumstance imposes a lower limit on the time step size for non-continuous noise histories. Naturally, a more general limitation restricts the use of larger $\Delta t$: the validity of the DFT is inherently tied to short timescales at which the average drift and diffusivity reliably follow from the linear and quadratic displacement. This newly added complication of a lower bound due to measurement noise can severely reduce the range of viable $\Delta t$. Using extremely small time steps can only be recommended for continuous noise histories, where most benefits are achieved only after crossing the correlation time threshold. Applications that seek to take advantage of the white-noise DFT need to mitigate this issue more effectively, thus a high signal-to-noise ratio becomes indispensable.

7. Discussion

In this work we have introduced a class of FT-like symmetries applicable to all kinds of Langevin dynamics. Contrary to established FTs central to the field of stochastic thermodynamics, these are not based on any form of time-reversal, but rather on the statistics of the underlying Gaussian white noise. Consequently, the measures appearing in the theorems are not associated with entropy or other typical thermodynamic quantities either. They instead relate to short-time displacements reweighted by the local competition between deterministic and stochastic forces. The noise-symmetry DFTs frequently hold even for systems that remain elusive to their time-reversal counterparts: they neither require any knowledge about stationary states nor the existence thereof and even apply to non-Markovian systems that feature long-time memory. The only requirement is that Langevin equations describe the evolution of one or more degrees of freedom.
This versatility leads to a powerful time-step—indeed independent inference algorithm that may be utilized to extract various parameters from computational and experimental trajectory data. The method is inherently connected to Bayesian inference, but it has built-in termination criteria and features information that may help accelerate convergence to optimal parameter sets.

Having focused primarily on computational examples, it would be interesting to see how the algorithm fares in extracting particularly elusive parameters from noisy experimental data—especially in setups where trajectories are readily available. Examples may include levitated nanoparticles, where one could measure nonlinearities of the confining potential, or protein-folding dynamics, to gauge the quality of reaction coordinates and extracted free energy landscapes. Such applications would greatly expand the preexisting catalog of FT-based analysis tools and high-precision measurement techniques [30–34]. Moreover, the white-noise DFT may also aid in assessing the relevance of selected parameters and variables on different timescales. For instance, the frequently employed overdamped representation of particle dynamics is bound to break down for sufficiently small time steps. The DFT can help in estimating the timescale below which the neglected degrees of freedom become relevant again to the description of the particle’s propagation. Such a procedure could also help in the selection of adequate integration time steps for numerical applications.

Finally, it should be noted that white noise itself breaks down as an approximation at sufficiently small timescales. This circumstance may motivate the theorem’s generalization to other forms of noise, be it white noise with non-Gaussian statistics or stochastic forces with finite memory—including generalized Langevin equations with non-trivial memory kernels [35, 36]. Furthermore, in its current form the noise-symmetry can only be applied to classical trajectories. It would be interesting to see whether comparable laws could be unearthed for noisy quantum systems, similar to the extension of time-reversal FTs to the quantum regime [37, 38].

Acknowledgments

This research was supported by the Austrian Science Fund (Grant No. I3163-N36) and by the Swiss National Science Foundation (SNF) through the NCCR-QSIT programme (Grant No. 51NF40-160591).

Data availability statement

The relevant datasets generated and analyzed throughout this work are available from the corresponding author upon reasonable request.

Appendix A. Beyond the Itô convention

In section 2 we have derived the white-noise DFT under the assumption of the Itô prescription. In this section, we seek to extend the generality and applicability of our results to arbitrary integration conventions. Integration conventions are specified by a parameter $\alpha \in [0, 1]$ that uniquely prescribes the discretization of equation (1) [26, 39]. The value of $\alpha$ determines where the functions $f_i$ and $g_{ij}$ are evaluated to compute the evolution of $q_i$ throughout a small time step $\Delta t$, imposing

$$\Delta q_i = f_i (q) \Delta t + \sum_j g_{ij} (q) \Delta W_j.$$  \hfill (A1)

