

Inference of a time delay in stochastic systems

Robin A. Kopp^{✉,*}, Sabine H. L. Klapp^{✉,†} and Deepak Gupta^{✉,‡}*Institut für Physik und Astronomie, Technische Universität Berlin, Hardenbergstraße 36, D-10623 Berlin, Germany*

(Received 1 August 2025; accepted 5 December 2025; published 5 February 2026)

Time delay is ubiquitous in many experimental and real-world situations. However, from measurable data it is often unclear whether time delay plays a significant role in the dynamics, and if it does, how long the time lag really is. This would be invaluable knowledge when analyzing and modeling such systems. Hitherto, no universal method is available by which the time delay can be inferred. To address this problem, we propose and demonstrate two different methods to infer time delay in overdamped Langevin systems with delayed feedback. In the first part, we focus on the power spectral density based on the positional data and use a characteristic signature of the time delay to infer the delay time. In limiting cases, we establish a direct relation between the observations made for nonlinear time-delayed feedback forces and analytical results obtained for the linear system. In other situations, despite the absence of this direct relation, the characteristic signature remains and can be exploited by a semiautomatic method to infer the delay time. Furthermore, it may not always be desirable or possible to observe a system for a long time to infer dependencies and parameters. Thus, in the second part, we propose a probing method combined with a neural network to infer the delay time, which requires only short observation time series. These proposed methods for inferring time delays in stochastic systems can be valuable tools for gaining deeper insight into the role of delay across a wide range of applications—from the behavior of individual colloidal particles under feedback control to emergent collective phenomena such as flocking and swarming.

DOI: [10.1103/PhysRevResearch.8.013129](https://doi.org/10.1103/PhysRevResearch.8.013129)

I. INTRODUCTION

Time delay can have a substantial impact on a wide range of experimental and real-world scenarios [1]. The presence of time delay makes the dynamics of a system dependent on the system's history itself. In some cases, it is introduced deliberately, whereas in other situations its presence is unavoidable. A prominent example for deliberate use of delay is Pyragas time-delayed feedback for the stabilization of otherwise unstable orbits [2,3], later applied to stabilization of satellite motion [4]. In experiments involving feedback control, researchers have explicitly considered time delay in design and modeling [5–7]. In colloidal systems exhibiting stochastic dynamics, time-delayed feedback control is employed to modify the transport properties [8,9]. Moreover, delay has been shown to have tremendous effect on thermodynamic observables [10–14], as demonstrated in recent experiments [7,15].

Another area where delay is important and, at the same time, difficult to control are systems with finite-time information propagation. Typical examples include communication lags while steering far objects such as satellites, the reaction

time humans experience when reacting to an external stimulus, or the time lag in experiments between processing image data and applying the desired steering or control impulse [16]. Moreover, time delay may have to be taken into consideration in complex systems, such as in swarms [17], in financial markets [18,19], in robotic systems [20], and in traffic flow [21,22]. Indeed, it has recently been shown in theoretical studies as well as simulations that the introduction of time delay and/or time-delayed feedback can induce intriguing emergent collective behavior. Examples range from polar active matter with delayed interactions [23,24], and Gaussian time-delayed feedback applied to Brownian particles [25–27], to ecological systems [28], and oscillator networks [29,30].

Stochastic systems with time delay belong to the class of non-Markovian systems since the dynamics depends on the systems' state at an earlier time; i.e., it is history dependent. Fundamentally, the time delay may be rationalized as a consequence of hidden degrees of freedom. From a Markovian embedding perspective [31], for a system with discrete time delay, an infinite number of degrees of freedom have been projected out [32,33]. Further, in the context of generalized Langevin equations, discrete time delay systems can be understood as systems with a δ -shaped memory kernel. Notice that this memory kernel, however, is not related to the environmental noise correlation [32].

However, in many experimental and real-world situations, it is not *a priori* clear whether time delay plays a significant role [16,34] or if it is present at all. Therefore, it is of fundamental importance to investigate the presence of time delay based on observable quantities to accurately design and understand physical systems.

*Contact author: r.kopp@tu-berlin.de†Contact author: sabine.klapp@tu-berlin.de‡Contact author: phydeepak.gupta@gmail.com

In this study, we focus on systems with a specific delay time rather than multiple delays or distributed delay [35]. Of course, assuming a single discrete time delay in a real system is a rather strong assumption. Broadly speaking, systems with several time delays, distributed delay, or a more general memory kernel have been studied. However, there are numerous systems where the discrete time delay assumption is well justified. In laser systems [36,37] and quantum-optical systems [38,39], time delay originates from light traveling through the optical setup at finite speed. Further examples include complex networks with finite signal propagation speeds (see Refs. [1,40] and references therein on the topics of neural and laser networks). In other cases, one can expect a “typical” time delay as part of a peaked distribution, e.g., when considering finite processing and reaction times. Therefore, gaining information about the presence of time delay as well as the “typical” or average time delay is of high value.

In recent years, ideas of parameter inference for dynamical systems and equation discovery (for ordinary as well as partial differential equations) from (noisy) data have received considerable interest. Methods range from Bayesian inference [41] and regression-based machine learning [42] to the application of neural network-based techniques [43]. For some types of system, methods for identifying governing equations [44,45] and inferring parameters for deterministic delay differential equations and linear stochastic delay differential equations (SDDEs) [46,47], specifically time delay in the underlying dynamics from time series data [48–50], exist. However, those methods are typically designed for application to deterministic systems or noisy measurement data [49]. For fully stochastic (non)linear systems, on the other hand, to the best of our knowledge, no universal approach for the inference of time delay exists. Motivated by this gap, we here propose and demonstrate, for the first time, two approaches to infer a time delay for Brownian systems.

The first approach is based on the power spectral density (PSD) of stochastic time series, which is often used to analyze experimental data [51–53], including systems with anomalous diffusion [54]. Here, we identify a characteristic signature of time delay, namely, oscillations with a delay-specific period, and demonstrate in a proof of concept how this signature can be exploited to infer the delay time. To do this, we start from Brownian systems subject to linear time-delayed feedback as well as specific delayed nonlinear forces, including respective limiting cases, and then consider highly nonlinear time-delayed feedback. We thereby demonstrate that the characteristic signature carries over from systems subject to linear feedback forces to nonlinear systems and, thus, that the PSD-based inference approach is agnostic; i.e., it does not rely on knowledge of the functional form of the underlying stochastic equations of motion.

The PSD approach, however, has a drawback. A large number (typically $N \gtrsim 100$) of relatively long [several Brownian times τ_B (defined in Sec. II)] trajectories is required to infer the time delay with sufficient precision. In real-world applications, having to record a large number of long trajectories may (depending on the scenario) not be desirable or not even possible.

Given the potential difficulties, we propose a powerful alternative approach employing a neural network. To establish

the concept, we utilize the fact that systems that involve time delay, rather than instantaneously responding to an external stimulus, react with some delay. Therefore, by applying a suitable perturbation and observing the dynamical response, we can obtain information about the underlying delay time of the system. In order to classify the system’s response, based on trajectory time series data, we make use of a convolutional neural network [55,56]. Such a network can be trained to output the most likely time delay for one or even several different functional forms of the underlying delayed and instantaneous forces. Compared to the PSD approach, the main advantage of this approach is that the observation window (i.e., the trajectory length) only has to be on the order of the time delay itself, and relatively few trajectories (typically $N \sim 25$) are sufficient for precise inference. In combination with state-of-the-art computer systems, this brings the inference of time delay in real time within the realm of possibilities.

The remainder of this paper is structured as follows. In Sec. II, we introduce and briefly discuss the model, that is, the stochastic delay differential equations that we use in our analysis and in the proof of concept. We then compute and analyze the power spectral density (defined in Sec. III A) for a system subject to linear time-delayed feedback in Sec. III B. After identifying the signature of time delay in linear time-delayed feedback systems, this signature serves as the basis for our general PSD-based approach to the inference of time delay. Next, after preliminary remarks (Sec. IV A) about the treatment of nonstationary time series, we present numerical results for linear (Sec. IV B) and specific nonlinear feedback forces (Sec. IV C) before demonstrating PSD-based inference for fully nonlinear systems in Sec. IV D. Subsequently, we move on to the perturbation-based approach to infer the time delay, where we first discuss the characteristic response of a system with time delay to an external perturbation. To this end, in Sec. V A we study the response of a deterministic linear system analytically using the method of steps and show that the response persists in stochastic systems and also carries over to nonlinear systems. We then discuss in Sec. V B how, using this response, a convolutional neural network can be employed to classify such perturbed trajectories of time-delayed stochastic systems. Subsequently, in Sec. VI we present results for time series classification with respect to the delay time by a suitably trained neural network, demonstrating the viability and accuracy of the perturbation-based approach. Again, we consider linear feedback forces in Sec. VI A as well as more general nonlinear feedback forces (Secs. VI B and VI C) in our proof of concept. This demonstrates the versatility of the neural network-based classification of time series from delayed stochastic systems. Finally, we present conclusions and an outlook to potential applications and future research directions in Sec. VII. Details about the simulation method, the inference procedures, data analysis, as well as analytical details, supplementary figures, and a brief discussion of colored noise systems with corresponding numerical data are given in Appendixes.

II. MODEL

To set the stage, we consider a one-dimensional system of a Brownian particle subject to time-delayed feedback (dFB).

The dynamics of the Brownian particle is governed by the stochastic delay differential equation

$$\gamma \frac{dx}{dt} = F_{\text{dFB}}[x(t), x(t - \tau)] + \xi(t). \quad (1)$$

The term $\xi(t)$ represents Gaussian white noise, with zero mean and correlation function $\langle \xi(t)\xi(t') \rangle = 2\gamma k_B T \delta(t - t')$, where $k_B T$ (with k_B being the Boltzmann constant and T the temperature) is the thermal energy. (We here focus on Gaussian white noise for simplicity. However, Appendix D discusses the applicability and numerical results based on the methods introduced below for systems with exponentially correlated noise.) The time-delayed feedback force, $F_{\text{dFB}}[x(t), x(t - \tau)]$, depends on the instantaneous position of the particle $x(t)$ and the earlier position $x(t - \tau)$, where τ represents the *delay* time. Instead of an initial condition, as it is required for nondelayed systems, the SDDE is complemented with an initial history function $\Phi(t)$ containing the particle's position for $t \in [-\tau, 0]$. In principle, the SDDE can additionally include the force $F_{\text{inst}}[x(t)]$ that depends *only* on the instantaneous position $x(t)$. In what follows, we take the particle diameter σ as a length scale and the Brownian time $\tau_B \equiv \sigma^2/D$ as a timescale (with the bare diffusion constant $D = k_B T/\gamma$).

A. Linear feedback force

For our analytical considerations, we focus on a linear time-delayed feedback force F_{dFB} of the form

$$\begin{aligned} F_{\text{lin}} &= -\nabla_x V_{\text{quad}}[x(t), x(t - \tau)] \\ &= -[k_a x(t) - k_b x(t - \tau)]. \end{aligned} \quad (2)$$

The corresponding delayed feedback potential, V_{quad} , can most intuitively be understood in the limit $k_a = k_b$, where it reduces to a harmonic potential centered at the earlier position $x(t - \tau)$. When $k_a \neq k_b$, the feedback force can be understood as competing instantaneous and delayed forces, where the dominant force is determined by the ratio of parameters k_a and k_b . Systems subject to such linear feedback forces have been studied extensively both, in deterministic [2,4,57,58] as well as stochastic systems [10,12,13,57,59]. In the context of deterministic dynamical systems, stability analysis is one of the central points of interest (see, e.g., Refs. [58,60]) and it has, for example, been shown in an early application that time-delayed feedback can be used to stabilize otherwise unstable orbits [2,3]. On the other hand, stability is also a central feature of stochastic systems and typically connected to stationary or steady states. For the present class of feedback forces, stability has been assessed for different parameter regimes [57] as well as in terms of extensions of the model [35]. Here, we distinguish two cases in terms of stability. First, we consider the case when $0 < k_b \leq k_a$, which we refer to as attractive feedback. Specifically, attractive feedback is marginally stable for $k_a = k_b$ (attractive harmonic feedback) [11,61], implying a constant mean position but a growing positional variance and, thus, a nonstationary time series in the language of time series analysis. Otherwise, it is stable [35,57], i.e., both positional mean and variance are constant in the long-time limit.

Second, the case when $k_a < 0$ and $k_b < 0$, we call repulsive feedback. Repulsive feedback exhibits a more complex stability behavior. In the marginally stable case, when $k_a = k_b$ (repulsive harmonic feedback), there exists a threshold value $k_{\text{thresh}}(\tau) \equiv \gamma/\tau$ in the parameter space (with respect to $|k_a|$ and $|k_b|$) [62] above which the feedback forces diverge (as expected for a force derived from an inverse parabola). Below this threshold, when $|k_a| = |k_b| < k_{\text{thresh}}(\tau)$, the system exhibits enhanced diffusion [62] or even transient superdiffusive behavior [61,62], depending on parameter combinations and the distance from the threshold. In the present article, we study the repulsive system only for parameter combinations below k_{thresh} . Furthermore, when $|k_a| < |k_b| < k_{\text{thresh}}(\tau)$ and $\tau < \arccos(-k_a/k_b)/\sqrt{k_b^2 - k_a^2}$, the system assumes a stationary state [57]. These conditions are satisfied in our considerations for repulsive feedback.

B. Nonlinear feedback forces

Among the broad class of nonlinear feedback forces, we here discuss two examples that have been studied in other contexts. Importantly, linearization of these forces yields the form given in Eq. (2).

The first force is based on a hyperbolic tangent function (see Ref. [8] for a similar feedback force in the context of feedback-controlled transport and Ref. [63] for a nondelayed system) that depends on the difference between instantaneous and delayed positions,

$$F_{\text{tanh}} = -A \tanh[B(x(t) - x(t - \tau))]. \quad (3)$$

This feedback force can be linearized for small magnitudes of B and/or small τ (in the spirit of a small delay approximation [64,65]), recovering the attractive/repulsive harmonic feedback system discussed above (Sec. II A). In its nonlinear form, it has the advantage of only producing finite feedback forces as opposed to the linear feedback force [Eq. (2)] introduced in Sec. II A.

Second, Gaussian time-delayed feedback has been shown to cause peculiar behavior and unusual features when applied to colloidal particles [7], most significantly persistent motion above a threshold in parameter space [7,62]. Here, we consider the force

$$\begin{aligned} F_{\text{Gauss}} &= -\nabla_x V_{\text{Gauss}}[x(t), x(t - \tau)] \\ &= \frac{a}{b^2} [x(t) - x(t - \tau)] e^{-\frac{[x(t) - x(t - \tau)]^2}{2b^2}}, \end{aligned} \quad (4)$$

where for $a > 0$, V_{Gauss} is a repulsive Gaussian delayed feedback potential as studied, e.g., in Refs. [7,25,62].

To understand the impact of F_{Gauss} , we first consider Eq. (1) with the feedback force [Eq. (4)] in the absence of a heat bath (i.e., temperature $T = 0$). Above a threshold in the parameter space, i.e., $a\tau/(\gamma b^2) > 1$ [62], the described particle moves at a constant velocity v_∞ in the long-time limit [62] given a nonzero initial history function.

In the stochastic system, above this threshold, persistent motion is observed at intermediate times, while the feedback leads to enhanced diffusion in the long-time limit. Further, below the threshold to persistent motion, i.e., $a\tau/(\gamma b^2) < 1$, a Brownian particle subject to repulsive Gaussian time-delayed feedback exhibits enhanced diffusion [62].

Linearizations of the repulsive Gaussian feedback force [Eq. (4)] [leading to variants of the two marginally stable linear feedback force systems discussed above; Sec. II A, Eq. (2) with $k_a = k_b$] are possible for two cases (when b is fixed). First, for small displacements [and $a\tau/(\gamma b^2) < 1$], we recover a repulsive harmonic feedback force [62] (see Appendix A 1 a). Second, for the case where the particle exhibits persistent motion [62] (see Appendix A 1 b), the linearization can be understood as a comoving attractive harmonic feedback trap.

III. POWER SPECTRAL DENSITY APPROACH—METHODOLOGY

In this paper, our main objective is to infer the time delay from noisy data. In the present section, we introduce and discuss the power spectral density approach.

A. Definitions

The power spectral density [66,67] is a well-established quantity for the analysis of noisy (measurement noise) or fully stochastic (e.g., Brownian) systems. It is frequently used in experimental and theoretical studies [51–54,68–72] as well as a wide range of (engineering) applications (e.g., signal analysis [67,73]). Its main advantages are accessibility (i.e., it can be directly applied to measurement data and computed using fast algorithms [74] such as the fast Fourier transform [75]) and robustness with respect to noise.

In the context of linear time-delayed feedback, the PSD has previously been defined and computed based on correlation functions [64] and in the context of response theory [76,77] (see also Appendix A 2). We here focus on a different definition (often used in signal analysis [67]). It reads

$$S(\omega) \equiv \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} |\tilde{Z}(\omega)|^2, \quad (5)$$

where $\mathcal{T}$ is the trajectory length, $\tilde{Z}$ is the Fourier transform of a stationary time series $Z(t)$ (i.e., with mean and variance approaching a constant value), and ω is the frequency. Here, we used the following definition for the Fourier transform:

$$\tilde{f}(\omega) = \int_{-\infty}^{+\infty} dt e^{-i\omega t} f(t). \quad (6)$$

For a stochastic time series $Z(t)$, the power spectral density (5) is a fluctuating quantity. To ensure that the data are suitably smooth for the purpose of inference of the delay time, we consider the noise-averaged power spectral density (from this point we use the acronym PSD as a shorthand notation for this quantity). It is defined as

$$\text{PSD} \equiv \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \langle |\tilde{Z}(\omega)|^2 \rangle, \quad (7)$$

where $\langle \dots \rangle$ indicates an ensemble average over noise realizations. To compute the PSD, we evaluate the ensemble average before taking the limit $\mathcal{T} \rightarrow \infty$. Swapping the average and the limit is possible as long as each trajectory is sufficiently long.

