OPEN ACCESS
IOP Publishing
Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **57** (2024) 235001 (23pp)
https://doi.org/10.1088/1751-8121/ad4752

Matrix numerical method for probability densities of stochastic delay differential equations

Nils Antary[1,2] **and Viktor Holubec**[3,*]

[1] Institute for Theoretical Physics, University of Leipzig, 04081 Leipzig, Germany
[2] Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, 14473 Potsdam, Germany
[3] Faculty of Mathematics and Physics, Department of Macromolecular Physics, Charles University, V Holešovičkách 2, CZ-180 00 Praha, Czech Republic

E-mail: viktor.holubec@mff.cuni.cz and nantary@protonmail.com

Received 31 December 2023
Accepted for publication 3 May 2024
Published 4 June 2024

Abstract

Stochastic processes with time delay are invaluable for modeling in science and engineering when finite signal transmission and processing speeds can not be neglected. However, they can seldom be treated with sufficient precision analytically if the corresponding stochastic delay differential equations (SDDEs) are nonlinear. This work presents a numerical algorithm for calculating the probability densities of processes described by nonlinear SDDEs. The algorithm is based on Markovian embedding and solves the problem by basic matrix operations. We validate it for a broad class of parameters using exactly solvable linear SDDEs and a cubic SDDE. Besides, we show how to apply the algorithm to calculate transition rates and first passage times for a Brownian particle diffusing in a time-delayed cusp potential.

Keywords: stochastic delay differential equations, probability density, numerical algorithm, numerical methods, transition rates, non-Markovian dynamics

1. Introduction

Consider the stochastic delay differential equation (SDDE)

$$\partial_t x(t) = f[x(t), x(t-\tau)] + \sigma\xi(t) \tag{1}$$

* Author to whom any correspondence should be addressed.

with a general time-delayed force $f[x(t),x(t-\tau)]$ and Gaussian white noise $\sigma\xi(t)$ with intensity σ. Similar equations are often encountered in modeling noisy feedback systems [1]. Examples range from feedback-controlled Brownian particles [2–5] and robots [6–8] over various biological [9–13] and physical [14–16] systems to finance [17]. SDDEs also successfully describe instabilities in car traffic [18] and recently, they have been applied to model other manybody nonequilibrium phenomena with time-delayed interactions such as flocking or swarming [19–27].

Unfortunately, SDDEs are notoriously difficult for analytical treatment if they are not linear [28, 29]. SDDEs with nonlinear forces can generally only be treated approximately, using a linearization, small or large delay approxiamation [30] or by expansions around an equilibrium dynamics [31]. Therefore, the main tools for studying SDDEs are numerical methods.

The stochastic processes described by SDDEs are non-Markovian, and it is impossible to write a closed dynamical equation for the probability density for the stochastic coordinate $x(t)$. For this reason, most available numerical methods focus on direct integration of the SDDE [32–34]. We are aware of only a single method devised to compute the coordinate probability density directly [35]. The suggested algorithm is based on a Gaussian approximation of a short-time propagator. For nonlinear problems, this approximation is valid only for small time steps, and thus, the suggested method becomes computationally expensive for long integration times.

In this work, we propose an alternative numerical algorithm that directly calculates the joint probability density for the system's history during time interval $[t-\tau,t]$ in discrete time and space. The dynamics is within the algorithm described by a master equation of the form $\mathbf{q}_{i+1}=U\mathbf{q}_i$, where the vector $\mathbf{q}_i$ contains the occupation probabilities of the individual states in the history space at discrete time i, and U is a time-independent matrix of corresponding transition probabilities. The presented method thus allows us to compute the time evolution of the joint probability density by multiplying the current state vector $\mathbf{q}_i$ by U and the stationary joint probability density as a solution to the eigenvalue problem $\mathbf{q}_\infty=U\mathbf{q}_\infty$. Furthermore, the transition probability matrix can be easily constructed to capture various boundary conditions in a way described in [36]. Hence, the presented method can also be applied to calculating first passage times and transition rates in bistable systems.

This work is structured into two main parts. In the first part, we present the proposed algorithm. In the second part, we calculate the variances for two linear and one non-linear SDDEs using this algorithm. Then, we apply it to predict the transition rate in a cusp potential. Afterwards, we conclude our findings.