Physical trajectories typically follow the Stratonovich convention $\alpha = 1/2$ rather than the Itô prescription $\alpha = 0$ as a consequence of the Wong–Zakai theorem [40]. Fortunately, it is possible to map Langevin equations from one integration convention to another [25, 26]: a system of Langevin equations as discretized in equation (A1) for arbitrary $\alpha$ is statistically equivalent with the system of (discretized) Itô-SDEs

$$\Delta q_i = \left[ f_i (q) + \alpha \sum_{j,k} g_{jk} (q) \frac{\partial g_{ik}}{\partial q_j} \right] \Delta t + \sum_j g_{ij} (q) \Delta W_j.$$  \hfill (A2)

We point to reference [25] for a detailed derivation of equation (A2), as it lies beyond the scope of this work. The equation demonstrates that the conversion between different $\alpha$-prescriptions simply entails an appropriate modification of the deterministic term, sometimes called the thermodynamic or spurious drift. It is proportional to $\alpha$, or more generally $\alpha$-difference, and vanishes in the case of additive noise due to the
derivatives in its expression. After modifying the drift term $f_i$ associated with convention $\alpha$ according to equation (A2), one can once again apply the white-noise DFT as delineated in the main text.

Alternatively, it is also possible to formulate DFTs for arbitrary $\alpha$ by employing the respective expressions of the transition probability. The corresponding infinitesimal propagators represent a pivotal element of the path integral formulation of Langevin dynamics and have been rigorously derived through various approaches [26, 39]. Although formulating the white-noise DFT for $\alpha \neq 0$ poses no additional conceptual difficulty, the expressions involved lose some simplicity and offer seemingly no noteworthy advantage over the Itô-based treatment.

Appendix B. Simulation methods

The simulations conducted throughout the course of the present study can be broadly split into three steps: trajectory generation, preparation of $Z_i$ statistics, and subsequent extraction of the $\Sigma(Z_i)$ symmetry function. The last step is performed either as consistency check or to extract the slope $R$ for the purpose of parameter optimization. In this section, we will briefly outline the methods we have employed to perform these tasks.

As mentioned in section 2.2, we can either generate an ensemble of trajectory segments or a single long trajectory to compute representative distributions of the marginal displacement $P(\Delta Z_i, w_i)$. In particular, the non-conservative system of section 3 relies on the trajectory ensemble approach, resetting the particle once it leaves a selected perimeter around the origin. Here, all trajectory segments begin at the origin with a velocity drawn from the Maxwell–Boltzmann distribution, prescribing $P(\mathbf{v}_0^A)$. In contrast, the two systems of section 4 follow the single trajectory approach. The initial state can be chosen arbitrarily in this case. All trajectories generated in this work use the symplectic OVRVO-integrator devised by Sivak et al [41].

We proceed to compute an array of weighted displacements $\Delta Z_i$ from the trajectory data, the precise expressions for each test system are given in the respective subsections. At this point subsequent realizations of $\Delta Z_i$ are still correlated by construction; the array is thus randomly shuffled before summing several $\Delta Z_i$ into a single realization of $Z_i$. For the extraction of $\Sigma(Z_i)$ it is numerically convenient to choose a number of addends $n$ that puts the average $\langle Z_i \rangle$ close to unity. The independent selection of addends from a sufficiently large sample implies the relation $\langle Z_i \rangle = n(\Delta Z_i)$, making the selection of $n$ straightforward. If desired, reducing the time step $\Delta t$ will result in an approximately inversely proportional increase of $n$ as determined from this criterion. In the presented applications, we have merely ensured that the average of $Z_i$ resides in an appropriate order of magnitude. Next, we compute the corresponding $Z_i$-histograms. The bins should be constructed symmetrically around the origin $Z_i = 0$. This alignment is necessary for the computation of the logarithmic probability ratio $\Sigma(Z_i)$. Finally, for parameter inference applications a linear curve of the form $RZ_i$ is fitted to $\Sigma(Z_i)$ to extract the slope $R$ as discussed in section 5. We have used a standard least-squares method to this end.