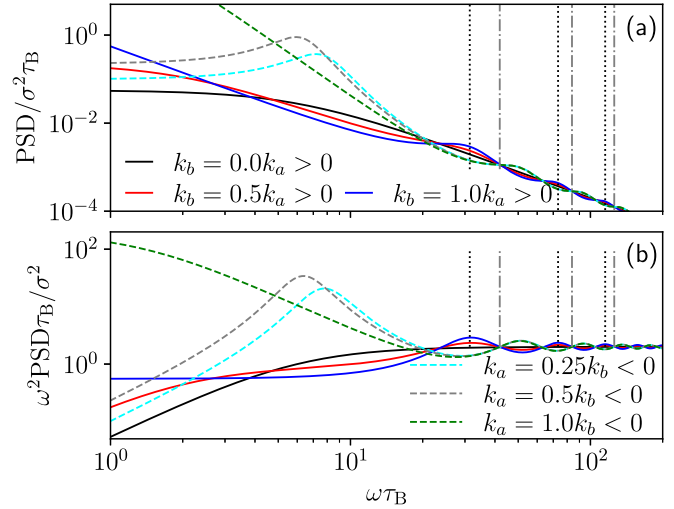


FIG. 1. Power spectral density as a function of the frequency ω for a Brownian particle subject to linear time-delayed feedback [Eq. (2)] at $\tau = 0.15\tau_B$. Solid/dashed lines: Analytical result [Eq. (10)] [panel (a)] and the PSD multiplied by ω^2 [panel (b)], for different values of $k_{a,b}$. Black solid curve represents a harmonically confined Brownian particle without feedback [i.e., $k_b = 0$ in Eq. (10)]. Dash-dotted gray vertical lines: Approximate inflection points at $\omega = 2\pi n/\tau$ with $n \in \mathbb{N}$. Dotted black vertical lines: Approximate maximum for $k_{a,b} > 0$ at $\omega = 3\pi/(2\tau) + 2\pi n/\tau$ for $n > 1$, where $n \in \mathbb{N}$.

B. Linear systems

Explicit results for the PSD can be obtained for the linear time-delayed feedback [Eq. (2)] where k_a and k_b are chosen so that the system is stable or marginally stable. For the repulsive case ($k_{a,b} < 0$), we choose the feedback parameters such that we stay below the threshold to instability, i.e., $|k_a| \leq |k_b| < k_{\text{thresh}}(\tau) \equiv \gamma/\tau$. Above the threshold, the system is unstable (resulting in a nonstationary time series due to diverging forces), and, thus, the PSD is not defined.

We are interested in the PSD (7) with $\tilde{Z} = \tilde{x}$. The Fourier transform of the linear dynamics [Eq. (1) with Eq. (2)] reads

$$\gamma i\omega \tilde{x}(\omega) = -(k_a - k_b e^{-i\omega\tau}) \tilde{x}(\omega) + \tilde{\xi}(\omega). \quad (8)$$

Solving for $\tilde{x}(\omega)$, we get

$$\tilde{x}(\omega) = \frac{\tilde{\xi}(\omega)}{\gamma i\omega + k_a - k_b e^{-i\omega\tau}}. \quad (9)$$

Substituting Eq. (9) in Eq. (7), we obtain

$$\langle S(\omega) \rangle_{\text{lin}} = \frac{2\gamma k_B T}{\gamma^2 \omega^2 - 2\gamma k_b \omega \sin(\omega\tau) - 2k_a k_b \cos(\omega\tau) + k_a^2 + k_b^2}. \quad (10)$$

Figure 1 shows analytical results [Eq. (10)] for the PSD of the linear system for different k_a, k_b and fixed τ .

We first consider the limit $k_b = 0$ in Eq. (10), i.e., no feedback. In this case, and if $k_a > 0$, the expression takes the shape of the PSD of a Brownian particle in a harmonic trap of stiffness k_a . It follows that $\langle S(\omega) \rangle_{\text{lin}} = \frac{2\gamma k_B T}{\gamma^2 \omega^2 + k_a^2}$, which decreases as $1/\omega^2$ in the large-frequency limit (see the nonoscillatory black solid curve in Fig. 1). (We note here that for $k_a < 0$,

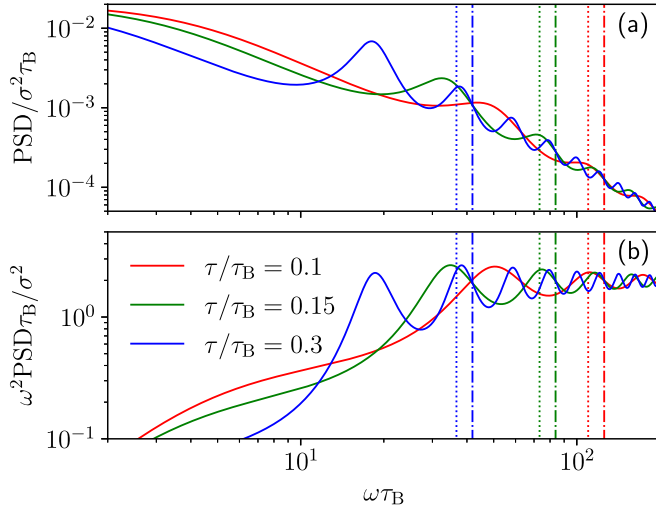


FIG. 2. Power spectral density as a function of the frequency ω for a Brownian particle subject to linear time-delayed feedback [Eq. (2)] for $k_b = 0.5k_a > 0$. Solid lines: Analytical results [Eq. (10)] [panel (a)] and the PSD multiplied by ω^2 [panel (b)] for different values of the delay time τ . Dash-dotted vertical lines: Approximate inflection points at $\omega = 2\pi n/\tau$ with $n \in \mathbb{N}$. Dotted vertical lines: Approximate maximum at $\omega = 3\pi/(2\tau) + 2\pi n/\tau$ for $n > 1$, where $n \in \mathbb{N}$.

the particle is subject to an instantaneous repulsive harmonic potential, which yields instability.)

When introducing time delay, i.e., $\tau > 0$, the τ -dependent oscillatory terms (sin and cos) in the denominator of Eq. (10) come into play. As a consequence, the PSD and, even more clearly, the function $\omega^2\text{PSD}$ [that is, Eq. (10) multiplied by ω^2] show oscillatory behavior, in contrast to the harmonically trapped particle (see Fig. 1). Oscillations induced by time delay are not unexpected. In deterministic systems, this can be an oscillating position, while for stochastic systems one would expect oscillations in the mean square displacement [61] or other correlation functions [12,76,77]. Furthermore, for similar linear feedback force systems, we know that at least in stationary states, oscillations can also occur in the PSD (see, e.g., Ref. [64]). Here, we analyze the oscillations in the linear system as a reference for our subsequent investigation of the nonlinear case.

From Fig. 1, one can observe that the minima and maxima for $k_{a,b} > 0$ and $k_{a,b} < 0$ are offset by half an oscillation period, where the offset is a consequence of the sign change in the prefactors of the trigonometric functions [Eq. (10)].

Importantly, the specific choice of the parameters $k_{a,b}$, only influences the small frequency behavior (see Appendix A3 for an analytical consideration of this limit and Appendix A4 for the special case of a harmonic feedback system) but does not affect the oscillatory part of the PSD on which we concentrate from now on.

To illustrate the τ dependence of the oscillations, we present the PSD and the function $\omega^2\text{PSD}$ at different values of the delay time (color coded) in Fig. 2 for a system with $k_a > k_b > 0$. While the term $2k_a k_b \cos(\omega\tau)$ in Eq. (10) has a maximum magnitude of $|2k_a k_b|$, the magnitude of $2\gamma k_b \sin(\omega\tau)$

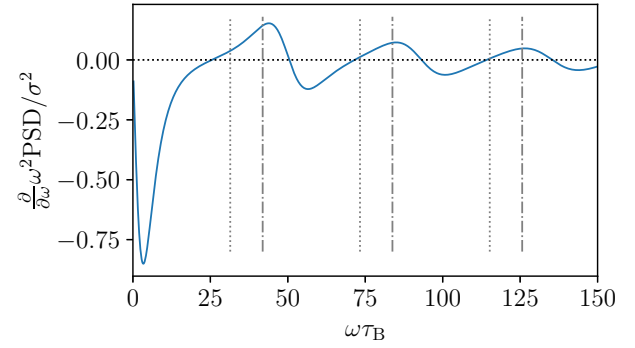


FIG. 3. The derivative of PSD [Eq. (10)] (i.e., $\partial/\partial\omega[\omega^2\text{PSD}(\omega)]$) as a function of ω . Here, $k_{a,b} > 0$ and $k_a \neq k_b$. Dotted horizontal line: Guide for the eye to indicate the zero crossings of the blue solid curve. Dash-dotted vertical line: Approximate inflection points of $\omega^2\text{PSD}$ at $\omega = 2\pi n/\tau$, for $n \in \mathbb{N}$. Dotted vertical line: Approximate maxima of $\omega^2\text{PSD}$ at $\omega = 3\pi/(2\tau) + 2\pi n/\tau$.

grows with ω . Therefore, as ω grows, specifically when $\omega > k_a/\gamma$, the term $2\gamma k_b \omega \sin(\omega\tau)$ [in Eq. (10)] becomes dominant. Indeed, we observe from Fig. 2 that the value of the delay time determines both, the location of the maxima and minima and the oscillation period. The first minimum/maximum after the sine term becomes dominant is the first relevant maximum/minimum for attractive/repulsive feedback. It is located at approximately $\omega = 3\pi/(2\tau)$. For $\omega \gg k_a/\gamma$, the oscillation period [due to the periodicity of the sin and cos terms in Eq. (10)] is given by $\Delta\omega = 2\pi/\tau$ in frequency space (dotted and dash-dotted vertical lines of the same color indicating the approximate location of the maxima and inflection points, respectively). We refer to this behavior, i.e., oscillations with period $\Delta\omega = 2\pi/\tau$ occurring at large ω , as the *signature* of time delay.

For longer delay times, we observe from Fig. 2 a shift of the delay-time-dependent positions of the minima and maxima to smaller frequencies and a shorter oscillation period, consistent with the prediction. Thus, the specific oscillatory behavior and delay-time-dependent oscillation period $\Delta\omega = 2\pi/\tau$ can, in principle, be used to infer time delay without exact knowledge about the other parameters in the system or fitting simulation data with the analytical result.

To extract the positions of the maxima and minima in the frequency space, it is instructive to consider the derivative of $\omega^2\text{PSD}$ with respect to ω . This is depicted in Fig. 3 for the case $k_{a,b} > 0$. The zero crossings correspond to the positions of the minima and maxima. For $k_{a,b} > 0$, the maxima of $\omega^2\text{PSD}$, which are roots of $\partial_\omega[\omega^2\text{PSD}]$, are located at $\omega \approx 3\pi/(2\tau) + 2\pi n/\tau$ (dotted vertical lines in Fig. 3) while the minima are offset by π/τ . The inflection points of $\omega^2\text{PSD}$ (maxima in Fig. 3) are located at $\omega \approx 2\pi n/\tau$, $n \in \mathbb{N}^+$ (indicated by dash-dotted vertical lines). We observe from Fig. 3, specifically from the second set of vertical lines at $\omega\tau_B \sim 75$, that the approximation is in good agreement with the actual positions of the roots and maxima of $\partial_\omega[\omega^2\text{PSD}]$, while for the first zero crossing and maximum there is still a significant deviation. At those frequencies, the term containing $\cos(\omega\tau)$ is not negligible.

IV. POWER SPECTRAL DENSITY APPROACH—NUMERICAL RESULTS

We now discuss numerical results (based on trajectories from Brownian dynamics (BD) simulations, see Appendix B 1 for details) for the PSD for various types of SDDEs. We first benchmark the numerics by considering strictly linear systems. We then proceed by considering the specific nonlinear systems introduced in Sec. II B, where linear approximations recover harmonic attractive/repulsive feedback systems. This allows us to obtain first insights into the impact of nonlinearity. Finally, we move on to more general nonlinear systems to see how the oscillations can be used to infer the delay time without explicit knowledge of the functional form of the underlying forces.

A. Preliminary remarks

For the computation of the PSD, it is crucial that the underlying time series is stationary (see Appendix B 2 for detailed discussion on this topic), which means that mean and variance of $Z(t)$ in Eq. (7) have to approach constant values, each as a function of time. Already a simple passive Brownian particle does not fulfill this requirement, since its positional variance grows in time. Similarly, the marginally stable linear feedback system effectively exhibits diffusive behavior in the long-time limit. To overcome this problem, we apply the differencing method (see Appendix B 2), i.e., we compute the PSD of the positional difference time series $\{\bar{x}(t)\} \equiv \{x(t + dt) - x(t)\}$ for small dt , instead of the position itself. As a rough analogy, this method can be understood as “taking the discretized derivative” of the positional data with respect to time. After applying the Fourier transform and dividing by dt as well as by the frequency (recall that derivatives become algebraic factors under the Fourier transform), this leaves us with the correct Fourier transform of the position, which can then be plugged into Eq. (7). Not applying this technique would lead to a significant deviation between numerical results and analytical prediction, which becomes particularly severe for the (nonstationary) repulsive feedback case, as shown in Appendix B 2, Fig. 12. For consistency, we apply the same procedure to all of our data, even for systems that are stationary in the long-time limit.

B. Linear systems

Figure 4 presents numerical results based on Eq. (1) with a feedback force as in Eq. (2) in the absence of additional instantaneous forces. As before, we distinguish attractive feedback [Fig. 4(a)] and repulsive feedback [Fig. 4(b)]. In the regime $\omega\tau_B \gtrsim 1$, we find almost perfect agreement with the analytical prediction for all cases considered. Only at small frequencies and $k_a \neq k_b$, we observe a deviation from the analytical predictions. The deviation can be shifted to smaller frequencies by using longer trajectories. We therefore understand this behavior as an effect connected to the finite length of the trajectory used to compute the PSD.

C. Nonlinear feedback systems and their linear approximations

Now we turn toward the nonlinear systems based on the hyperbolic tangent [Eq. (3)] and Gaussian feedback forces

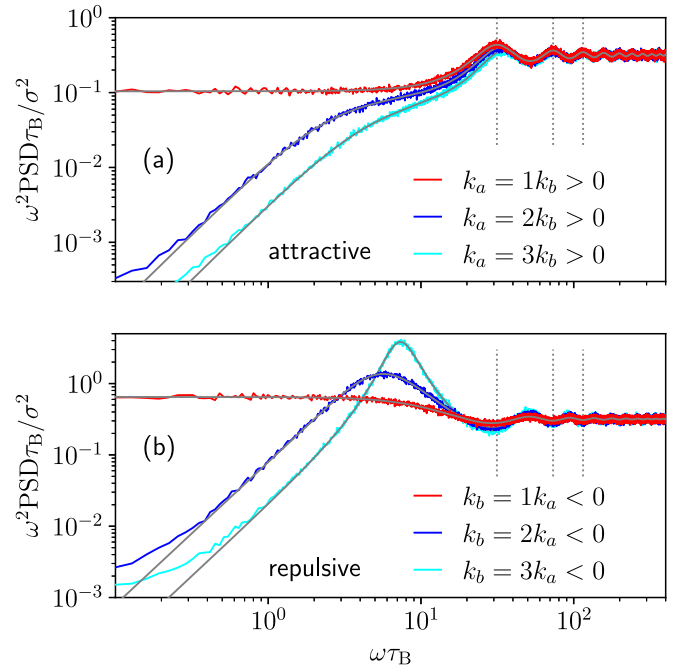


FIG. 4. Comparison of analytical power spectral density (solid gray curves) [Eq. (10)] with numerically simulated noise-averaged PSD as a function of frequency ω for the linear system (2) at $\tau = 0.15\tau_B$. (a) $k_{a,b} > 0$. (b) $k_{a,b} < 0$, where we take $|k_b| < k_{\text{thres}}(\tau)$ to ensure finite feedback forces. Vertical dotted gray lines: Approximate analytical locations of maxima/minima for attractive/repulsive feedback. Numerical results are obtained by noise averaging over 400 noise realizations.

[Eq. (4)] that reduce to attractive/repulsive harmonic feedback systems (discussed in Sec. III B) in certain limits. Our goal here is to investigate whether the oscillatory behavior persists in nonlinear systems and under what conditions deviations from linear behavior are observed in $\omega^2 \text{PSD}$. Numerical data for $\omega^2 \text{PSD}$ are presented in Fig. 5. Solid lines represent analytical results obtained for the different limits discussed in Sec. II B and Appendix A 1 b.

First, we consider the hyperbolic tangent feedback force [Eq. (3)]. Figures 5(a) and 5(b) present numerical data of the PSD and the corresponding analytical approximations based on the attractive harmonic feedback system [Eq. (2) with $k_a = k_b > 0$], for different τ and A , respectively. For both cases, we observe good agreement between numerical data and the analytical prediction for larger ω (i.e., $\omega\tau_B \gtrsim 10$ in Fig. 5), especially with respect to the oscillations. For small ω , we observe noticeable deviations from the analytical linear prediction. These deviations grow as τ is increased or A is decreased. This is because we are approaching the limit where the linear approximation becomes invalid. Similar to linear systems (Fig. 2), we find for this nonlinear system that the delay time determines the oscillation period, as well as the location of the minima and maxima.