2. Numerical method

2.1. Preliminaries

The dynamical equation for probability density $\rho(x,t)$ corresponding to the SDDE (1) can be written in the form [28]

$$\partial_t\rho\left(x,t|\phi\right)=\left[\frac{\sigma^2}{2}\partial_x^2-\partial_x\bar{f}(x,t|\phi)\right]\rho\left(x,t|\phi\right), \tag{2}$$

where ϕ denotes an initial condition for $x(t)$, i.e. $x(t)=\phi(t)$ for $t\in[-\tau,0]$. The function $\bar{f}$ is defined as the average $\bar{f}(x,t|\phi)=\int \mathrm{d}y f(x,y)\rho_2(y,t-\tau|x,t,\phi)$ over the two-time probability

density $\rho_2(y,t-\tau|x,t,\phi)$ that a trajectory residing at coordinate x at time t was at coordinate y at time $t-\tau$ under the given initial condition ϕ. Since equation (2) provides no recipe for calculating the two-time probability density ρ_2, it does not represent a closed dynamical equation for $p(x,t|\phi)$. In fact, no such a closed dynamical equation is known. Nevertheless, we will show below how equation (2) can be numerically integrated to yield the joint probability density for x during the time interval $[t-\tau,t]$.

Our strategy is based on rewriting equation (1) using the Markovian embedding [30]

$$\begin{aligned} \dot{x}_0(t) &= f[x_0(t),x_n(t)] + \sigma\xi(t), \\ x_i(t) &= x_{i-1}(t-\Delta_t), \quad i=1,\ldots,n, \end{aligned} \tag{3}$$

where $\Delta_t=\tau/n$, $x(t)=x_0(t)$, and the embedding becomes exact in the limit of the infinite number n of auxiliary variables $x_i(t)=x(t-i\Delta_t)$. For a fixed history $\{x_i(t)\}_{i=1}^n$, the first equation in (3) corresponds to the Fokker–Planck equation

$$\partial_t\rho(x,t|x_1,\ldots,x_n) = \left[\frac{\sigma^2}{2}\partial_x^2 - \partial_x f(x,x_n)\right]\rho(x,t|x_1,\ldots,x_n), \tag{4}$$

which can be solved, e.g. using the numerical method described in [36], to yield the probability density $p(x,t+\Delta_t,x_0,\ldots,x_{n-1}|x_0,t,x_1,\ldots,x_n)$ for reaching the coordinate x at time $t+\Delta_t$ with history $x_0,\ldots,x_{n-1}$ from the coordinate x_0 at time t with history $x_1,\ldots,x_n$. The remaining equations in (3) represent a deterministic history update, implying that only transition probabilities between history states $\{x_0,x_1,\ldots,x_n\}$ and $\{x,x_0,\ldots,x_{n-1}\}$ with arbitrary x can be nonzero.

Working in a discrete time and space, the probability densities p yield a finite set of transition probabilities between discrete states with coordinates labeled by possible values of the $(n+1)$tuples $\{x(t),x(t-\Delta_t),\ldots,x(t-n\Delta_t)\}$. Organizing the occupation probabilities of these history states at a discrete time i inside a vector $\boldsymbol{q}_i$ and the corresponding transition probabilities p inside a matrix U, the time evolution of the joint probability density $\rho(x,t;x_1,t-\Delta t;\ldots;x_n,t-\tau)$ can be characterized by the Markovian dynamical equation $\boldsymbol{q}_{i+1}=U\boldsymbol{q}_i$ for a discrete-time random walk on a lattice. In the next subsection, we describe how to use these insights to find the probability density for x in practice.

2.2. Algorithm

As outlined above, we start with discretizing space and time. Assuming we aim to find the probability density for x on an interval $x\in[x_-,x_+]$, we choose the discrete space-step $\Delta_x=(x_+-x_-)/(N_x-1)$ to discretize the interval into N_x discrete space points $x_l=x_-+l\Delta_x$, $l=0,\ldots,N_x-1$. Similarly, we choose the discrete time step $\Delta_t=\tau/N_t$ so that each of the points in the history space $\{x(t),x(t-\Delta_t),\ldots,x(t-N_t\Delta_t)\}$ is described by N_t+1 coordinates.