ORCID iDs

M Innerbichler https://orcid.org/0000-0003-3719-3793
A Militaru https://orcid.org/0000-0002-5083-317X
L Novotny https://orcid.org/0000-0002-9970-8345
C Dellago https://orcid.org/0000-0001-9166-6235

References

[1] Seifert U 2019 From stochastic thermodynamics to thermodynamic inference Annu. Rev. Condens. Matter Phys. 10 171–92
[2] Seifert U 2012 Stochastic thermodynamics, fluctuation theorems and molecular machines Rep. Prog. Phys. 75 126001
[3] Harris R J and Schütz G M 2007 Fluctuation theorems for stochastic dynamics J. Stat. Mech. P07020
[4] Crooks G E 1999 Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences Phys. Rev. E 60 2721
[5] Jarzynski C 1997 Nonequilibrium equality for free energy differences Phys. Rev. Lett. 78 2690
[6] Hatano N and Sasa S 2001 Steady-state thermodynamics of Langevin systems Phys. Rev. Lett. 86 3463
[7] Langenvin P 1908 Sur la théorie du mouvement Brownien C. R. Acad. Sci., Paris 146 530–3
[8] van Kampen N G 1981 Stochastic Processes in Physics and Chemistry (Amsterdam: North-Holland)
[9] Risken H 1989 The Fokker–Planck Equation 2nd edn (Berlin: Springer)
[10] Coffey W T, Kalmykov Y P and Waldron J T 2004 The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering 2nd edn (Singapore: World Scientific)
[11] Schimansky-Geier L, Mieth M, Rosé H and Malchow H 1995 Structure formation by active Brownian particles Phys. Lett. A 207 140–6
[12] Romanczuk P, Bär M, Ebeling W, Lindner B and Schimansky-Geier L 2012 Active Brownian particles Eur. Phys. J. Spec. Top. 202 1–162
[13] Fodor É, Nardini C, Cates M E, Tailleur J, Visco P and van Wijland F 2016 How far from equilibrium is active matter? Phys. Rev. Lett. 117 038103
[14] Bereshkovskii A M and Szabo A 2013 Diffusion along the splitting/commitment probability reaction coordinate J. Phys. Chem. B 117 13115–9
[15] Peters B, Bolhuis P G, Mullen R G and Shea J-E 2013 Reaction coordinates, one-dimensional Smoluchowski equations, and a test for dynamical self-consistency J. Chem. Phys. 138 054106
[16] Seifert U 2008 Stochastic thermodynamics: principles and perspectives Eur. Phys. J. B 64 423–31
[17] Jayannavar A M and Sahoo M 2007 Charged particle in a magnetic field: Jarzynski equality Phys. Rev. E 75 032102
[18] Wang C and Feldman D E 2014 Fluctuation theorems without time-reversal symmetry Int. J. Mod. Phys. B 28 1430003
[19] West D K, Olmsted P D and Paci E 2006 Free energy for protein folding from nonequilibrium simulations using the Jarzynski equality J. Chem. Phys. 125 204910
[20] Hummer G and Szabo A 2010 Free energy profiles from single-molecule pulling experiments Proc. Natl Acad. Sci. USA 107 21441–6
[21] Liphardt J, Dumont S, Smith B, Tinoco I and Bustamante C 2002 Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski’s equality Science 296 1832–5
[22] Hayashi K, Ueno H, Inoue K and Noji H 2010 Fluctuation theorem applied to F1-ATPase Phys. Rev. Lett. 104 218103
[23] Camunas-Soler J, Alemany A and Ritort F 2017 Experimental measurement of binding energy, selectivity, and allostery using fluctuation theorems Science 355 412–5
[24] Kubo R 1966 The fluctuation–dissipation theorem Rep. Prog. Phys. 29 255
[25] Campisi M, Talkner P and Hänggi P 2007 The Tasaki–Crooks quantum fluctuation theorem J. Phys. A: Math. Theor. 40 F569
[26] Campisi M, Talkner P and Hänggi P 2009 Fluctuation theorem for arbitrary open quantum systems Phys. Rev. Lett. 102 210401
[27] Cugliandolo L F and Lecomte V 2017 Rules of calculus in the path integral representation of white noise Langevin equations: the Ornager–Machlup approach J. Phys. A: Math. Theor. 50 345001
[28] Groisman B, Fraenkel L and Ritort F 2010 The fluctuation theorem and the information content of probability distributions J. Phys. A: Math. Theor. 43 375002
[29] Sivak D A, Chodera J D and Crooks G E 2014 Time step rescaling recovers continuous-time dynamical properties for discrete-time Langevin integration of nonequilibrium systems J. Phys. Chem. B 118 6466–74