Moving on to the system subject to Gaussian time-delayed feedback [Eq. (4)], we have to distinguish two limits based on the deterministic threshold, i.e., $a\tau/(\gamma b^2) = 1$. Far below this threshold (in particular $a = 2k_B T$ and $a = 3k_B T$), in the limit of small displacements the feedback force can be

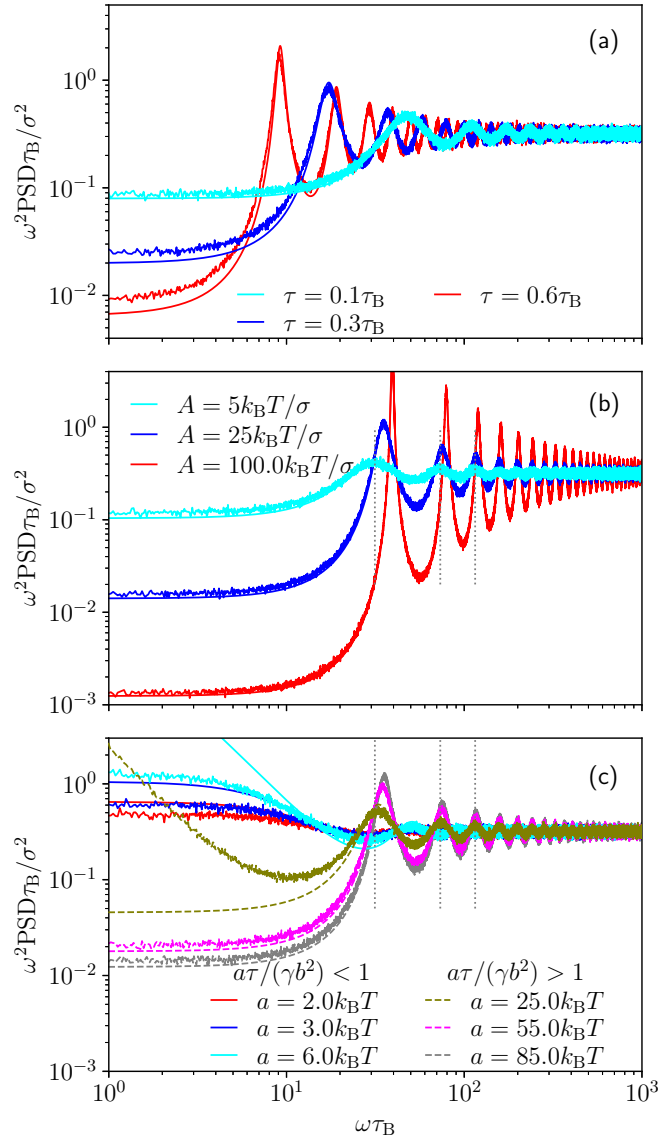


FIG. 5. Comparison of numerically simulated (colored points, obtained from averaging over 400 noise realizations) power spectral density of nonlinear systems with the linear approximation (solid/dashed curves). (a) tanh feedback force for various τ ; solid curves: linear approximation with $k_a = k_b = 10k_B T/\sigma$. (b) tanh feedback force [Eq. (3)] for different A ; solid curves: linear approximation with corresponding $k_a = k_b > 0$. (c) Gaussian feedback force [Eq. (4)]; solid (dashed) curves: analytical result for the linear approximation with repulsive harmonic feedback ($k_a = k_b < 0$) and modified attractive harmonic feedback for the constant velocity state [see Appendix A 1 b, Eq. (A1) for more details]. (a) $B = 1/\sigma$. (b), (c) $\tau = 0.15\tau_B$. Vertical gray dotted lines: Approximate locations of maxima/minima according to the analytical result for the linear system in panels (b) and (c).

approximated as a repulsive harmonic feedback [62]. On the other hand, for feedback parameters far beyond the threshold to persistent motion ($A = 55k_B T$ and $A = 85k_B T$), the particle can be considered as moving at an approximately constant velocity while being subject to attractive harmonic feedback [62] (see Appendix A 1 b). From Fig. 5(c), far above and below the threshold, we see agreement between the data obtained

by numerical simulations of the full nonlinear problem and the expression Eq. (10) and the result based on Eq. (A4), respectively. Particularly, we observe oscillations with the same characteristics as in the linear system. Only at low frequencies, i.e., before the onset of oscillatory behavior, do the simulation data deviate from the analytical approximation.

Interestingly, as we approach the threshold, $a\tau/(\gamma b^2) = 1$, from above or below ($a = 6k_B T$ and $a = 25k_B T$) we see a significant deviation from the analytical prediction at small ω , indicating that the simple linear approximation is no longer valid but a crossover behavior becomes relevant. However, the locations of the maxima and minima remain unaffected, from which we interpret that the signature of time delay persists even outside the linear regime.

From the results discussed so far, we conclude that the characteristic oscillations with period $\Delta\omega = 2\pi/\tau$ are robust against variation of the feedback force, even in the nonlinear regime.

D. General nonlinear time-delayed feedback

We now turn our attention to general nonlinear feedback forces and nonlinear instantaneous forces.

While for the previously discussed systems, in principle, a fit with the exact analytical expression is a way to extract the delay time, for a more general functional form there may not be a direct correspondence.

However, we can make use of the observation for the linear system that, at large ω , the minima and maxima are $\Delta\omega \approx 2\pi/\tau$ apart. We hypothesize that oscillations with this oscillation period can also be found in the function $\omega^2\text{PSD}$ of nonlinear systems; i.e., the signature of time delay carries over from linear to more general nonlinear systems. On this basis, we can design a numerical procedure based on a polynomial fit in order to extract the positions of the minima and maxima in frequency space from numerical data and subsequently compute an approximate inferred delay time (for details of the procedure, see Appendix B 3).

In the following, we present numerical results where we apply this procedure to three combinations of nonlinear feedback forces and nonlinear instantaneous forces.

The corresponding SDDEs read

$$\gamma \dot{x}(t) = -ax^3(t) + bx(t) - cx(t - \tau) + \xi(t), \quad (11)$$

$$\gamma \dot{x}(t) = -ax^3(t) + bx(t) - cx^2(t - \tau) + \xi(t), \quad (12)$$

$$\gamma \dot{x}(t) = -a \cos[x(t)] - b + c \cos[x(t - \tau)] + \xi(t). \quad (13)$$

Figure 6 presents exemplary fits of the resulting function $\omega^2\text{PSD}$ including the obtained positions of the minima and maxima, which were used to infer the delay time. The extracted locations of the minima and maxima are in good agreement with the data as shown in Fig. 6, where the extracted locations are marked as circles.

Table I summarizes the inference results, where we compare true τ and inferred delay times τ_{inferred} for simulation data based on Eqs. (11), (12), and (13) for three different delay times.

Overall, we find robust and accurate inference results for various delay times with an uncertainty on the order of 1%

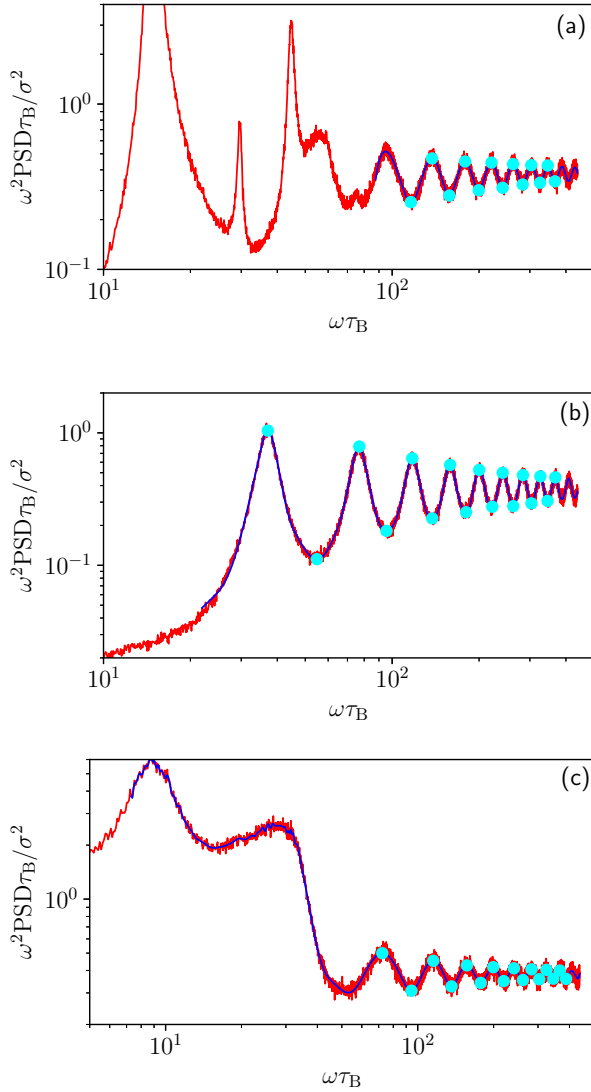


FIG. 6. Numerically simulated PSD (red) as a function of ω for nonlinear systems (11) [panel (a)], (12) [panel (b)], and (13) [panel (c)]. In each panel: Polynomial fit (blue) using the quartic splines and automatically extracted positions (cyan circles) of maxima and minima using the roots of the derivative of the polynomial (see Appendix B 3). See Table I for the corresponding inference results.

or less. The accuracy of the inference is similar for the different linear and nonlinear feedback and instantaneous forces.

TABLE I. Delay times τ_{inferred} in units of τ_B inferred from PSD with 1 standard deviation as uncertainty: (a) Data corresponding to Eq. (11) with $a = 10k_B T/\sigma^4$ and $b = c = 10k_B T/\sigma^2$; (b) data corresponding to Eq. (12) with $a = 10k_B T/\sigma^4$, $b = 10k_B T/\sigma^2$, and $c = 20k_B T/\sigma^3$; and (c) data corresponding to Eq. (13) with $a = 10k_B T/\sigma$, $b = 5k_B T/\sigma$, and $c = 20k_B T/\sigma$.

(a)	Equation (11)	(b)	Equation (12)	(c)	Equation (13)
τ	τ_{inferred}	τ	τ_{inferred}	τ	τ_{inferred}
0.1	0.100 ± 0.002	0.1	0.101 ± 0.001	0.1	0.100 ± 0.005
0.15	0.151 ± 0.001	0.15	0.152 ± 0.003	0.15	0.149 ± 0.003
0.3	0.303 ± 0.014	0.3	0.301 ± 0.009	0.3	0.300 ± 0.007

Crucially, for the described procedure, knowing the functional form of the forces (instantaneous and delayed) is not required. Therefore, this procedure can be applied to a wide range of systems with discrete time delay. Furthermore, the procedure is not limited to Gaussian white noise systems only. In Appendix D, we present numerical results for systems with exponentially correlated noise (see Fig. 16 and Table III).

V. PERTURBATION AND DEEP LEARNING-BASED APPROACH—METHODOLOGY

Up to this point, we have demonstrated that the power spectral density features a characteristic signature of time delay for a broad class of feedback forces (linear and nonlinear) in stochastic systems. This allowed us to infer time delay even if the exact functional form of the underlying forces is unknown. As a drawback, this methodology requires a large number of long trajectories. This can be a challenge in experiments, where one may only have fewer and shorter trajectories at ones disposal.

To circumvent this issue, we here propose an alternative procedure to infer time delay. This procedure is based on the observation of the response of the system under the influence of a perturbative force. Broadly speaking, for a delay-driven system subjected to a perturbation, we expect the system's response to occur not instantaneously but rather with some delay, as we shall see explicitly in Sec. V A. Assuming the perturbation to be a force that is applied instantaneously and that is sufficiently strong (i.e., strong compared to noise, feedback forces, and instantaneous forces), we expect a distinct delayed reaction, a *characteristic response*, as we explain in detail in the following.

A. Characteristic response of a system with time delay to a perturbation

To develop an analytical understanding of the characteristics of the response, we take one step back and consider a deterministic system (i.e., the limit $T \rightarrow 0$) with linear time-delayed feedback force and an additional perturbative force $F_{\text{perturb}}(t)$. Its equation of motion reads

$$\gamma \dot{x} = -[k_a x(t) - k_b x(t - \tau)] + F_{\text{perturb}}(t). \quad (14)$$

Intuitively, in the limit $k_a = k_b > 0$ (attractive harmonic), the linear feedback force can be understood as the consequence of a delayed harmonic potential centered at $x(t - \tau)$. Figure 7 shows its schematic representation. We assume that the particle was at rest at $x(t) = 0$ in the time interval $t \in [-\tau, 0]$. We then apply a perturbative force, F_{perturb} , starting at $t = 0$ for a duration $T_{\text{perturb}} > \tau$. The particle will first be confined to a harmonic potential centered at $x(t - \tau) = 0$. Following dynamics (14) after $t = \tau$, the delayed position $x(t - \tau)$ (and hence the location of the harmonic trap) will also start to change, resulting in a change in dynamical behavior. To make analytical progress, we now assume a perturbation protocol given by a step function, i.e., $F_{\text{perturb}}(t) = F \Theta(t)$, where $\Theta(t)$ is the Heaviside function. We calculate the dynamics resulting from Eq. (14) using the so-called *method of steps* [78]. This method can be used to obtain an analytical solution in intervals of one delay time iteratively,

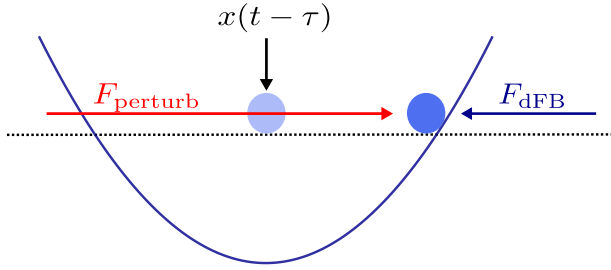


FIG. 7. Schematic depiction: A particle (dark blue circle), subject to attractive harmonic time-delayed feedback with feedback force [Eq. (2)] for $k_a = k_b > 0$ in one dimension. The feedback potential is centered at the delayed position $x(t - \tau)$ (i.e., at the faint blue circle). A perturbing force F_{perturb} is applied [Eq. (14)].

provided that the history function $\Phi(t)$ containing the position of the particles in the previous time interval is known, i.e., $\Phi(t) = x(t)$ for $t \in [t_0 - \tau, t_0]$. Formally, the piecewise

$$x(t) = \begin{cases} 0, & -\tau < t < 0, \\ \frac{F}{k_a}(1 - e^{-\frac{k_a t}{\gamma}}), & 0 \leq t < \tau, \\ \frac{F}{k_a}[2 - e^{-\frac{k_a t}{\gamma}} - e^{-\frac{k_a}{\gamma}(t-\tau)} - \frac{k_b}{\gamma}e^{\frac{k_a \tau}{\gamma}}(t - \tau)], & \tau \leq t < 2\tau. \end{cases} \quad (16)$$

For reasons that will become clear later, we also take the time derivative of this piecewise solution (16), yielding

$$\dot{x}(t) = \begin{cases} 0, & -\tau < t < 0, \\ \frac{F}{\gamma}e^{-\frac{k_a t}{\gamma}}, & 0 \leq t < \tau, \\ \frac{F}{k_a}[\frac{k_a}{\gamma}e^{-\frac{k_a t}{\gamma}} + \frac{k_a}{\gamma}e^{-\frac{k_a}{\gamma}(t-\tau)} - \frac{k_b}{\gamma}e^{\frac{k_a \tau}{\gamma}}], & \tau \leq t < 2\tau. \end{cases} \quad (17)$$

One first observation from Eqs. (16) and (17) is that the piecewise solution (16) and also its derivative (17) are proportional to the perturbative force F . Thus, in the absence of a perturbative force, the particle stays at rest at the initial position $x(0) = 0$, as expected in the deterministic case.

Second, Eqs. (16) and (17) directly reveal the role of the delay time τ . Indeed, in the first interval $[0, \tau]$, the system behaves as if there was no delay at all. This changes when we consider the regime $t \geq \tau$: While the first term encodes the system's instantaneous reaction to the applied perturbation, the second and third terms can be understood as the time-delayed response. Those terms can lead to a significant change in dynamical behavior, where for each specific τ the first change in dynamical behavior occurs immediately at $t = \tau$. This means (given a zero history) that from this point on memory effects become important.

Figure 8 presents corresponding exemplary numerical data including the analytical results from Eqs. (16) and (17). We perturb the system by applying a constant force to the particle starting at $t = 0$ for a duration T_{perturb} , which has to be, at least, longer than one delay time (i.e., $\tau < T_{\text{perturb}}$) and observe the system. The observation continues after the perturbation is set to zero at $t = T_{\text{perturb}}$ for a further duration of T_{perturb} [Fig. 8(c)]. We see from Fig. 8 that after applying the perturbation, but also after the perturbation is set to zero, there

solution as a function of time during the interval $[t_0, t_0 + \tau]$ reads

$$x_{[t_0, t_0 + \tau]}(t) = \Phi(t_0)e^{-\frac{k_a}{\gamma}(t-t_0)} + \frac{k_b}{\gamma} \int_{t_0}^t dt' \Phi(t' - \tau)e^{-\frac{k_a}{\gamma}(t-t')} + \frac{F}{\gamma} \int_{t_0}^t dt' \Theta(t')e^{-\frac{k_a}{\gamma}(t-t')}. \quad (15)$$

Similarly, the particle's position $x(t)$ can be evaluated for $t \geq t_0 + \tau$ iteratively. However, for our analytical considerations, we focus on the dynamical behavior during the first two delay times τ after switching on the perturbation at $t = 0$.