From now on, we will refer to the discrete spatial coordinates x_l by their labels l. In the discretized time, the coordinates of the system in the history space can be written as vectors $\boldsymbol{s}_i=(l_0,\ldots,l_{N_t})^T$ with elements $s_{i,j}=l_j\equiv l[(i-j)\Delta t]$, $j=0,\ldots,N_t$. The total number of states in the history space is $(N_x)^{N_t+1}$. An example of the system's evolution through the history space is shown in figure 1.

The transition probabilities $p(\boldsymbol{s}_{i+1}|\boldsymbol{s}_i)=p(x,t+\Delta_t,x_0,\ldots,x_{n-1}|x_0,t,x_1,\ldots,x_n)$ between the points in the history space vanish whenever $s_{i+1,j}\neq s_{i,j+1}$, $j=0,\ldots,N_t-1$. The nonzero

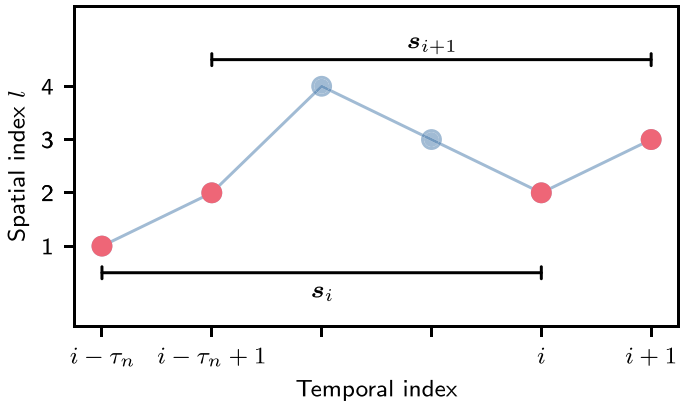

Figure 1. Sketch illustrating two consecutive states in the history space for $N_t = 4$.

values can be calculated from equation (4) using the method of [36]. Denoting by $\boldsymbol{\rho}(t)$ the vector of probabilities $\rho_i(t)$ corresponding to the density $\rho(l,t|l_1,\ldots,l_{N_t})$, the Fokker–Planck equation can be approximately rewritten as the Master equation

$$\partial_t \rho_l(t) = r_{l-1\to l}\rho_{l-1}(t) + r_{l+1\to l}\rho_{l+1}(t) - (r_{l\to l-1} + r_{l\to l+1})\,\rho_l(t) \tag{5}$$

with the transition rates $r_{k\to l} = r_{k\to l}(l_1,\ldots,l_{N_t})$

$$r_{l\to l\pm1} = \frac{\sigma^2}{2\Delta_x^2}\exp\left(\pm\frac{f_{l\to l\pm1}\Delta_x}{\sigma^2}\right), \quad r_{l\pm1\to l} = \frac{\sigma^2}{2\Delta_x^2}\exp\left(\mp\frac{f_{l\pm1\to l}\Delta_x}{\sigma^2}\right), \tag{6}$$

which depend on the history through the forces $f_{k\to l} = f_{k\to l}(l_1,\ldots,l_{N_t})$. In our numerical simulations, we use the form

$$f_{k\to l}(l_1,\ldots,l_{N_t}) = f\left[(x_k + x_l)/2, \left(x_{l_{N_t}} + x_{l_{N_t-1}}\right)/2\right]. \tag{7}$$

This choice is motivated by the fact that we need to integrate the dynamical equation (5) over a single time-step Δ_t when the coordinate changes from x_i to x_j and the relevant history coordinate from $x_{l_{N_t}}$ to $x_{l_{N_t-1}}$. Since we have the whole history during the time interval of length τ available, the presented strategy can, in principle, also capture more complicated memory dependencies of the force in equation (1) than $f(x(t), x(t-\tau))$.

The transition rates (6) are valid in the bulk of the discrete state space. On its boundaries, one can cautiously choose from various boundary conditions [36]. However, we recommend implementing absorbing boundaries even when solving for a stationary density in an infinite domain when one normally introduces a reflecting boundary condition equivalent to an infinite potential barrier far enough from the expected maximum of the probability density. This is because the reflecting boundary conditions in the delayed dynamics lead to oscillations, which would not be observed in the infinite system. When modeling a system with natural boundaries, we thus impose absorbing boundaries far away from the most likely occupied states and normalize the occupation probabilities to one after each iteration in time to correct for the probability leaking through the boundary.

Having all transition rates, one can rewrite the Master equation (5) as $\partial_t \boldsymbol{\rho}(t) = R(l_1,\ldots,l_{N_t})\boldsymbol{\rho}(t)$, and formally integrate it over single time-step. Denoting as $\boldsymbol{\rho}_i$ the probability vector at the discrete time i, we find $\boldsymbol{\rho}_{i+1} = G(l_1,\ldots,l_{N_t})\boldsymbol{\rho}_i$, with the propagator

$$G\left(l_1,\ldots,l_{N_t}\right) = \exp\left[R\left(l_1,\ldots,l_{N_t}\right)\Delta_t\right]. \tag{8}$$

Elements $G_{i,j}(l_1,\ldots,l_{N_t})$ of this matrix are transition probabilities from state $\{i,l_1,\ldots,l_{N_t}\}$ to state $\{j,i,l_1,\ldots,l_{N_t-1}\}$ in the history space, which represent the nonzero transition probabilities $p(\boldsymbol{s}_{i+1}|\boldsymbol{s}_i)$. Altogether, we thus have

$$p\left(\boldsymbol{s}_{i+1}|\boldsymbol{s}_i\right) = \begin{cases} 0 & s_{i+1,j} \neq s_{i,j+1} \text{for any } j \in [0,N_t-1] \\ G_{s_{i,0},s_{i+1,0}}\left(s_{i,1},\ldots,s_{i,N_t}\right) & \text{otherwise} \end{cases}. \tag{9}$$

In the final step, we introduce a vector $\boldsymbol{q}_i$ of length $(N_x)^{N_t+1}$ so that its element q_{i,k_i},

$$k_i = \sum_{j=0}^{N_t} s_{i,j}\left(N_x\right)^j, \tag{10}$$

contains the probability to find the system at discrete time i in the history state $\mathbf{s}_i$[4]. After organizing all the transition probabilities (9) into an $(N_x)^{N_t+1}\times(N_x)^{N_t+1}$ matrix U with elements $U_{k^{i+1},k^i} = p(\mathbf{s}_{i+1}|\mathbf{s}_i)$, the system evolution through the history space can be described by the discrete time Markov process

$$\mathbf{q}_{i+1} = U\mathbf{q}_i. \tag{11}$$

The matrix multiplication takes care of summation over all history states $\boldsymbol{s_i}$, which contribute to the next state $\boldsymbol{s_{i+1}}$. Although the matrix U can become very large, only $N_x^{N_t+2}$ of its $(N_x)^{2(N_t+1)}$ entries are nonzero. While this number can still be very large for denser discretizations, we tested that one can work efficiently with U using functions for sparse matrices implemented in Python [37] or similar.

The dynamical equation (11) can be used for solving the complete time-resolved evolution of the system through the history space from an initial condition $\mathbf{q}_0$ as $\mathbf{q}_i = U^i\mathbf{q}_0$[5]. When $\mathbf{q}_i$ converges to a steady state $\boldsymbol{q}$, the latter can directly be determined by solving the eigenvalue problem $\boldsymbol{q} = U\boldsymbol{q}$. If the absorbing boundaries are imposed to supplement natural boundaries as suggested above, the steady state should be identified with the state corresponding to the eigenvalue of U closest to one from below. In the example section below, we refer to the time-resolved solution as the *dynamical* algorithm and the steady-state solution as the *steady-state* algorithm.

The states $\mathbf{q}_i$ contain the complete history of the coordinate during one delay interval. The one time probability density corresponding to the discretized version $\rho(x_l, i\Delta_t|\mathbf{s}_0)$ of ρ in equation (2) follows from the marginalization

$$\rho\left(l,i\right) = \sum_{k=0}^{(N_x)^{N_t+1}} q_{i,k}\delta_{\lfloor k/N_x^{N_t}\rfloor,l}, \tag{12}$$

[4] The inverse map is $s_{i,j} = (k^i \bmod N_x, \lfloor k^i/N_x\rfloor \bmod N_x, \ldots., \lfloor k^i/N_x^{N_t-1}\rfloor \bmod N_x, \lfloor k^i/N_x^{N_t}\rfloor)^T$, where mod is the remainder of a division, and $\lfloor\rfloor$ denotes the floor operation.