We assume that the initial history function represents a particle at rest in position $x = 0$, i.e., we choose $\Phi(t) = x(t) = 0$ for $t \in [-\tau, 0]$ (which can easily be realized by pinning the particle). The analytical expressions for the position as a function of time during the time interval $[-\tau, 2\tau]$ following Eq. (15) then reads

is a characteristic change in dynamical behavior after one τ [indicated by color-coded vertical dotted lines in Figs. 8(a) and 8(b)]. This change consists in “kinks” and other deviations from a system without delay. We also find that the change is most distinct in $\dot{x}$ [Fig. 8(b)], even though this is not immediately obvious from the analytical expression (17). Importantly, such a characteristic response is not limited to linear systems but carries over to more complex nonlinear instantaneous and delayed forces; see Fig. 14 (Appendix) for exemplary data.

But how can this knowledge be exploited in a noisy Brownian system with time delay? Does the characteristic response described before carry over to noisy systems? To this end, in Fig. 8(d) we show how a linear stochastic system based on feedback forces as in Eq. (2) reacts to a perturbation by a constant external force [Fig. 8(e)] following a step protocol, where we consider the discretized derivative of positional data. We observe that also in the noisy data the characteristic response of a system with time delay (black lines represent the deterministic response) becomes apparent [Fig. 8(d)].

In summary, inspecting Figs. 8 and 14, we notice that all observed systems herein show a characteristic response whose details are unique to the specific functional form of the forces involved. In particular, a characteristic response is also present and visible in systems with nonlinear delayed and instantaneous forces. We emphasize that it is important to have a

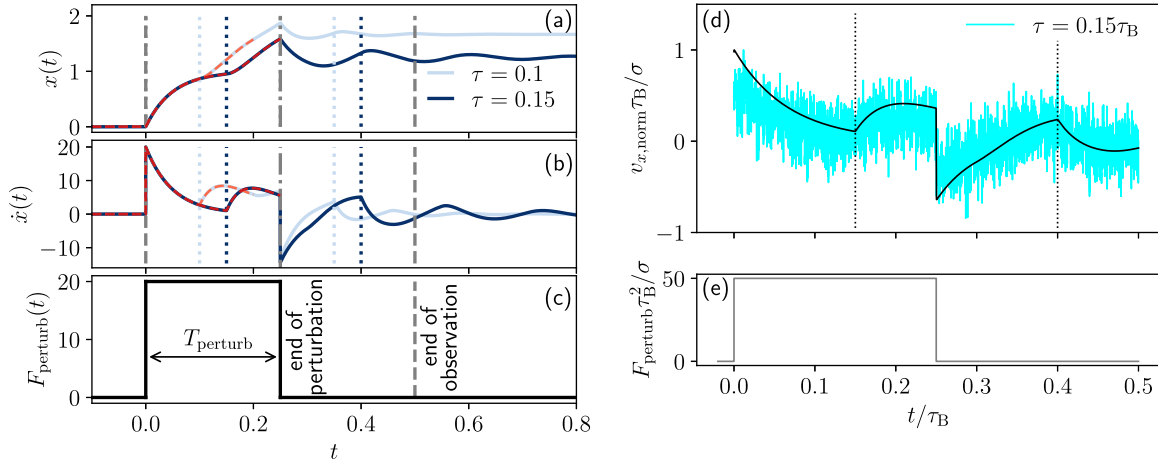


FIG. 8. Effect of a perturbation on deterministic [panels (a)–(c)] and stochastic systems [panels (d) and (e)]. (a)–(c) The particle subject to attractive harmonic time-delayed feedback and external perturbation [Eq. (14)]. Red dashed lines: Analytical solutions for position [Eq. (16)] [panel (a)] and velocity [Eq. (17)] [panel (b)] up to $t = 2\tau$ for $\tau = 0.1$ and up to the end of the perturbation for $\tau = 0.15$. Dash-dotted vertical line indicates end of perturbation; dashed gray vertical lines represent the beginning and end of observation; and vertical dotted lines (color coded) serve as a guide for the eye at one delay time τ after the onset, and then, after the end of the perturbation. Panel (d) shows the discretized “velocity” $v_{x,\text{norm}}$ after noise averaging over 25 realizations, applying a simple moving average over three time steps and normalization with the maximum velocity v_{max} as a function of time during the observation window; the deterministic velocity [black, from panel (b)] is shown as a reference. Dotted vertical lines serve as a guide for the eye at $t = \tau$ and $t = T_{\text{perturb}} + \tau$. (c), (e) The perturbing force F_{perturb} as a function of time. The system is perturbed for a duration of $T_{\text{perturb}} = 0.25 > \tau$ [panels (a)–(c)] and $T_{\text{perturb}} = 0.25\tau_B > \tau$ [panels (d) and (e)], and the observation continues for another T_{perturb} .

sudden onset of a sufficiently strong perturbing force to cause a sufficiently strong reaction, compared to noise in stochastic systems. Crucially, and in contrast to the PSD-based approach to inference, instead of observing the trajectory for a long time, the necessary observation duration only has to be on the order of the time delay itself.

Now the question remains, how this feature of systems with time delay can be used to infer the delay time, specifically in noisy Brownian systems. Depending on the parameters, it can be challenging to visually distinguish the characteristic response from noise, despite an average over noise realizations. (In Appendix B 7, Fig. 15, trajectory data for non-linear systems are presented to give the reader some intuition concerning the complexity of responses and the obfuscation of the change in dynamical behavior through thermal noise.) To this end, we propose the use of a neural network to classify trajectories, or rather stochastic time series, in terms of the underlying delay time. We introduce this technique in the following subsection.

B. Convolutional neural network—time series classification

One popular approach for classification of time series data is to employ a supervised machine learning algorithm using (deep) convolutional neural networks [55,56]. The idea behind this approach is to train the network on labeled training data.

A simple analogy can be found in image classification, where a network is fed labeled images of cats and dogs. Here, the goal is to correctly classify cats as cats and dogs as dogs. For this purpose, the weights and biases within the network are iteratively updated during training. The objective in training is to minimize an appropriate loss function and, at the same time, maximize classification accuracy.

This simple binary example can, of course, be extended to many more, even related species (e.g., distinguish between dog and wolf), making the classification more fine grained.

Here, we use this idea to classify one-dimensional time series data (where the time series was produced by application of the perturbation method described in Sec. V A and in Appendix B 4) in terms of the underlying delay time. We make use of the characteristic response of a system with time delay; that is, we assume that the τ -dependent change of dynamical behavior within the time series is a feature the neural network can be trained on. We focus on the discretized derivative of the positional data as shown in Fig. 8(d). In addition to the more distinct response compared to that seen in $x(t)$ [see Figs. 8(a) and 8(b)], this quantity has the advantage that standardization and normalization (see Appendix B 4) are significantly better achievable compared to pure trajectory data.

Figure 9 depicts schematically a simplified network architecture for the classification task. It consists of an input layer, which receives the data obtained from the perturbation method in terms of the discretized derivative. Several hidden convolutional and pooling layers follow. The weights and biases associated with the connections within this part of the network are modified in training to minimize an appropriate loss function. Finally, the hidden layers are connected to the output layer, where a softmax activation function is used such that the values of the output layer represent a discrete probability distribution $P(\tau)$ (see rightmost part of Fig. 9) over an array of τ_{node} , i.e., the delay times used in training. In other words, each node of the output layer represents one specific delay time τ_{node} corresponding to one of the delay times the network was trained on. The node can assume probability values $P(\tau)$ between zero and one with $\sum_{\tau_{\text{node}}} P(\tau) = 1$. Details

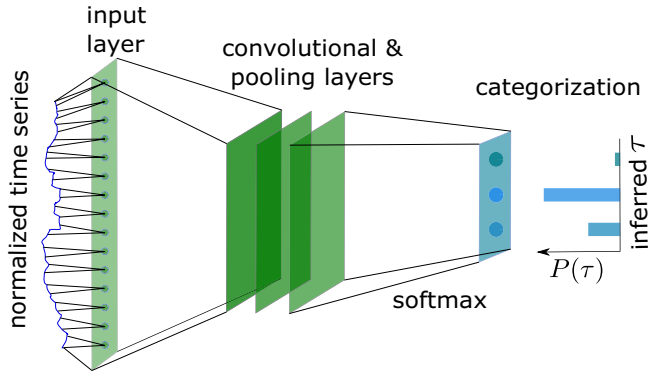


FIG. 9. Visualization of time series classification through convolutional neural networks. Time series data (discretized derivative, see Appendix B 4 for the input data preparation) are fed into the neural network via the input layer. Then, the data go through a sequence of convolutional and pooling layers. In our specific architecture, before the second last layer the network is flattened to be one-dimensional and subsequently fed to the last hidden layer, a densely connected layer. Finally, a densely connected layer is used as an output layer, where a softmax activation function is employed, in order to return a probability distribution for an array of different τ_{node} . The result is a classification of the time series in terms of delay times.

on network architecture, hyperparameter tuning, and training are given in Appendix B 5. For more information on how to train, test, and predict parameters using the convolution neural networks, see Refs. [55,56].

When training the network, we choose a range of delay times as well as a range of remaining parameters. For the linear system, this implies a large number of combinations of the parameters k_a , k_b , and τ , i.e., each parameter set consists of a tuple $(\tau_{\text{node}}, k_a, k_b)$, where all are discrete values within the parameter range. The delay times are represented in the output nodes of the neural network. For each τ_{node} , the training is then carried out for many combinations of the remaining parameters. This makes the inference results robust, with the consequence that the specific combination of parameters is of minor importance for the inference accuracy of the trained network [56].

Since the goal is to be able to infer time delay from noisy time series data, we train the network on several noise realizations for each of the parameter combinations. Specifically, we employ on the order of 100 realizations depending on the complexity of the involved delayed and instantaneous forces.

VI. PERTURBATION AND DEEP LEARNING-BASED APPROACH—NUMERICAL RESULTS

In the following, we discuss results of the perturbation-based approach to infer time delay using neural networks for the trajectory data obtained from Brownian dynamics simulations (see Appendix B 1 for details).

We will first focus on inference by means of a convolutional neural network trained on a linear time-delayed feedback force [see Eq. (2)]. With this network, we infer time delay for trajectory data obtained from numerical simulations of the linear feedback force [see Eq. (2)] as a proof of concept. Subsequently, the tanh feedback force [see Eq. (3)]

is considered to demonstrate that the network trained on a linear feedback force is able to generalize. In principle, the perturbation-based approach to infer time delay using convolutional neural networks trained on linear feedback forces may also be applicable to other systems that exhibit a stationary or nonequilibrium steady state such as Gaussian time-delayed feedback [Eq. (4)]; however, an extensive analysis is beyond the scope of this study.

Finally, we will discuss the inference results obtained from a convolutional neural network trained on trajectories based on Eqs. (11)–(13). We use the trained network to classify trajectory data obtained from numerical simulations of the same equations of motion. In this way, we demonstrate that it is possible to train such a neural network on several functional forms of the (nonlinear) delayed and instantaneous forces simultaneously while maintaining similar prediction accuracy as before. In practical applications, this allows for inference despite not knowing the exact dynamics beforehand.

A. Network trained on linear model applied to linear force data

For our numerical study, we first consider the attractive linear feedback force system based on Eqs. (1) and (2) with $k_a = k_b > 0$. Here, we use a neural network that was trained on the same model to infer time delay. In Fig. 10(a), we show inference results represented by the probability distribution $P(\tau)$ (see Sec. V B). We averaged over the classification probabilities for different values of the parameters k_a and k_b at each data point, as well as over ten noise realizations. Each of the color-coded curves corresponds to one of the exit nodes and represents the probability of a time series to be based on the delay time value τ_{node} . Markers correspond to the underlying delay times of test data, i.e., the time series we obtain from the perturbation approach and feed into the neural network (after applying the preparation procedure discussed in Appendix B 4). We classify time series based on delay times of $\tau = 0\tau_B$ (passive system) up to $\tau = 0.2\tau_B$.

The highest classification probability is observed if a time series is based on a delay time close to one of the τ_{node} the network was explicitly trained on. In those cases, we find $P(\tau) \approx 1$; that is, the network is almost certain. In close proximity to those delay times, the probability remains high where the resulting inference curve for each exit node approximately follows a Gaussian close to the delay values τ_{node} , as shown in Fig. 10(b). For time series where τ lays inbetween two τ_{node} , the exit nodes corresponding to τ_{node} above and below show finite probabilities. The finer the time delay resolution in training, the better the inference results will become, up to a certain point where the network can no longer distinguish between adjacent values of τ_{node} .

B. Network trained on linear model applied to tanh feedback force data

For the red data points in Fig. 10, we use the same network (trained on the linear feedback force with $k_a = k_b > 0$) to infer the delay time for trajectories from simulations of the tanh feedback force system, where we averaged the results for several different values of the parameter A of the tanh feedback force [Eq. (3)] as well as ten noise realizations. We ensured that the linear feedback force system was trained in

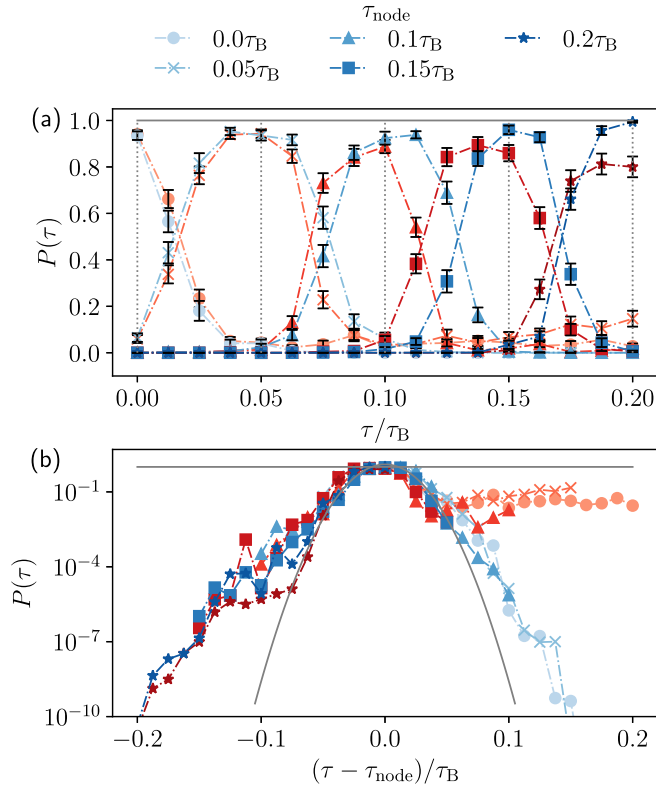


FIG. 10. Classification probabilities for the delay time. (a) Colored symbols (color intensity increasing with τ_{node}): The probability of inferring a delay time at an exit node corresponding to τ_{node} . Blue: Data correspond to Eqs. (1) and (2) for $k_a = k_b > 0$. Red: Data correspond to tanh feedback force [Eq. (3)] with $A > 0$. The neural network is trained on the linear model [Eq. (2)] with parameter range $k_a\tau_B^2 \in [2.5, 15]$, where each input trajectory is prepared as discussed in Appendix B 4. We then infer time delay for both models, linear [Eq. (2)] (blue) with same range of k_a and the tanh model [Eq. (3)] (red) with parameter range $A\tau_B^2 \in [4, 16]$. Average over inference results for ten different trajectories (prepared as discussed in Appendix B 4) for each parameter combination. Each data point is an average over six different values of $k_{a,b}$ and seven different values of A , respectively. Error bars indicate one standard error of the mean. Vertical dotted lines indicate τ_{node} . (b) Same data (but log-scaled y axis) as in panel (a) represented by scaling and shifting the data to reflect the deviation from the output node τ_{node} . Gray Gaussian curve centered around zero as well as horizontal gray line (indicating probability one) act as guides for the eye.

a parameter range that is similar to that of the tanh feedback force, if it was linearly approximated.

Overall, we find accurate inference results, with classification probabilities in excess of 80% for delay times close to those used in training. Looking closely at Fig. 10(a), we observe that the peaks still approximately follow a Gaussian but are shifted slightly to the left; that is, the network slightly overestimates the delay time in the test trajectories, especially for the longer delay times considered. Another observation concerns larger deviations that we find at the highest values of the delay time τ . We interpret this observation as follows. For long delay times, the tanh feedback force [Eq. (3)] becomes constant at large displacements. On the other hand, for small

amplitudes A , the change in the feedback force in response to the perturbation is no longer clearly distinguishable from thermal noise. Therefore, the underlying reason for the observed deviations may be a dominance of diffusive behavior when the tanh feedback force assumes a constant value (beyond the linear regime) for large displacements in combination with a small feedback force amplitude compared to noise. In those cases, the neural network can no longer clearly distinguish between different delay times leading to an increased probability [see Fig. 10(b)] for inferring short delay (where $\tau = 0$ corresponds to diffusive behavior).

In summary, these results indicate that a network trained in linear time-delayed feedback is able to generalize to non-linear systems with the underlying tanh feedback force. For suitable parameter ranges, similar results can be achieved for the Gaussian feedback force discussed above. This indi-

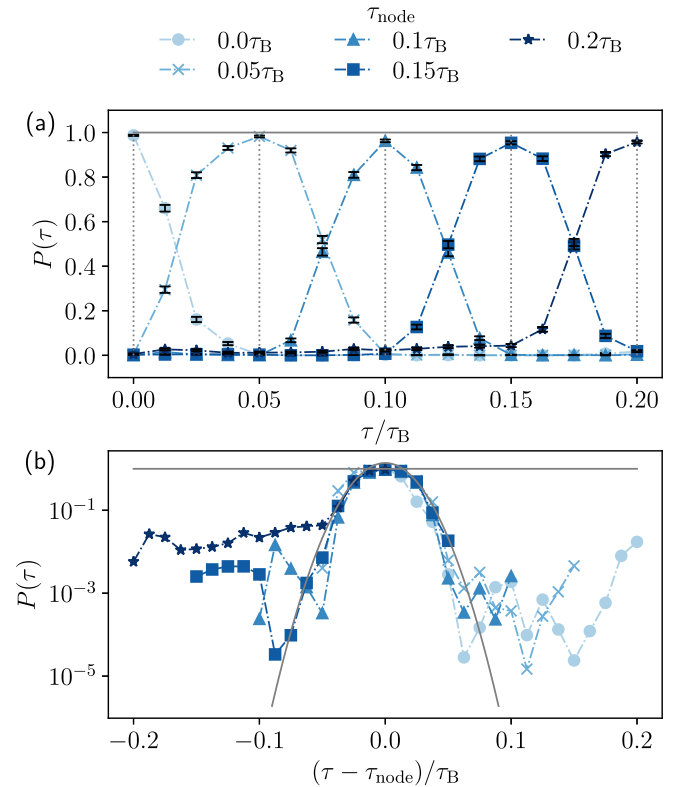


FIG. 11. Classification probabilities for the delay time. Inference results for network trained on Eqs. (11)–(13), where each input trajectory is prepared as discussed in Appendix B 4. We then infer time delay for trajectories obtained from the same model. (a) Symbols (color intensity increasing with τ_{node}): The probability of inferring a delay time at an exit node corresponding to τ_{node} . Average over inference results for ten different trajectories (prepared as discussed in Appendix B 4) for each parameter combination. Each data point is an average over 27 different combinations of the system parameters. Error bars indicate one standard error of the mean. Vertical dotted lines indicate τ_{node} . (b) Same data (but log-scaled y axis) as in panel (a) represented by scaling and shifting the data to reflect the deviation from the output node τ_{node} . Gray Gaussian curve centered around zero as well as horizontal gray line (indicating probability one) act as guides for the eye.

cates that systems subject to time delay that behave linearly in some limit may allow classification via a neural network trained on linear feedback forces, without any further knowledge about the functional form of the feedback forces in the systems.

C. Network trained simultaneously on several nonlinear feedback forces

So far, we have trained our model on only one specific feedback force in a range of parameters. However, in practical applications, it may be desirable to have a more versatile network that can infer the delay time for a wide variety of feedback forces, where the underlying functional form cannot be specified with certainty. To demonstrate that, in principle, this is also possible, we trained a convolutional neural network on the different nonlinear equations of motion in Eqs. (11)–(13). The corresponding results are presented in Fig. 11. Each data point represents the classification probability averaged over ten noise realizations, three different functional forms of forces, and 27 combinations of the system parameters that are different from those used in training. From Fig. 11(a), we observe good inference accuracy with classification probability around 90% close to the delay times on which the network was trained. Furthermore, we again observe activation curves that resemble a Gaussian [see Fig. 11(b)]. Further away from those delay times, we see an elevated activation of other output nodes; i.e., the probability of misclassification increases between two values of τ_{node} but overall remains small (below 10%). Although it is not possible to specify the exact reason for this behavior, it might be due to the systems' dynamical response in some parameter regimes being less sensitive to the perturbation applied.

We note that while not strictly necessary, to improve inference accuracy, we here applied a confining attractive harmonic instantaneous potential for a time interval longer than the delay time before applying the perturbing force to ensure a consistent history.

This proof of concept shows that, in principle, the neural network can be trained on several time-delayed feedback and instantaneous forces without sacrificing inference accuracy. Finally, in principle, this method can also be used on systems with colored noise as demonstrated in Appendix D, Fig. 17.

VII. CONCLUSIONS

In this article, we have proposed and demonstrated in a proof of concept two methods to infer discrete time delay in a wide range of stochastic systems with linear as well as nonlinear time delay dependence.

Starting from linear time-delayed feedback, we have identified and analytically analyzed a characteristic signature of time delay in the PSD, which we have shown to carry over to nonlinear systems. On this basis, we have developed a method to infer time delay and the specific delay time. The method requires no knowledge about the functional form of the of the underlying forces.

Since this PSD-based method requires many long stochastic trajectories for accurate inference, we asked the question

(in the view of experimental situations) whether it is possible to infer time delay based on a short observation of only few realizations of a stochastic system. To this end, we developed a perturbation-based approach that utilizes a characteristic response a system subject to time delay exhibits in the presence of an external perturbation. We then used convolutional neural networks to classify the resulting trajectory data in order to infer time delay for several linear and nonlinear stochastic systems subject to time delay. This method only requires a short duration of the observation on the order of the time delay itself. For this reason, we expect that this method, in principle, also allows inference in the absence of a nonequilibrium steady state or even in unstable systems, provided that underlying forces do not diverge during the observation time.

We emphasize that in experiments, such a perturbation can be realized in principle using state-of-the-art techniques [79,80]. Specifically, for colloidal particles the perturbing force can be applied using optical tweezers [6,81] or external fields [82].

In recent years, several parameter inference techniques for dynamical systems and equation discovery (for ordinary as well as partial differential equations) from (noisy) data have been developed, methods ranging from Bayesian inference [41] to the application of neural network-based techniques [43]. However, to the best of our knowledge, only few such methods exist for systems with time delay, with highly limited applicability to stochastic systems. Our methods to infer time delay, which can be applied to linear as well as nonlinear systems or even without explicit knowledge of the underlying forces, substantially advance the toolbox of parameter inference in stochastic systems.

In this article, we have restricted our analysis and proofs of concept to simple one-dimensional systems consisting of one Brownian particle. However, we expect that the presented method can, in principle, be applied to more complex systems, which would open up many possible avenues for future research and path the way to further extending the typical single colloidal particle experiments [7,9,83], as done in Ref. [61] where feedback with time delay is applied to several interacting colloidal particles.

Staying in one dimension, an interesting extension could be the application of the presented inference methods to partially observed systems. One possible scenario would be a passive probe particle coupled to a kinesin walker (see Appendix C). Here, the question would be if, by observation and control of the probe particle alone, potential time delay in the propulsion of the kinesin walker [84] could be observed.

In this study, we focused on white noise systems for our analytical and numerical considerations but also demonstrated the applicability of our inference methods to systems with exponentially correlated noise. However, in real-world systems more complex types of noise may have to be taken into consideration (e.g., different types of colored noise, multiplicative noise, spatially inhomogeneous noise). For those systems, it remains an open challenge to develop strategies to infer time delay by extending and generalizing the methods discussed here.

Furthermore, from a more fundamental perspective, it would be interesting to investigate whether one could exploit

fluctuation-dissipation-like relations [85–87] to learn about the delay. We note that for the present equations of motion, there is no conventional fluctuation-dissipation relation between noise correlation and memory kernel since the systems are subject to white noise (i.e., zero correlation time), whereas the “memory kernel” corresponding to discrete delay is given by a delta distribution around the delay time τ [33,35]. Still, it would be interesting to explore whether other types of fluctuation relations exist (e.g., when comparing positional moments [88]), and if they could provide us with information about the time delay in a stochastic system.

Furthermore, it would be of great interest to apply the proposed methods to infer time delay in interacting colloidal many-particle systems as well as in systems from related fields, such as systems consisting of motile agents that exhibit collective behavior like flocking or swarming. In recent years, several studies have proposed that time delay may play a role in emergent collective phenomena [23,24,89], where some have explicitly included time delay in their mathematical modeling, leading to interesting effects and insights. On the other hand, advances on the experimental side now make controlling many interacting colloidal particles [34,90] and observing and tracking motile agents [91–93] possible. Having a reliable method to infer the presence of time delay and even the delay time itself would greatly help progressing our knowledge and understanding of the mechanisms behind the emergence of collective phenomena.

Beyond systems with one discrete delay time like those discussed here, systems with several delay times or with distributed delay are receiving considerable interest (see Ref. [35] and references therein). Extending and generalizing the methods discussed here to such systems would be of great value and an open challenge for future research.

In summary, the methods presented in this article can be the first step toward an understanding of the presence and role of time delay in many experiments and real-world applications, since knowledge about the presence of time delay is essential in modeling as well as in controlling dynamical behavior in such situations.

ACKNOWLEDGMENTS

We are grateful to Sandro Azale (University of Padova) for many insightful and stimulating discussions. We also thank Diego Rengifo (Uniandes, TU Berlin) for carefully reading the manuscript and providing valuable feedback. This work was supported in its early stages by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project No. 163436311—SFB 910. D.G. gratefully acknowledges support from the Alexander von Humboldt Foundation.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: ANALYTICAL DETAILS

1. Gaussian time-delayed feedback—analytical details

a. Expansion of the nonlinear feedback force in the limit of small displacements

For small displacements, that is, small a and τ , the feedback force can be linearized [62], analogously to a small delay approximation [64], resulting in a force of the form of Eq. (2), where $k_a = k_b = a/b^2$. Corresponding numerical results are presented in Sec. IV C.

b. Expansion of the nonlinear feedback force in the constant velocity state

We can exploit the marginally stable constant velocity state [62] this kind of feedback force can produce in the deterministic limit. We rewrite the equation of motion in terms of position $x(t)$ relative to the position of a particle moving at constant velocity v_∞ . We can then linearly approximate the feedback force, assuming that the deviation is small. This leaves us with an approximate equation of motion that describes the stochastic dynamics in the “constant velocity state” [62]:

$$\gamma \dot{x}(t) \approx \gamma v_\infty + K[x(t) - x(t - \tau) - v_\infty \tau] + \xi(t), \quad (\text{A1})$$

where we define

$$v_\infty = \frac{\sqrt{2}b}{\tau} \sqrt{-\ln\left(\frac{\gamma b^2}{a\tau}\right)}, \quad (\text{A2})$$

$$K \equiv \frac{a}{b^2} \exp\left[-\frac{(v_\infty \tau)^2}{2b^2}\right] \left(1 - \frac{(v_\infty \tau)^2}{b^2}\right). \quad (\text{A3})$$

Finally, based on Eq. (A1) we can also obtain the Fourier transform for such a system; it reads (with $K < 0$)

$$\tilde{x}_{v_\infty}(\omega) = \frac{(\gamma + K\tau)v_\infty \delta(\omega) + \tilde{\xi}}{\gamma i\omega - K + Ke^{-i\omega\tau}}. \quad (\text{A4})$$

In our simulations, we cannot resolve the zero-frequency contribution, which we therefore neglect in the following. Under this assumption, the Fourier transform of the position would be the same as for the linear feedback force with $k_a = k_b > 0$, which was discussed in the main text (see Sec. III B). We note that persistent motion cannot be considered a stationary state but rather a nonequilibrium steady state. The variance of the position in such systems grows over time, even if the mean particle position goes to zero for long observation times or many noise realizations. For this reason, the differencing method has to be applied in order to obtain the correct numerical results. Corresponding numerical results are presented in Sec. IV C.

2. Definition of the power spectral density via correlation functions

To put our definition of the PSD [Eq. (5)] into a broader context, we note that the power spectral density is often defined via the Fourier transform of the corresponding autocorrelation function [94] under the assumption of stationarity of the underlying stochastic process. We recall that for one

trajectory, the power spectral density is defined as

$$S_{xx}(f) = \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} |\tilde{x}_{\mathcal{T}}(f)|^2, \quad (\text{A5})$$

with frequency f and observation time window $\mathcal{T}$. The quantity $\tilde{x}_{\mathcal{T}}$ is the Fourier transform of the position for some observation time window $\mathcal{T}$,

$$\tilde{x}_{\mathcal{T}}(f) = \int_{-\infty}^{\infty} x_{\mathcal{T}}(t) e^{-i2\pi ft} dt, \quad (\text{A6})$$

$$= \int_{t_0 - \mathcal{T}/2}^{t_0 + \mathcal{T}/2} x(t) e^{-i2\pi ft} dt. \quad (\text{A7})$$

Assuming ergodicity, the right-hand side of Eq. (A7) can be rewritten as [94]

$$\begin{aligned} & \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} |\tilde{x}_{\mathcal{T}}(f)|^2 \\ &= \int_{-\infty}^{\infty} \left[\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_{-\infty}^{\infty} x_{\mathcal{T}}^*(t - t') x_{\mathcal{T}}(t) dt \right] e^{-i2\pi ft'} dt', \end{aligned} \quad (\text{A8})$$

$$= \int_{-\infty}^{\infty} R_{xx}(t') e^{-i2\pi ft'} dt', \quad (\text{A9})$$

with R_{xx} being the autocorrelation function of the position $x(t)$. For example, for a confined Brownian particle, the mean and variance of the position are constant after some time. Under this condition, the power spectral density can be computed via the Fourier transform of the corresponding positional autocorrelation function as well as via the modulus of the Fourier transform of the position. We fully rely on the latter since the autocorrelation is only of limited use in analytical computations for the time delay systems considered here. Furthermore, the Fourier transform of the position of a linear system can be easily computed both analytically and numerically.

3. Linear time-delayed feedback—small frequency behavior

We consider the limit of $\omega \ll 1/\tau$ for $k_a \neq k_b$, and expanding to second order in $\omega\tau$ gives

$$\langle S(\omega) \rangle_{\text{lin}} \approx \frac{2\gamma k_B T}{\omega^2 k_a k_b (\tau - \tau_+)(\tau - \tau_-) + (k_a - k_b)^2}, \quad (\text{A10})$$

where we defined

$$\tau_{\pm} \equiv \gamma \frac{1 \pm \sqrt{1 - k_a/k_b}}{k_a}. \quad (\text{A11})$$

From the approximation equation [Eq. (A10)], we can see that the PSD scales as $1/\omega^2$. For $k_{a,b} > 0$, this behavior can be interpreted as a reduced ability of the feedback trap to confine compared to the instantaneous harmonic trap, resulting in a corresponding deviation from the black curve in Fig. 1 at small ω due to diffusive behavior in the long-time limit. Analogously, for $k_{a,b} < 0$ the frequency dependency is connected to enhanced diffusion due to repulsive time-delayed feedback.

4. Linear time-delayed feedback—marginally stable system

The marginally stable case, where $k_a = k_b$, is special and has been the subject of several previous studies [61,62]. Its peculiar properties are reflected by the PSD (solid blue and

dashed green lines in Fig. 1). For the attractive feedback, in the limit of $k_a = k_b > 0$, the system essentially shows diffusive behavior for all frequencies ($\text{PSD} \propto 1/\omega^2$), indicating that in this case the ability of the feedback to trap the particle has been lost. For repulsive feedback, on the other hand, i.e., for $k_{a,b} < 0$ when $k_a = k_b$, we observe an enhanced diffusion (higher value of the PSD at small frequencies), in agreement with previous studies on this type of system [61,62]. Otherwise, the repulsive feedback can have a confining effect, albeit less strict than a harmonic trap, as can be seen for $|k_a| < |k_b|$ at small frequencies (cyan and gray dashed lines in Fig. 1), where the PSD resembles that of a harmonically trapped particle.

APPENDIX B: NUMERICAL PROCEDURE

We perform numerical simulations, data analysis, and data visualization using the Python programming language. For numerical computations and data processing, we use numpy [95] and scipy [96], and for visualization, we use matplotlib [97]. The deep learning computations are implemented in keras [98], based on tensorflow [99].

1. Numerical solution of the overdamped delay Langevin equation

To solve the SDDE (1), we perform BD simulations based on the Euler-Maruyama integration scheme [100,101]. The resulting discretized version of Eq. (1) reads

$$x_{n+1} = x_n + \gamma^{-1} F(x_n, x_{n-N_\tau}) \Delta t + \gamma^{-1} \sqrt{2k_B T \gamma \Delta t} \eta_n, \quad (\text{B1})$$

where $N_\tau = \tau/\Delta t$ is the number of time steps within one delay time τ , for the discretization time step Δt . η_n is a uncorrelated Gaussian random number with zero mean and unit variance, i.e., $\langle \eta_n \eta_m \rangle = 0$ for $n \neq m$. Equation (B1) is solved provided an initial history for the time interval $[-\tau, 0]$ via the discretized history function $\Phi(t_n) = \Phi_{t_n}$; that is, the trajectory in the discretized time interval $[-\tau, 0]$ has to be known. For all stochastic simulations in this study, we used the trajectories of free Brownian particles as the initial history. Furthermore, we used time steps of $\Delta t = 10^{-4} \tau_B$ (much smaller than the smallest delay time considered) to ensure sufficient numerical precision in all the cases considered.

2. The differencing method

For the computation of the (noise-averaged) PSD [Eq. (7)], a stationary time series is required, which is quite often not the case in physical systems. Already, for a passive Brownian particle this problem has to be addressed, while for feedback-driven systems it can become even more pronounced.

Figure 12 presents exemplary data for systems with attractive and repulsive harmonic time-delayed feedbacks, which both behave diffusively, especially in the long-time limit. The cyan and red data correspond to PSD computed by directly applying the Fourier transform to the nonstationary positional data. We observe a shift compared to the analytical expectation that is hardly visible in the PSD but becomes apparent in $\omega^2 \text{PSD}$. While when averaging over a small number of noise realizations, this shift is lost in noise, when averaging over a larger number of noise realizations, the shift becomes obvious.