[5] Since U^i is for $i>1$ not sparse anymore, it is preferable to evaluate the product $U^i\mathbf{q}_0$ as i multiplications of the initial vector from the left.

assuming that the discrete process is initialized at time 0 with unit probability in state $\mathbf{s}_0$ and $\lfloor\rfloor$ denotes the floor operation. Other joint probability distributions during the time interval $[t-\tau,t]$ can be calculated similarly. Besides, the jump process in the history space is Markovian, and thus, arbitrary joint probability density can be calculated using the matrix U. The numerical method can also be generalized to calculate characteristic functions of various functionals over the stochastic trajectories of the stochastic process $x(t)$ as described in [36].

3. Numerical examples

To test the proposed method, we apply it to three specific examples. First, we show that it can reproduce known analytical solutions of linear SDDEs. As a second example, we use the method to calculate probability density and variance for a diffusion in a delayed cubic potential. In the last example, we compute probability density and rates for transitions between the wells of a delayed cusp potential.

In all the studied situations, we compare the numerical results to solutions found using Brownian dynamics (BD) simulations of the corresponding SDDEs. For the linear case, the reason is to provide a performance comparison of the two methods. In the simulations, we simultaneously iterated 2000 trajectories using a vectorized version of the Euler algorithm with timestep $dt=\tau/60$. We performed ten such runs in parallel to estimate simulation errors and speed up the computation. For further implementation details, interested readers can directly inspect our codes [38].

In the numerics, we tested the precision of the results for several values of space and time discretization N_x and N_t. Interestingly, we found that in most examples, the precision significantly increases only up to $N_t=2$, and for finer time discretizations, it barely changes. The only exception is the case with a combination of delayed and nondelayed harmonic potential in section 3.1.2, where increasing N_t to three improves the results significantly.

There is a stark difference between the meaning of the time discretization in BD simulations and the numerical algorithm. In simulations, dt is an integration timestep over which one assumes a constant deterministic component of the right-hand side of the SDDE. In the numerics, Δ_t discretizes the history interval $[t-\tau,t]$ and approximates the drift term f in (4) within the individual intervals of duration Δ_t by corresponding average values. With this approximate f, the spatially discretized partial differential equation (5) is solved by the exact matrix exponential. Using a fine space discretization with a large N_x is thus important for obtaining precise results. For the usually optimal time discretization $N_t=2$, the number of nonzero elements of the rate matrix U increases as N_x^4 and the computation time grows accordingly.

3.1. Linear SDDEs

Consider equation (1) with a linear function $f[x(t),x(t-\tau)]=-ax(t)-bx(t-\tau)$. The resulting SDDE

$$\dot{x}(t)=-ax(t)-bx(t-\tau)+\sigma\xi(t) \tag{13}$$

is linear, and its properties, including the probability densities for $x(t)$, are known [30]. In particular, linearity implies that the probability density for $x(t)$, evolved from a Gaussian initial condition, is Gaussian and thus determined solely by its mean and variance. For completeness, we derive these two moments analytically, including their small delay approximations, in appendix A.

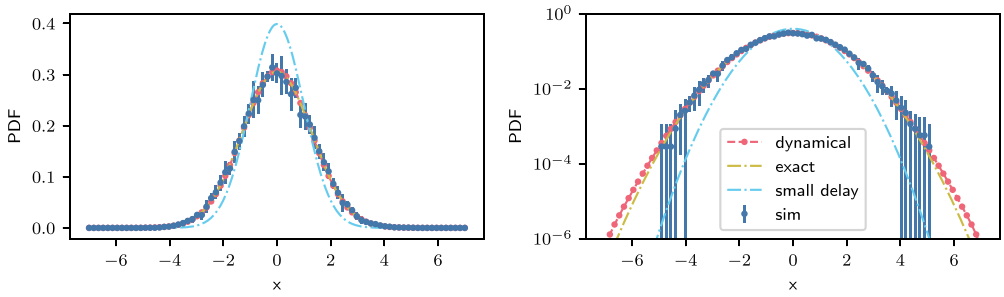

Figure 2. Probability density for $x(t)$ in SDDE (14) for $\tau = 1$ and $t = 10$ in normal (left) and logarithmic (right) scale. We used $N_x = 81$ and $N_t = 2$ in numerics. The errors of the BD simulation were calculated from ten parallel runs.