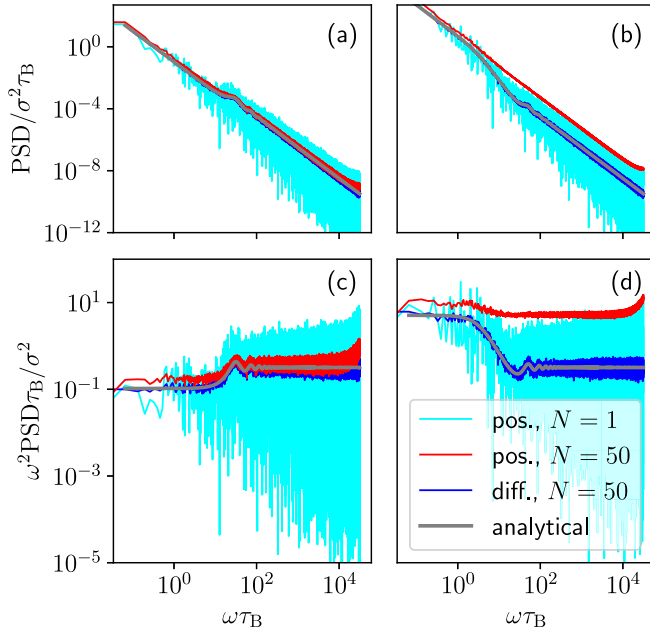


FIG. 12. Comparison of position-based and differencing-based computation of PSD and $\omega^2\text{PSD}$ for a system with linear feedback force [Eq. (2)]. (a), (c) Marginally stable system with attractive harmonic feedback; panels (b) and (d) are for repulsive harmonic feedback below the deterministic threshold, i.e., $|k_a| = |k_b| < \gamma/\tau$.

The reason behind this behavior can be understood when considering that we are taking the Fourier transform of a (within the observation period) not strictly ergodic and certainly nonstationary quantity. More specifically, the time series is nonstationary since the variance of the position grows over time. Changing the temporal resolution or the length of the trajectory only has a small effect on the observed shift. The observed shift can intuitively be understood as follows: In a stationary system (assuming zero mean and constant variance), the position will (more or less) fluctuate around zero, given that the observation window is long enough. However, if the process is not stationary in the sense that the variance grows as a function of time, in some realizations, the position will be positive (or negative) for a significant part of the observation time. When taking the Fourier transform of such a trajectory, it will be somewhat biased. If we took the average over noise realizations now, we could expect this bias to average out over different realizations. However, following the definition of the PSD [Eq. (7)], we compute the square modulus of the Fourier transform of the position before applying the noise average. For a stationary process, this does not matter. However, in our case, the bias cannot be averaged out but rather accumulates and depends on the variance of the process and how much of the observation period will be uninterruptedly spent on one side of the mean position. We therefore have to approach the problem differently and make the time series stationary before even computing the Fourier transform in order to obtain results that match our analytical expectation.

To this end, we use the differencing method; that is, instead of using the positions as time series, we consider the positional difference over a fixed time interval Δt . This is a textbook

technique from time series analysis [102] to make nonstationary time series stationary. For our case, instead of computing the Fourier transform of the position we rather compute the Fourier transform of the positional difference, i.e.,

$$\frac{\widetilde{\Delta x(t)}}{\Delta t} = \frac{\mathcal{F}[x(t) - x(t - \Delta t)]}{\Delta t}, \quad (\text{B2})$$

where the operator $\mathcal{F}$ represents the Fourier transform. Intuitively, this looks like the Fourier transform over a temporal derivative (which we already exploited in the analytical considerations) or essentially (for small Δt) the Fourier transform of the right-hand side of the discretized Langevin equation. Applying this method, we can obtain a numerical result that matches exactly the analytical expectation, as shown by the dark blue curves in Fig. 12.

The differencing method was used for all results discussed in Sec. IV.

3. Extraction of frequencies corresponding to minima and maxima of the PSD

In the following, we describe a procedure that allows us to extract the delay time from simulation data, which was used in Sec. IV D. It is based on the analytical results obtained for the linear system, but can also be applied to systems that depend on $x(t)$ and/or $x(t - \tau)$ in a nonlinear fashion. This procedure relies on the spacing of the positions of maxima and minima in the quantity $\omega^2\text{PSD}$, and is applicable as long as the signal-to-noise ratio in the system allows for reliable determination and extraction of the position of these minima and maxima. As such, the larger the number of available noise realizations, the better the inference results. For example, Fig. 12 illustrates the level of noise that has to be dealt with in this kind of system in the case of a marginally stable time delay system. In order to obtain a good polynomial fit over multiple peaks, we need to average over a sufficient number of noise realizations; for the figures in the main text [Figs. 6(a)–6(c)], 400 noise realizations were used.

While, in principle, the positions of the minima can be obtained manually, it is more reliable and convenient to use a spline interpolation for a subsequent semiautomatic extraction of the positions of minima and maxima. It is semiautomatic in the sense that one has to determine for how many minima and maxima a reliable fit can be obtained. Specifically, we use quartic splines for fitting the numerical data, which allows us to extract the roots of its derivative automatically. As the data are noisy and somewhat irregular, special care has to be taken when adjusting the smoothing parameter in the spline interpolation, such that the interpolation faithfully represents the location of the minima and maxima in the numerical data. Furthermore, limiting the range of the interpolation can significantly improve the results since at low frequencies additional oscillations with different oscillation period are observed in some cases and at high frequencies the oscillations are lost in noise.

Specifically, the extraction of the delay time requires the following steps:

- (1) Apply spline interpolation to $\omega^2\text{PSD}$ using quartic splines implemented within `scipy`.

(2) Take derivative of the resulting polynomial (built-in function of spline interpolation in scipy).

(3) Compute the roots ω_n of the derivative of the polynomial corresponding to the minima and maxima in ω^2 PSD (built-in function of spline interpolation in scipy).

(4) Discard the first N_{discard} minima and maxima (until the minima and maxima have approximately equal spacing) (see Sec. III B).

(5) Determine N_{ext} , the number of maxima and minima that are reliably represented by the polynomial; terminate the fit before the oscillations are lost in noise.

The inferred delay time is then given by

$$2\pi/\tau_{\text{inferred}} = \frac{\sum_{n=1}^{N_{\text{ext}}-2} (\omega_{n+2} - \omega_n)}{(N_{\text{ext}} - 2)}, \quad (\text{B3})$$

where τ_{inferred} is the inferred delay time, N_{ext} is the number of minima and maxima that can be reliably identified, and we averaged the distances between all available pairs of consecutive minima and maxima.

4. Deep learning: data preparation

For the perturbation-based deep learning approach, we first need to prepare the data. We apply the following steps:

(1) Take an ensemble average over 25 noise realizations over the positional data at each point in time.

(2) Apply the differencing method, $\{\bar{x}(t)\} \equiv \{x(t + dt)\} - \{x(t)\}$, to render the time series stationary.

(3) Apply a simple moving average over three subsequent time steps for denoising.

(4) Normalize by dividing by the maximum value during the observation period.

The data are prepared in this way for training the neural network as well as for inference using the trained network (see Sec. VI).

5. Deep learning: network architecture and hyperparameter tuning

For the classification of trajectory data, we use a convolutional deep neural network (see, e.g., Refs. [55,56]) implemented using keras [98] based on tensorflow [99]. The neural network (see Sec. V B, especially Fig. 9 for a simplified schematic depiction) consists of an input layer (the number of nodes equals the number of relevant time steps), hidden layers, and an output layer (the number of nodes equals the number of classes, i.e., delay times the network is trained on). For hidden layers, considerations such as complexity, regularization, and available resources play a crucial role. Here, we focus on three convolutional layers and pooling layers to limit complexity. We use a hyperparameter tuning scheme in order to achieve the desired accuracy for classification by the network. This means that over how many nodes pooling is applied in a pooling layer, filter and kernel size in convolutional layers, and the optimization method and the activation function are chosen automatically. Before the final hidden layer, the network is flattened. The final hidden layer is then a densely connected layer followed by the densely connected output layer. The number of nodes of the output layer is equal to the number of different delay times on which the model

TABLE II. Network structure by layers; hyperparameters, i.e., over how many nodes pooling is applied in a pooling layer, filter and kernel size in convolutional layers, and the optimization method and the activation function are optimized by hyperparameter tuning. For the input data preparation, see Appendix B 4.

Layer	Role/structure	Activation
1	Input	—
2	Normalization	—
3	Convolutional	ReLU
4	Max pooling	—
5	Convolutional	ReLU
6	Max pooling	—
7	Convolutional	ReLU
8	Max pooling	—
9	Flatten	—
10	Dense	ReLU
11	Output/dense	softmax

is trained. A *softmax* activation function is used to obtain a classification in terms of the probability that the classified trajectory is based on a certain delay time. The network is then constructed as shown in Table II. As the loss function, which is minimized during training, the sparse categorical cross entropy was used. Figure 13 presents exemplary accuracy [Fig. 13(a)] and loss [Fig. 13(b)] as a function of training epochs for the training of the neural network on several nonlinear feedback forces [Eqs. (11)–(13)] simultaneously. We observe good convergence already at as little as 20 training epochs.

6. Characteristic response of a nonlinear feedback force system

Figure 14 presents exemplary data for the response of systems with nonlinear feedback forces. In all cases, a complex but still characteristic response is observed. As for linear

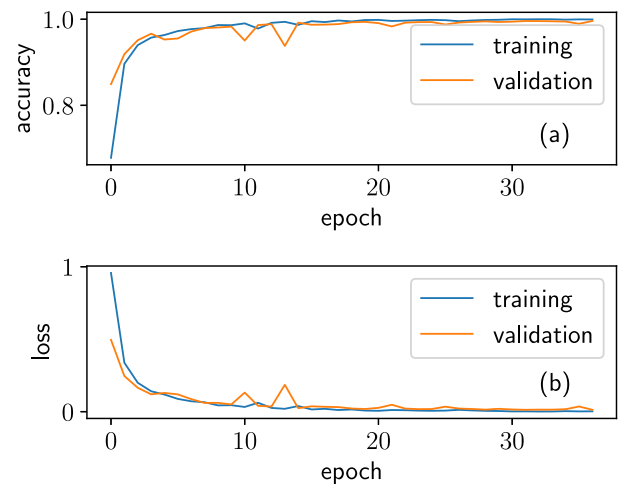


FIG. 13. Training the convolutional neural network on nonlinear feedback forces [Eqs. (11)–(13)]: Panel (a) shows the classification accuracy, and panel (b) shows the training loss (computed via sparse categorical crossentropy), both plotted as functions of training epochs for the training and validation datasets.

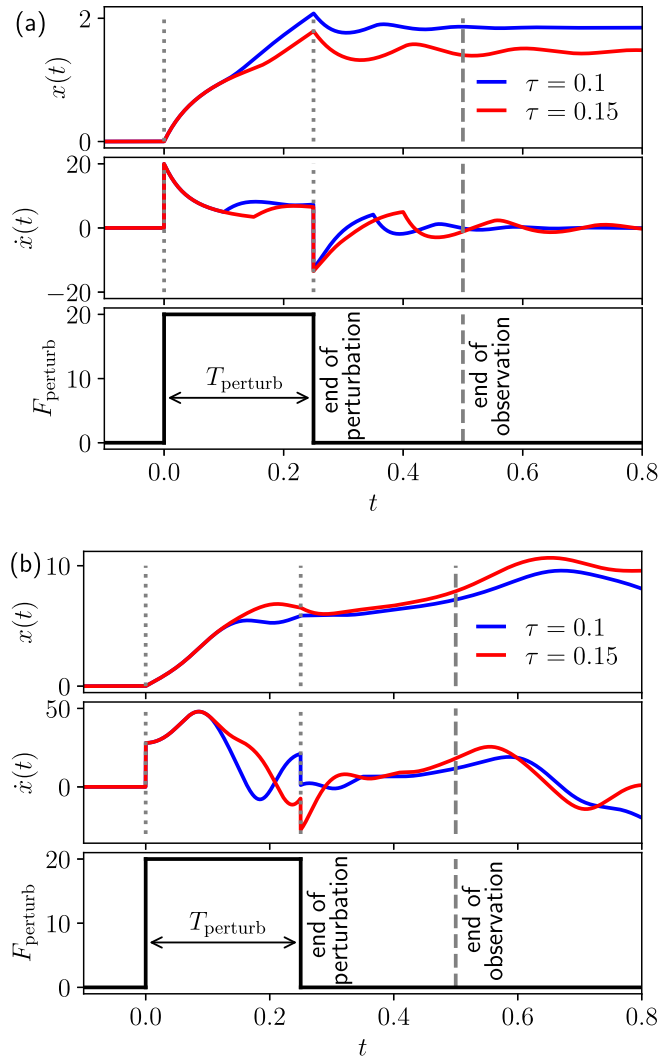


FIG. 14. Characteristic response of a nonlinear system with time delay to an external perturbation (see also Sec. V A). (a) Response of the tanh feedback force [Eq. (3)] to a constant external force. (b) The response of a system based on Eq. (13) to the same type of perturbation. Vertical dotted lines indicate the onset and termination of the perturbation, while the vertical dashed line marks the end of the observation window.

systems, a change in dynamical behavior is observed one delay time after switching on the perturbation as well as one delay time after switching off the perturbation.

7. Characteristic response of a stochastic system with nonlinear feedback force

Figure 15 presents exemplary data for the response of stochastic systems (11)–(13) subject to nonlinear and delayed forces with corresponding inference results. In contrast to the linear system (for comparison see Fig. 8), the response of a nonlinear system can be more complex, as indicated by the solid lines representing the deterministic response in Fig. 15. In the stochastic data, the change in dynamical behavior that constitutes the characteristic response is obfuscated by ther-

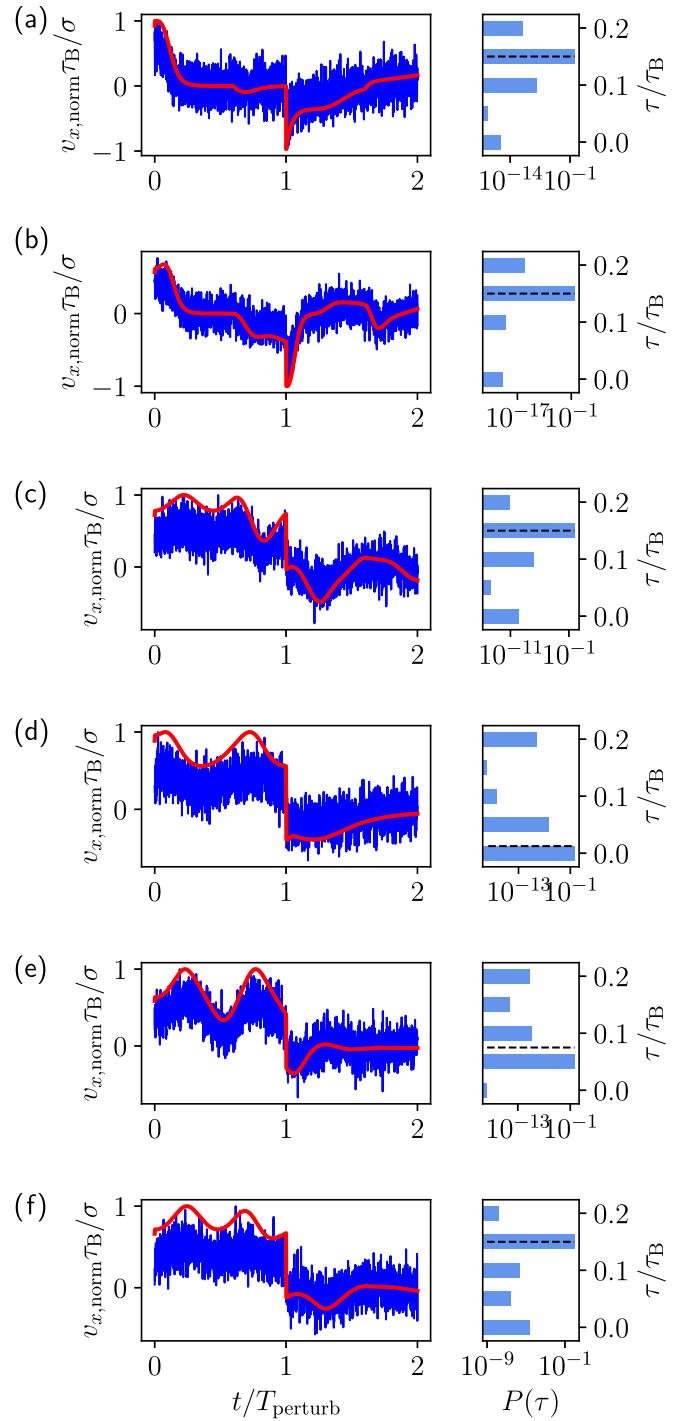


FIG. 15. Characteristic response of a stochastic nonlinear system with time delay to an external perturbation (see also Sec. V A) and corresponding inference results (right panels). Trajectories prepared as discussed in Appendix B 4. Solid line: corresponding deterministic response. Dashed horizontal line: true delay time. (a)–(c) $\tau = 0.15\tau_B$. (a) Equation (11) at $a = 13k_B T/\sigma^4$, $b = 3k_B T/\sigma^2$, and $c = 11k_B T/\sigma^2$; (b) Eq. (12) at $a = 13k_B T/\sigma^4$, $b = 18k_B T/\sigma^2$, and $c = 17k_B T/\sigma^3$; and (c) Eq. (13) at $a = 7k_B T/\sigma$, $b = 10k_B T/\sigma$, and $c = 17k_B T/\sigma$. (d)–(f) Trajectories based on Eq. (13) at (d) $\tau = 0.0125\tau_B$, $a = 7k_B T/\sigma$, $b = 10k_B T/\sigma$, and $c = 17k_B T/\sigma$; (e) $\tau = 0.075\tau_B$, $a = 13k_B T/\sigma$, $b = 3k_B T/\sigma$, and $c = 11k_B T/\sigma$; (f) $\tau = 0.15\tau_B$, $a = 9k_B T/\sigma$, $b = 3k_B T/\sigma$, and $c = 7k_B T/\sigma$.

mal noise. Still, the neural network predicts delay times with high accuracy, if these are close to the delay times the network was trained on, and if the underlying trajectories are prepared as described in Appendix B 4. For delay times in-between two subsequent τ_{node} , the inference results tend to be less accurate.

APPENDIX C: EXTENSION TO A SYSTEM WITH HARMONICALLY COUPLED PROBE PARTICLE

Here, we consider a one-dimensional system of a Brownian particle subject to time-delayed feedback, which is coupled to a passive probe particle via an attractive harmonic potential. We refer to the feedback-driven particle using the superscript fb and to the probe particle using the superscript pr. The dynamics of the feedback-driven particle is given by an overdamped Langevin equation. The tracer particle is coupled to the feedback driven particle via the position x_{fb} , which enters the harmonic coupling potential. The coupled equations of motion then read

$$\gamma_{\text{fb}} \frac{dx_{\text{fb}}}{dt} = F_{\text{fb}}(x_{\text{fb}}(t), x_{\text{fb}}(t - \tau)) - F_c + \xi_{\text{fb}}(t), \quad (\text{C1})$$

$$\gamma_{\text{pr}} \frac{dx_{\text{pr}}}{dt} = F_c + \xi_{\text{pr}}(t), \quad (\text{C2})$$

with $F_c = F_c(x_{\text{pr}}(t), x_{\text{fb}}(t))$. The constants γ_{fb} and γ_{pr} are the respective friction constants with $\gamma_{\text{fb}} = \gamma_{\text{pr}} = \gamma$ for simplicity. The coupling potential U reads

$$U(x_{\text{pr}}(t), x_{\text{fb}}(t)) = \frac{k}{2} (x_{\text{pr}}(t) - x_{\text{fb}}(t))^2, \quad (\text{C3})$$

resulting in the force $F_c = -\nabla_{x_{\text{pr}}} U$ and analogously $-\nabla_{x_{\text{fb}}} U = -F_c$ that acts on the coupled probe particle. Furthermore, ξ_{fb} and ξ_{pr} represent a one-dimensional Gaussian white noise, respectively, with zero mean and correlation function $\langle \xi_{\text{fb/pr}}(t) \xi_{\text{fb/pr}}(t') \rangle = 2\gamma_{\text{fb/pr}} k_B T \delta(t - t')$. Then, the Fourier transform of the dynamics reads as follows:

$$\gamma i\omega \tilde{x}_{\text{fb}}(\omega) = -[k_a \tilde{x}_{\text{fb}}(\omega) - k_b e^{-i\omega\tau} \tilde{x}_{\text{fb}}(\omega)] - k[\tilde{x}_{\text{fb}}(\omega) - \tilde{x}_{\text{pr}}(\omega)] + \tilde{\xi}_{\text{fb}}(\omega), \quad (\text{C4})$$

$$\gamma i\omega \tilde{x}_{\text{pr}}(\omega) = k[\tilde{x}_{\text{fb}}(\omega) - \tilde{x}_{\text{pr}}(\omega)] + \tilde{\xi}_{\text{pr}}(\omega). \quad (\text{C5})$$

We can now solve for $\tilde{x}_{\text{fb}}$ and $\tilde{x}_{\text{pr}}$, respectively, and obtain

$$\tilde{x}_{\text{fb}} = \frac{\frac{k \tilde{\xi}_{\text{pr}}}{\gamma i\omega + k} + \tilde{\xi}_{\text{fb}}}{\gamma i\omega + k_a - k_b e^{-i\omega\tau} + k - \frac{k^2}{\gamma i\omega + k}}, \quad (\text{C6})$$

$$\tilde{x}_{\text{pr}} = \frac{(k \tilde{x}_{\text{fb}} + \tilde{\xi}_{\text{pr}})}{\gamma i\omega + k}. \quad (\text{C7})$$

The analytical expression for the PSD then follows from the definition as for the simple linear system. From Eqs. (C6) and (C7), we can already see that the characteristic oscillations, and, thus, the signature of time delay extend to the probe particle. Therefore and in principle, by observing the probe particle alone, inference of time delay is possible with a method similar to that proposed in Sec. V.

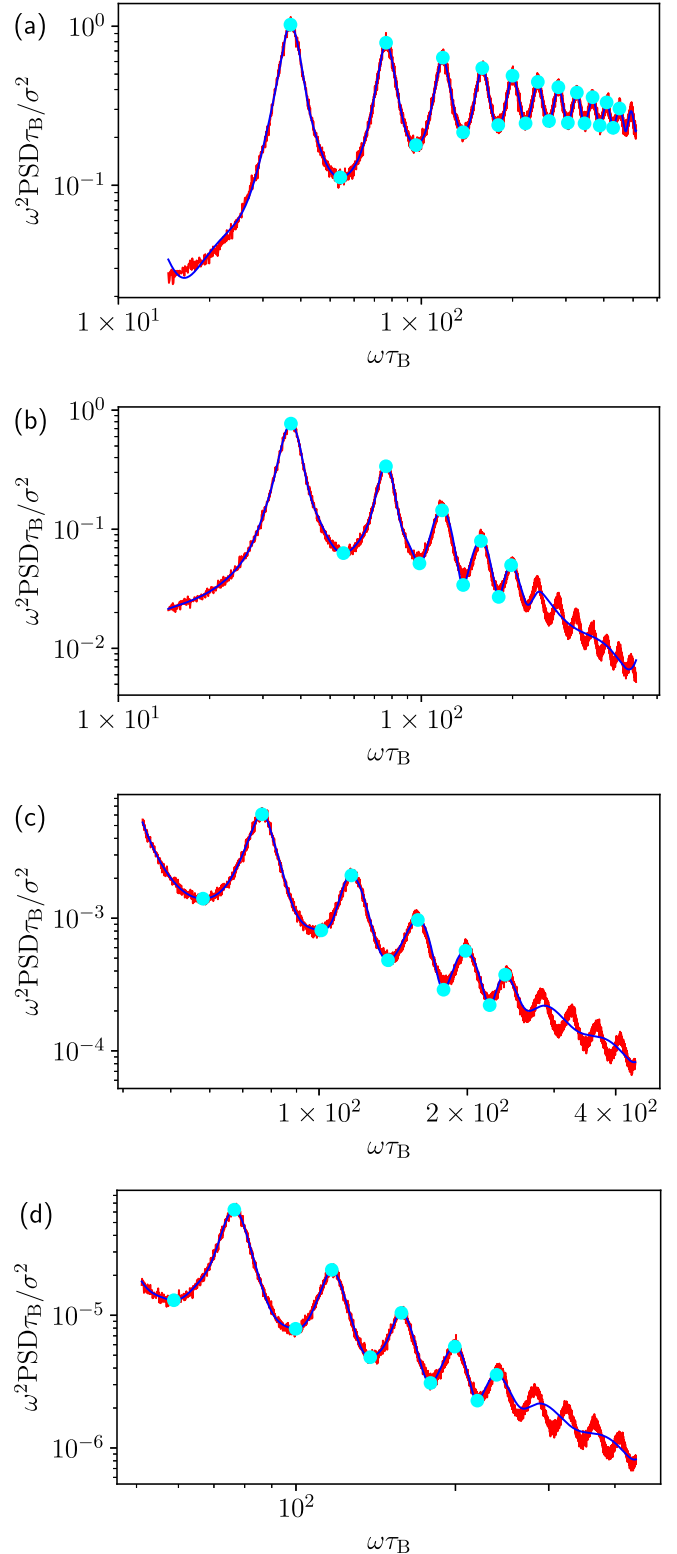


FIG. 16. Numerically simulated PSD (red) as a function of ω for nonlinear systems (12) with white noise replaced by colored noise. In each panel: Polynomial fit (blue) using the quartic splines and automatically extracted positions (cyan circles) of maxima and minima using the roots of the derivative of the polynomial (see Appendix B 3). See Table III for the corresponding inference results. Colored noise at correlation times: (a) $t_{\text{corr}} = 0.01\tau$, (b) $t_{\text{corr}} = 0.1\tau$, (c) $t_{\text{corr}} = 1\tau$, and (d) $t_{\text{corr}} = 10\tau$, other parameters as in Fig. 6 and Table I.

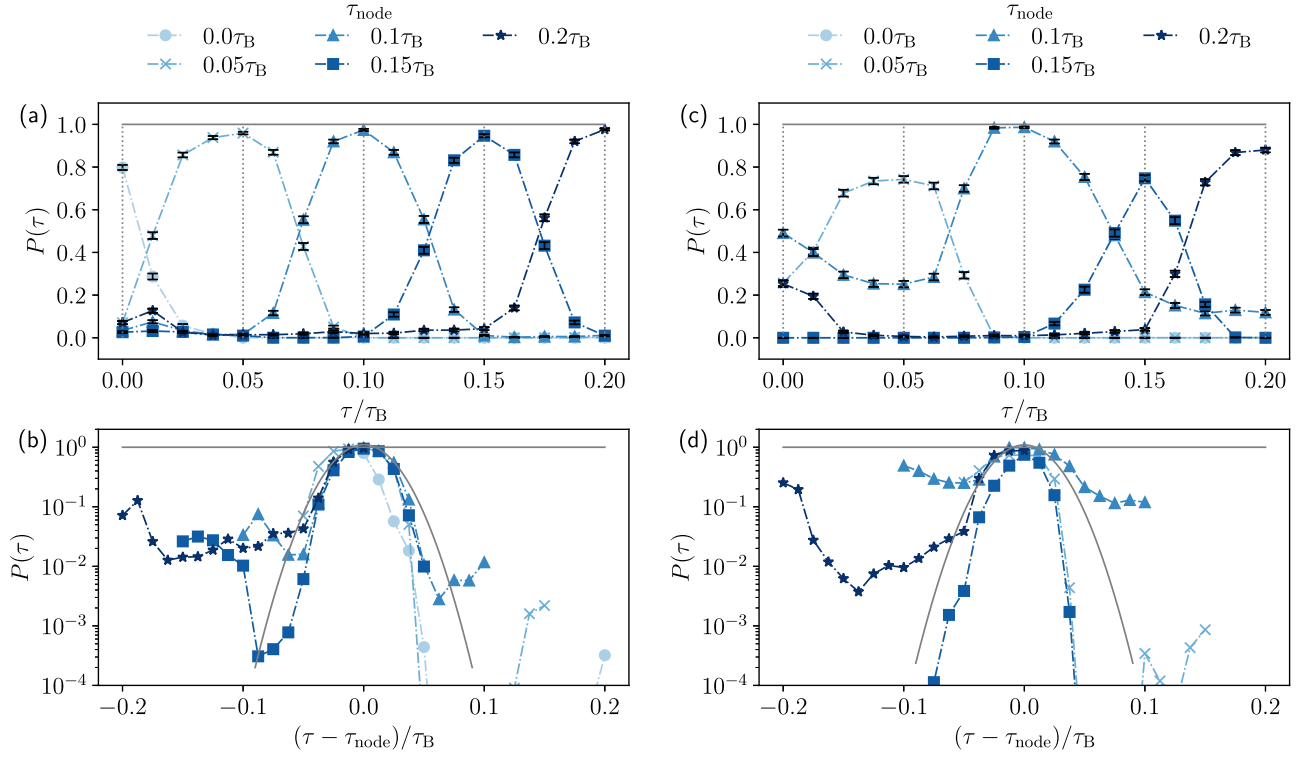


FIG. 17. Inference results for a network trained on nonlinear systems (11)–(13) with white noise, for trajectories from colored noise systems based on Eqs. (11)–(13) at correlation times $t_{\text{corr}} = 0.001\tau$ [panels (a) and (b)] and $t_{\text{corr}} = 0.1\tau$ [panels (c) and (d)]. (a), (c) Symbols (color intensity increasing with τ_{node}): The probability of inferring a delay time at an exit node corresponding to τ_{node} . Average over inference results for ten different trajectories (prepared as discussed in Appendix B 4) for each parameter combination. Each data point is an average over 27 different combinations of the system parameters. Error bars indicate one standard error of the mean. Vertical dotted lines indicate τ_{node} . (b), (d) Same data (but log-scaled y axis) as in panels (a) and (c) represented by scaling and shifting the data to reflect the deviation from the output node τ_{node} . Gray Gaussian curve centered around zero as well as horizontal gray line (indicating probability one) act as guides for the eye.

APPENDIX D: SYSTEMS WITH EXPONENTIALLY CORRELATED NOISE—NUMERICAL RESULTS

In this Appendix, we demonstrate that our methods of inferring the time delay are not limited to white noise systems but can be applied or extended to systems with colored noise. In other words, the signatures of time delay remain intact even in colored noise systems.

To this end, we first introduce exponentially correlated colored noise and subsequently present exemplary data for the PSD-based inference approach as well as the neural network-based approach for nonlinear systems. Specifically, we consider the nonlinear systems in Eqs. (11)–(13) and replace the Gaussian white noise (present on their right-hand side) by an additive colored noise [103]:

$$\dot{\zeta}(t) = -\frac{1}{t_{\text{corr}}} \zeta(t) + \sqrt{\frac{2\gamma k_B T}{t_{\text{corr}}^2}} \eta(t), \quad (\text{D1})$$

where t_{corr} is the correlation time and $\eta(t)$ is a white noise (delta correlated) with zero mean and unit variance.

It turns out that in the long-time limit (i.e., the stationary state), the colored noise $\zeta(t)$ is exponentially correlated:

$$\langle \zeta(t) \zeta(t') \rangle_{\text{ss}} = \frac{\gamma k_B T}{t_{\text{corr}}} e^{-|t-t'|/t_{\text{corr}}}. \quad (\text{D2})$$

Before proceeding further, we would like to highlight its two limiting behaviors. (1) In the limit $t_{\text{corr}} \rightarrow 0$, for $t = t'$, the covariance (D2) diverges; otherwise, the correlation goes to 0 for $t \neq t'$. Further, integrating the right-hand side of Eq. (D2) from $t \in (-\infty, +\infty)$ gives $2\gamma k_B T$. Therefore, this correlation is a delta function in this particular limit (with a prefactor $2\gamma k_B T$). (2) In the limit $t_{\text{corr}} \rightarrow \infty$, $\dot{\zeta} \rightarrow 0$, i.e., ζ becomes a quenched random Gaussian variable with zero mean and variance $\gamma k_B T / (t_{\text{corr}}) \rightarrow 0$ (as $t_{\text{corr}} \rightarrow \infty$).

Figure 16 shows the PSD of the extracted positions of minima and maxima. Table III presents the corresponding inference results for one delay time at different values of the correlation time t_{corr} . Interestingly, while the behavior of the PSD as a function of the frequency changes as the colored noise is added to the system and the correlation time t_{corr}

TABLE III. Inference of the delay times corresponding to system in Eq. (12) based on the PSD data shown in Fig. 16.

t_{corr}/τ	τ/τ_B	$\tau_{\text{inferred}}/\tau_B$	σ_τ/τ_B
0.01	0.15	0.151	± 0.003
0.1	0.15	0.152	± 0.007
1	0.15	0.153	± 0.009
10	0.15	0.152	± 0.010

is increased, the characteristic signature of time delay, i.e., oscillations with a delay-specific oscillation period, remains.

For the neural network approach, Fig. 17 shows data for inference results of the modified nonlinear systems in Eqs. (11)–(13) at two different correlation times using a network trained solely on white noise systems. In the white noise limit, i.e., $t_{\text{corr}} \ll \tau$, the inference results resemble those obtained for white noise systems. However, when the correlation

time is on the order of the delay time or longer, the resulting trajectories look significantly different than before. Thus, as we expect, a neural network trained on systems with only white noise may not be able to accurately infer the time delay for systems with longer correlation times. Therefore, one requires to train the neural network with different correlation times (including different combinations of other parameters) to have better accuracy when inferring the time delay.