3.1.1. Time-delayed harmonic potential. We start by studying the system

$$\dot{x}(t) = -x(t-\tau) + \xi(t), \tag{14}$$

with the initial condition $x(t) = 0$ for $t \leqslant 0$. It describes the diffusion of a particle in a time-delayed harmonic potential. Setting $a = 0$ in equation (13), the coefficients b and σ can be absorbed into units of t and x. Equation (14) thus represents a general linear SDDE containing only delayed terms on the right-hand side. Depending on the value of delay, solutions to (14) averaged over the noise exhibit three different dynamical regimes [29]: (I) $\tau < 1/\mathrm{e}$, monotonous exponential decay; (II) $1/\mathrm{e} < \tau < \pi/2$, oscillatory exponential decay; (III) $\pi/2 < \tau$, oscillatory exponential increase.

In figure 2, we demonstrate that the probability density for x obtained numerically (red) and by BD simulation (dark blue) indeed match the exact Gaussian distribution (yellow). This example shows that the numerics can be expected to outperform BD simulations when applied to either calculate tails of probability densities or higher moments of x, whose precise evaluation requires sufficient sampling of rare events and thus many simulation runs. The disagreement between the exact solution and numerics for far tails can be decreased by increasing N_x. We also show the prediction obtained using the small delay approximation in light blue for completeness.

The Gaussian probability density is solely determined by its mean $\langle x \rangle$ and variance $\nu = \langle x^2 \rangle - \langle x \rangle^2$. The former is, for our initial condition, 0, and thus, in figure 3, we show the variance computed as a function of delay time τ. In the left column, we present ν for finite times numerically obtained by the dynamical algorithm, i.e. by iterating equation (11). To depict behavior for multiple finite times in a single plot, we show the time-averaged variance $\nu = \frac{1}{2.5}\int_{7.5}^{10} \nu(t')\mathrm{d}t'$ (data points) together with the minimum and maximum values of $\nu(t)$ during the time interval (7.5, 10) (colored areas). BD simulation results (blue) overlap with exact analytical results (yellow) in all three dynamical regimes. Panels (a), (c) and (e) differ only in space and time discretization used for the numerics (red). They show that increasing N_t beyond two does not improve the precision, while changing N_x from 41 to 81 yields significantly better agreement with exact results.

The right column of figure 3, shows stationary variance $\nu = \lim_{t\to\infty} \nu(t)$ in dynamical regimes I and II, where it is finite. Panels (b), (d) and (f) differ just in the precision of the numerics in the same way as in the left column. The small delay approximation (light blue)

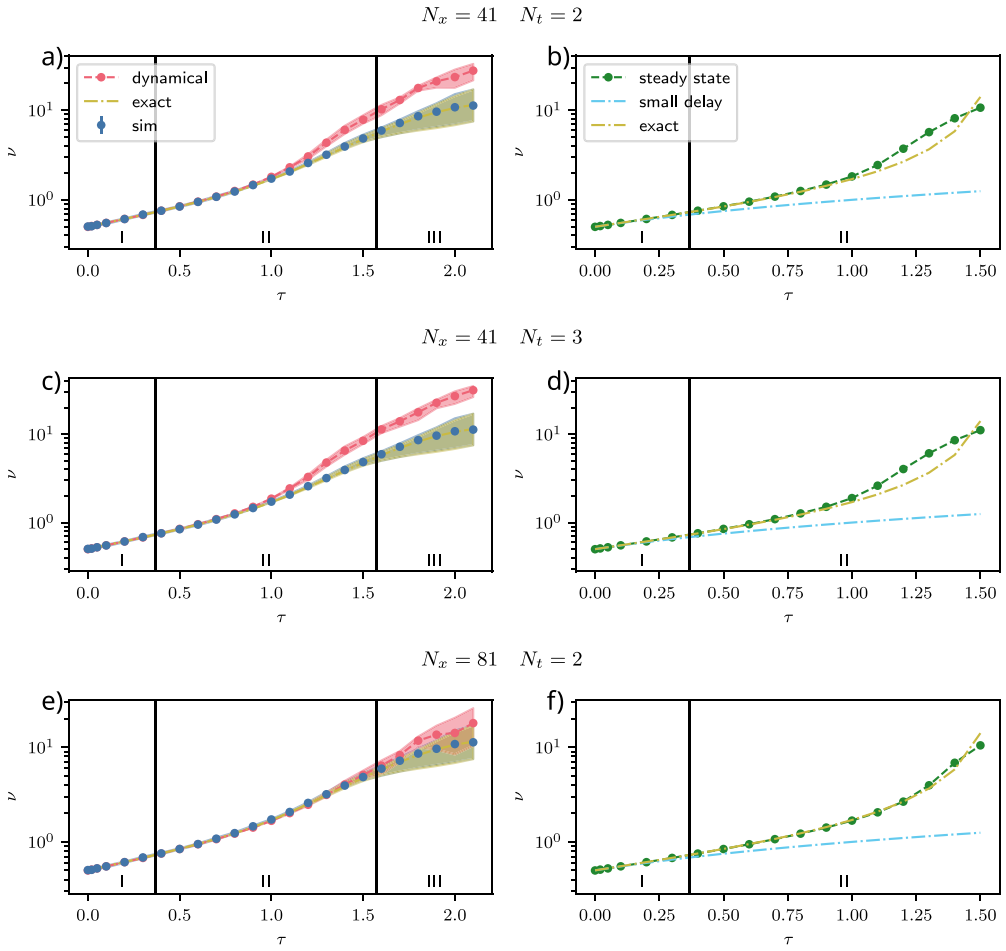

Figure 3. Variance ν for $x(t)$ in equation (14) as a function of delay time τ. In the left panels, the data points show ν calculated as the time average over the interval (7.5, 10) numerically (red), by BD simulation (blue), and by exact analytics (yellow). The colored areas mark the interval's minimal and maximal values of $\nu(t)$. The error bars in simulations were calculated from ten parallel runs and are smaller than the data points. In the right panels, we show stationary variance attained at long times calculated numerically (green), by exact analytics (yellow), and using the small delay approximation (light blue). The numerical solver uses space and time discretization N_t and N_x shown above the individual panels and $x_+ - x_- = 4\sqrt{\nu}$.

predicts the exact variance (yellow) relatively well in the first dynamical regime. The numerical results (green) obtained using the steady state algorithm, i.e. by solving $U\mathbf{q} = \mathbf{q}$, agree best with the analytical results for $N_x = 81$ and $N_t = 2$.

The main reason the error of the numerical results increases in figure 3 with the delay time is the increase of ν with τ. We set the boundaries of the $x_\mp$ interval as $x_+ - x_- = 4\sqrt{\nu}$ so that the probability at the boundaries is negligible. However, for a fixed N_x, the discrete space step $\Delta_x = (x_+ - x_-)/(N_x - 1)$ increases with the delay, leading to decreased precision for the same computation time. Besides, the history's sampling $\Delta_t = \tau/2$ also worsens with increasing delay.

Table 1. Computation times in seconds on an Ubuntu laptop with AMD Ryzen 5 2500U CPU (four 2GHz cores) and 18.4 GiB RAM for numerics with specific discretizations and the simulations. The simulations were performed for ten parallel runs, each with 2000 trajectories. The *Setup* time measures the initialization period of the program before any iteration in time is performed. The *Iteration* time is the duration of one Δ_t update in numerics and dt update in the simulation. The τ *interval* time is the computation time of one delay period. It is given by the Iteration time $\times$ N_t for the numerics and $\times$ 60 for the simulation. The *Steady state overall* time is the runtime of the steady state algorithm for the specific discretization.

	Numerics			Simulation
N_x	41	41	81	
N_t	2	3	2	60
Setup	0.828	5.255	8.218	—
Iteration	0.004	0.147	0.055	0.003
τ *interval*	0.007	0.441	0.11	0.207
Steady state overall	1.316	34.414	15.252	

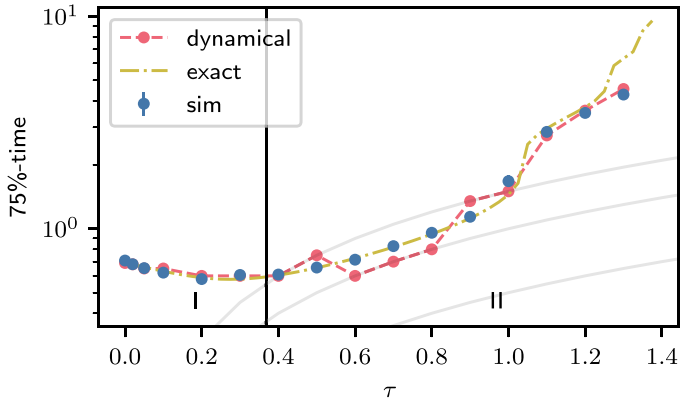

Figure 4. Time when the variance of the probability density for $x(t)$ in (14) reaches 75% of its stationary value as a function of delay obtained from exact analytical expression (yellow), BD simulations (blue), and numerically (red). The grey lines depict from bottom to top the first three iteration times $\Delta_t = \tau/2, 2\Delta_t, 3\Delta_t$ of the numerical algorithm with $N_t = 2$ and $N_x = 81$.