-
- [1] E. Schöll, S. H. L. Klapp, and P. Hövel, *Control of Self-Organizing Nonlinear Systems* (Springer, Cham, 2016).
 - [2] K. Pyragas, Continuous control of chaos by self-controlling feedback, *Phys. Lett. A* **170**, 421 (1992).
 - [3] K. Pyragas, Control of chaos via extended delay feedback, *Phys. Lett. A* **206**, 323 (1995).
 - [4] A. P. Tsui and A. J. Jones, The control of higher dimensional chaos: Comparative results for the chaotic satellite attitude control problem, *Physica D* **135**, 41 (2000).
 - [5] M. Gavrilov and J. Bechhoefer, Feedback traps for virtual potentials, *Philos. Trans. R. Soc. A* **375**, 20160217 (2017).
 - [6] A. Kumar and J. Bechhoefer, Nanoscale virtual potentials using optical tweezers, *Appl. Phys. Lett.* **113**, 183702 (2018).
 - [7] M. C. R. Bell-Davies, A. Curran, Y. Liu, and R. P. A. Dullens, Dynamics of a colloidal particle driven by continuous time-delayed feedback, *Phys. Rev. E* **107**, 064601 (2023).
 - [8] K. Lichtner and S. H. L. Klapp, Feedback-controlled transport in an interacting colloidal system, *Europhys. Lett.* **92**, 40007 (2010).
 - [9] B. J. Lopez, N. J. Kuwada, E. M. Craig, B. R. Long, and H. Linke, Realization of a feedback controlled flashing ratchet, *Phys. Rev. Lett.* **101**, 220601 (2008).
 - [10] S. A. M. Loos and S. H. L. Klapp, Heat flow due to time-delayed feedback, *Sci. Rep.* **9**, 2491 (2019).
 - [11] R. A. Kopp and S. H. L. Klapp, Heat production in a stochastic system with nonlinear time-delayed feedback, *Phys. Rev. E* **110**, 054126 (2024).
 - [12] T. Munakata, S. Iwama, and M. Kimizuka, Linear stochastic system with delay: Energy balance and entropy production, *Phys. Rev. E* **79**, 031104 (2009).
 - [13] T. Munakata and M. L. Rosinberg, Entropy production and fluctuation theorems for Langevin processes under continuous non-Markovian feedback control, *Phys. Rev. Lett.* **112**, 180601 (2014).
 - [14] M. Debiossac, D. Grass, J. J. Alonso, E. Lutz, and N. Kiesel, Thermodynamics of continuous non-Markovian feedback control, *Nat. Commun.* **11**, 1360 (2020).
 - [15] M. Debiossac, M. L. Rosinberg, E. Lutz, and N. Kiesel, Non-Markovian feedback control and acausality: An experimental study, *Phys. Rev. Lett.* **128**, 200601 (2022).
 - [16] M. Fränz, S. Muñoz-Landin, V. Holubec, and F. Cichos, Fully steerable symmetric thermoplasmonic microswimmers, *ACS Nano* **15**, 3434 (2021).
 - [17] J. Hindes, K. Szwajkowska, and I. B. Schwartz, Hybrid dynamics in delay-coupled swarms with mothership networks, *Phys. Rev. E* **94**, 032306 (2016).
 - [18] G. Stoica, A stochastic delay financial model, *Proc. Am. Math. Soc.* **133**, 1837 (2005).
 - [19] J. L. Callen, M. Khan, and H. Lu, Accounting quality, stock price delay, and future stock returns, *Contemp. Account. Res.* **30**, 269 (2013).
 - [20] M. Mijalkov, A. McDaniel, J. Wehr, and G. Volpe, Engineering sensorial delay to control phototaxis and emergent collective behaviors, *Phys. Rev. X* **6**, 011008 (2016).
 - [21] R. Sipahi, F. M. Atay, and S.-I. Niculescu, Stability of traffic flow behavior with distributed delays modeling the memory effects of the drivers, *SIAM J. Appl. Math.* **68**, 738 (2008).
 - [22] L. Yu, T. Li, and Z.-K. Shi, Density waves in a traffic flow model with reaction-time delay, *Physica A* **389**, 2607 (2010).
 - [23] V. Holubec, D. Geiss, S. A. M. Loos, K. Kroy, and F. Cichos, Finite-size scaling at the edge of disorder in a time-delay vicsek model, *Phys. Rev. Lett.* **127**, 258001 (2021).
 - [24] F. Pakpour and T. Vicsek, Delay-induced phase transitions in active matter, *Physica A* **634**, 129453 (2024).
 - [25] R. A. Kopp and S. H. L. Klapp, Spontaneous velocity alignment of Brownian particles with feedback-induced propulsion, *Europhys. Lett.* **143**, 17002 (2023).
 - [26] S. Tarama, S. U. Egelhaaf, and H. Löwen, Traveling band formation in feedback-driven colloids, *Phys. Rev. E* **100**, 022609 (2019).
 - [27] S. Tarama, Engineering living worms and active crystals with colloids propelled by attractive time-delayed feedback, *J. Chem. Phys.* **163**, 044903 (2025).
 - [28] E. Pigani, D. Sgarbossa, S. Suweis, A. Maritan, and S. Azaele, Delay effects on the stability of large ecosystems, *Proc. Natl. Acad. Sci. USA* **119**, e2211449119 (2022).
 - [29] M. G. Earl and S. H. Strogatz, Synchronization in oscillator networks with delayed coupling: A stability criterion, *Phys. Rev. E* **67**, 036204 (2003).
 - [30] T. Dahms, J. Lehnert, and E. Schöll, Cluster and group synchronization in delay-coupled networks, *Phys. Rev. E* **86**, 016202 (2012).
 - [31] P. Siegle, I. Goychuk, P. Talkner, and P. Hänggi, Markovian embedding of non-Markovian superdiffusion, *Phys. Rev. E* **81**, 011136 (2010).
 - [32] S. A. M. Loos and S. H. L. Klapp, Fokker–Planck equations for time-delayed systems via Markovian embedding, *J. Stat. Phys.* **177**, 95 (2019).
 - [33] T. J. Doerries, S. A. M. Loos, and S. H. L. Klapp, Correlation functions of non-Markovian systems out of equilibrium: Analytical expressions beyond single-exponential memory, *J. Stat. Mech.* (2021) 033202.
 - [34] M. Fränz and F. Cichos, Active particle feedback control with a single-shot detection convolutional neural network, *Sci. Rep.* **10**, 12571 (2020).

- [35] S. A. M. Loos, S. Hermann, and S. H. L. Klapp, Medium entropy reduction and instability in stochastic systems with distributed delay, *Entropy* **23**, 696 (2021).
- [36] C. Otto, K. Lüdge, and E. Schöll, Modeling quantum dot lasers with optical feedback: Sensitivity of bifurcation scenarios, *Phys. Stat. Sol. (B)* **247**, 829 (2010).
- [37] C. Otto, B. Globisch, K. Lüdge, E. Schöll, and T. Erneux, Complex dynamics of semiconductor quantum dot lasers subject to delayed optical feedback, *Int. J. Bifurcation Chaos* **22**, 1250246 (2012).
- [38] U. Dorner and P. Zoller, Laser-driven atoms in half-cavities, *Phys. Rev. A* **66**, 023816 (2002).
- [39] A. Carmele, J. Kabuss, F. Schulze, S. Reitzenstein, and A. Knorr, Single photon delayed feedback: A way to stabilize intrinsic quantum cavity electrodynamics, *Phys. Rev. Lett.* **110**, 013601 (2013).
- [40] *Complex Time-Delay Systems: Theory and Applications*, edited by F. M. Atay, Understanding Complex Systems (Springer, Berlin, Heidelberg, 2010).
- [41] U. von Toussaint, Bayesian inference in physics, *Rev. Mod. Phys.* **83**, 943 (2011).
- [42] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proc. Natl. Acad. Sci. USA* **113**, 3932 (2016).
- [43] Z. Chen, Y. Liu, and H. Sun, Physics-informed learning of governing equations from scarce data, *Nat. Commun.* **12**, 6136 (2021).
- [44] Á. T. Köpeczi-Bócz, H. Sykora, and D. Takács, Data-driven delay identification with SINDy, in *Advances in Nonlinear Dynamics*, edited by W. Lacarbonara, NODYCON Conference Proceedings Series, Vol. III (Springer, Cham, Switzerland, 2024), pp. 481–491.
- [45] G. Leylaz, S. Wang, and J.-Q. Sun, Identification of nonlinear dynamical systems with time delay, *Int. J. Dyn. Control* **10**, 13 (2022).
- [46] Y. Zhao and S. W. K. Wong, Inference for delay differential equations using manifold-constrained Gaussian processes, *arXiv:2406.15170*.
- [47] U. Kuchler and M. Sørensen, Statistical inference for discrete-time samples from affine stochastic delay differential equations, *Bernoulli* **19**, 409 (2013).
- [48] H. Voss and J. Kurths, Reconstruction of non-linear time delay models from data by the use of optimal transformations, *Phys. Lett. A* **234**, 336 (1997).
- [49] W. Horbelt, J. Timmer, and H. Voss, Parameter estimation in nonlinear delayed feedback systems from noisy data, *Phys. Lett. A* **299**, 513 (2002).
- [50] M. J. Bünner, M. Popp, Th. Meyer, A. Kittel, and J. Parisi, Tool to recover scalar time-delay systems from experimental time series, *Phys. Rev. E* **54**, R3082(R) (1996).
- [51] T. M. Muenker, G. Knotz, M. Krüger, and T. Betz, Accessing activity and viscoelastic properties of artificial and living systems from passive measurement, *Nat. Mater.* **23**, 1283 (2024).
- [52] K. Berg-Sørensen and H. Flyvbjerg, Power spectrum analysis for optical tweezers, *Rev. Sci. Instrum.* **75**, 594 (2004).
- [53] K. Berg-Sørensen, E. J. G. Peterman, T. Weber, C. F. Schmidt, and H. Flyvbjerg, Power spectrum analysis for optical tweezers. II: Laser wavelength dependence of parasitic filtering, and how to achieve high bandwidth, *Rev. Sci. Instrum.* **77**, 063106 (2006).
- [54] V. Sposini, D. Krapf, E. Marinari, R. Sunyer, F. Ritort, F. Taheri, C. Selhuber-Unkel, R. Benelli, M. Weiss, R. Metzler, and G. Oshanin, Towards a robust criterion of anomalous diffusion, *Commun. Phys.* **5**, 305 (2022).
- [55] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, 2nd ed. (Cambridge University Press, Cambridge, 2022).
- [56] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, Deep learning for time series classification: A review, *Data Min. Knowl. Discov.* **33**, 917 (2019).
- [57] U. Kuchler and B. Mensch, Langevins stochastic differential equation extended by a time-delayed term, *Stoch. Stoch. Rep.* **40**, 23 (1992).
- [58] S. Yanchuk and P. Perlikowski, Delay and periodicity, *Phys. Rev. E* **79**, 046221 (2009).
- [59] T. D. Frank and P. J. Beek, Stationary solutions of linear stochastic delay differential equations: Applications to biological systems, *Phys. Rev. E* **64**, 021917 (2001).
- [60] S. Yanchuk, M. Wolfrum, P. Hövel, and E. Schöll, Control of unstable steady states by long delay feedback, *Phys. Rev. E* **74**, 026201 (2006).
- [61] D. Saha, S. Tarama, H. Löwen, and S. U. Egelhaaf, Cybloids—Creation and control of cybernetic colloids, *Soft Matter* **20**, 8112 (2024).
- [62] R. A. Kopp and S. H. L. Klapp, Persistent motion of a Brownian particle subject to repulsive feedback with time delay, *Phys. Rev. E* **107**, 024611 (2023).
- [63] T. Lequy and A. M. Menzel, Stochastic motion under nonlinear friction representing shear thinning, *Phys. Rev. E* **108**, 064606 (2023).
- [64] S. Guillouezic, I. L'Heureux, and A. Longtin, Small delay approximation of stochastic delay differential equations, *Phys. Rev. E* **59**, 3970 (1999).
- [65] S. Guillouezic, I. L'Heureux, and A. Longtin, Rate processes in a delayed, stochastically driven, and overdamped system, *Phys. Rev. E* **61**, 4906 (2000).
- [66] M. S. Bartlett, Periodogram analysis and continuous spectra, *Biometrika* **37**, 1 (1950).
- [67] R. B. Blackman and J. W. Tukey, The measurement of power spectra from the point of view of communications engineering—Part I, *Bell Syst. Tech. J.* **37**, 185 (1958).
- [68] A. Squarcini, A. Solon, and G. Oshanin, Spectral density of individual trajectories of an active Brownian particle, *New J. Phys.* **24**, 013018 (2022).
- [69] D. Krapf, N. Lukat, E. Marinari, R. Metzler, G. Oshanin, C. Selhuber-Unkel, A. Squarcini, L. Stadler, M. Weiss, and X. Xu, Spectral content of a single non-Brownian trajectory, *Phys. Rev. X* **9**, 011019 (2019).
- [70] D. Krapf, E. Marinari, R. Metzler, G. Oshanin, X. Xu, and A. Squarcini, Power spectral density of a single Brownian trajectory: What one can and cannot learn from it, *New J. Phys.* **20**, 023029 (2018).
- [71] R. Metzler, Brownian motion and beyond: First-passage, power spectrum, non-Gaussianity, and anomalous diffusion, *J. Stat. Mech.* (2019) 114003.
- [72] S. A. McKinley and H. D. Nguyen, Anomalous diffusion and the generalized Langevin equation, *SIAM J. Math. Anal.* **50**, 5119 (2018).

- [73] L. He and B. Feng, *Fundamentals of Measurement and Signal Analysis* (Springer, Singapore, 2022).
- [74] P. Welch, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoust.* **15**, 70 (1967).
- [75] J. W. Cooley and J. W. Tukey, An algorithm for the machine calculation of complex Fourier series, *Math. Comp.* **19**, 297 (1965).
- [76] M. L. Rosinberg, T. Munakata, and G. Tarjus, Stochastic thermodynamics of Langevin systems under time-delayed feedback control: Second-law-like inequalities, *Phys. Rev. E* **91**, 042114 (2015).
- [77] M. L. Rosinberg, G. Tarjus, and T. Munakata, Stochastic thermodynamics of Langevin systems under time-delayed feedback control. II. Nonequilibrium steady-state fluctuations, *Phys. Rev. E* **95**, 022123 (2017).
- [78] T. Erneux, *Applied Delay Differential Equations*, Surveys and Tutorials in the Applied Mathematical Sciences (Springer, New York, NY, 2009).
- [79] S. Ciliberto, Experiments in stochastic thermodynamics: Short history and perspectives, *Phys. Rev. X* **7**, 021051 (2017).
- [80] S. Dago, J. Pereda, S. Ciliberto, and L. Bellon, Virtual double-well potential for an underdamped oscillator created by a feedback loop, *J. Stat. Mech.* (2022) 053209.
- [81] S. Toyabe, H.-R. Jiang, T. Nakamura, Y. Murayama, and M. Sano, Experimental test of a new equality: Measuring heat dissipation in an optically driven colloidal system, *Phys. Rev. E* **75**, 011122 (2007).
- [82] M. A. Fernandez-Rodriguez, F. Grillo, L. Alvarez, M. Rathlef, I. Buttinoni, G. Volpe, and L. Isa, Feedback-controlled active Brownian colloids with space-dependent rotational dynamics, *Nat. Commun.* **11**, 4223 (2020).
- [83] Y. Jun and J. Bechhoefer, Virtual potentials for feedback traps, *Phys. Rev. E* **86**, 061106 (2012).
- [84] B. Ertel, J. Van Der Meer, and U. Seifert, Waiting time distributions in hybrid models of motor–bead assays: A concept and tool for inference, *Int. J. Mol. Sci.* **24**, 7610 (2023).
- [85] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics*, edited by M. Cardona, P. Fulde, K. Von Klitzing, H.-J. Queisser, and H. K. V. Lotsch, Springer Series in Solid-State Sciences (Springer, Berlin, Heidelberg, 1991), Vol. 31.
- [86] U. Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani, Fluctuation–dissipation: Response theory in statistical physics, *Phys. Rep.* **461**, 111 (2008).
- [87] M. S. Gomes-Filho, L. C. Lapas, E. Gudowska-Nowak, and F. A. Oliveira, The fluctuation–dissipation relations: Growth, diffusion, and beyond, *Phys. Rep.* **1141**, 1 (2025).
- [88] A. V. Chechkin and R. Klages, Fluctuation relations for anomalous dynamics, *J. Stat. Mech.* (2009) L03002.
- [89] X. Wang, P.-C. Chen, K. Kroy, V. Holubec, and F. Cichos, Spontaneous vortex formation by microswimmers with retarded attractions, *Nat. Commun.* **14**, 56 (2023).
- [90] T. Bäuerle, R. C. Löffler, and C. Bechinger, Formation of stable and responsive collective states in suspensions of active colloids, *Nat. Commun.* **11**, 2547 (2020).
- [91] N. Miller and R. Gerlai, From schooling to shoaling: Patterns of collective motion in zebrafish (*Danio rerio*), *PLoS ONE* **7**, e48865 (2012).
- [92] J. D. Partridge, Surveying a swarm: Experimental techniques to establish and examine bacterial collective motion, *Appl. Environ. Microbiol.* **88**, e01853-21 (2022).
- [93] B. Koger, A. Deshpande, J. T. Kerby, J. M. Graving, B. R. Costelloe, and I. D. Couzin, Quantifying the movement, behaviour and environmental context of group-living animals using drones and computer vision, *J. Anim. Ecol.* **92**, 1357 (2023).
- [94] S. L. Miller and D. Childers, Power spectral density, in *Probability and Random Processes* (Elsevier, Amsterdam, 2012), pp. 429–471.
- [95] C. R. Harris, *et al.*, Array programming with NumPy, *Nature (London)* **585**, 357 (2020).
- [96] P. Virtanen, *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nat. Methods* **17**, 261 (2020).
- [97] J. D. Hunter, Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).
- [98] F. Chollet, *et al.*, Keras, 2015, <https://keras.io>.
- [99] M. Abadi, *et al.*, TENSORFLOW: Large-scale machine learning on heterogeneous systems, 2015, <https://www.tensorflow.org/>.
- [100] G. Maruyama, Continuous Markov processes and stochastic equations, *Rend. Circ. Mat. Palermo* **4**, 48 (1955).
- [101] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Stochastic Modelling and Applied Probability (Springer, Berlin, Heidelberg, 1992).
- [102] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications Book Subtitle: With R Examples*, Springer Texts in Statistics (Springer, Cham, 2025).
- [103] P. Hänggi and P. Jung, Colored noise in dynamical systems, *Adv. Chem. Phys.* **89**, 239 (1994).