In table 1 we provide run times of the BD simulations, implemented with 10×2000 realization of the stochastic process with $dt = \tau/60$, and the two numerical procedures with different space and time discretizations on our specific hardware described in the caption. The table shows that the overall runtime of numerics with the optimal discretization $N_x = 81$ and $N_t = 2$ in figure 3 is approximately half of the runtime of the simulations. While the simulation is more precise when predicting the variance for large delay times, the numerics is superior when computing probability density for x, which requires much better statistics than simple averaging.

To further assess the ability of the numerics to capture the relaxation dynamics of the system, we show in figure 4 the time when the probability density for $x(t)$ initialized at time 0 with probability one at $x = 0$ reached 75% of the stationary variance. While the time obtained by BD simulation (blue) nicely agrees with the exact analytical results (yellow), the numerics

(red) show nonnegligible discrepancies in the dynamical regime II. These errors are caused by the fact that the numerics use quite large time steps $\Delta_t = \tau/2$ (depicted by grey lines in the figure) for long delays and could be partly decreased by interpolation.

3.1.2. Combination of time-delayed and standard harmonic potentials. Let us now turn to the case

$$\dot{x}(t) = -x(t) - x(t-\tau) + \xi(t) \tag{15}$$

with the initial condition $x(t) = 0$ for $t \leqslant 0$. It follows from equation (13) by setting $a = b = \sigma = 1$ and might be interpreted as a dynamical equation for a particle diffusing in a combination of delayed and standard harmonic potentials. In this case, the small delay approximation gives the same result for coordinate variance ν and thus also the steady-state probability density as the exact solution. For specific expressions, see appendix A.

In figure 5, we present the results for variance in the same way as in figure 3. The left panel thus shows ν averaged over the time interval (7.5, 10), and the right panel depicts the stationary variance. The individual rows differ in the discretization used in the numerics. Contrary to what we observe in other numerical examples above and below, in this case, the best agreement between numerics (red on the left and green on the right), simulation (blue on the left), and the exact result (yellow) is obtained for a modest space discretization $N_x = 41$ and finer time-discretization $N_t = 3$. Even better agreement between numerics and exact results would be obtained for a larger N_x. But this would cost a lot in terms of computation time and working memory.

3.2. Quartic potential

Let us now turn to nonlinear SDDEs, where no exact analytical results are known, and thus, numerical or simulation techniques are the only way of obtaining reliable results for non-negligible delays. Specifically, we consider the SDDE

$$\dot{x}(t) = -x(t-\tau)^3 + \xi(t), \tag{16}$$

which can be interpreted as a diffusion in a delayed quartic potential. As in the case of equation (14), one can consider the potential stiffness and noise intensity in (16) to be absorbed in units of t and x. The analysis for $\xi = 0$ in appendix B reveals, that equation (16) without noise is unstable for the initial conditions $x(t) = 0$ for $t < 0$, and $x(0) = x_0$ with $|x_0| > \sqrt{3/(4\tau)}$. Otherwise, solutions to the deterministic part of equation (16) converge to a steady state at $x = 0$. However, this does not mean the same applies to the case with $\xi(t) \neq 0$. Indeed, the noise can always excite the coordinate above the unstable initial distance from the origin corresponding to the given delay and thus create diverging trajectories. This effect is nonnegligible for an arbitrary nonzero delay, and hence, the dynamics described by SDDE (16) converges to a steady state only for $\tau = 0$.

In appendix B, we show that a small delay approximation of equation (16) yields a well-defined stationary distribution for arbitrary τ. Given the above discussion, the small delay approximation thus captures the dynamics correctly only for $\tau = 0$ and fails qualitatively for any $\tau > 0$.

To avoid divergencies caused by the inevitable instabilities, we analyzed the dynamics numerically and by BD simulations with absorbing boundaries at $x_\pm = \pm 4$. We imposed reflecting boundaries at the same positions for the small delay approximation. This only affects

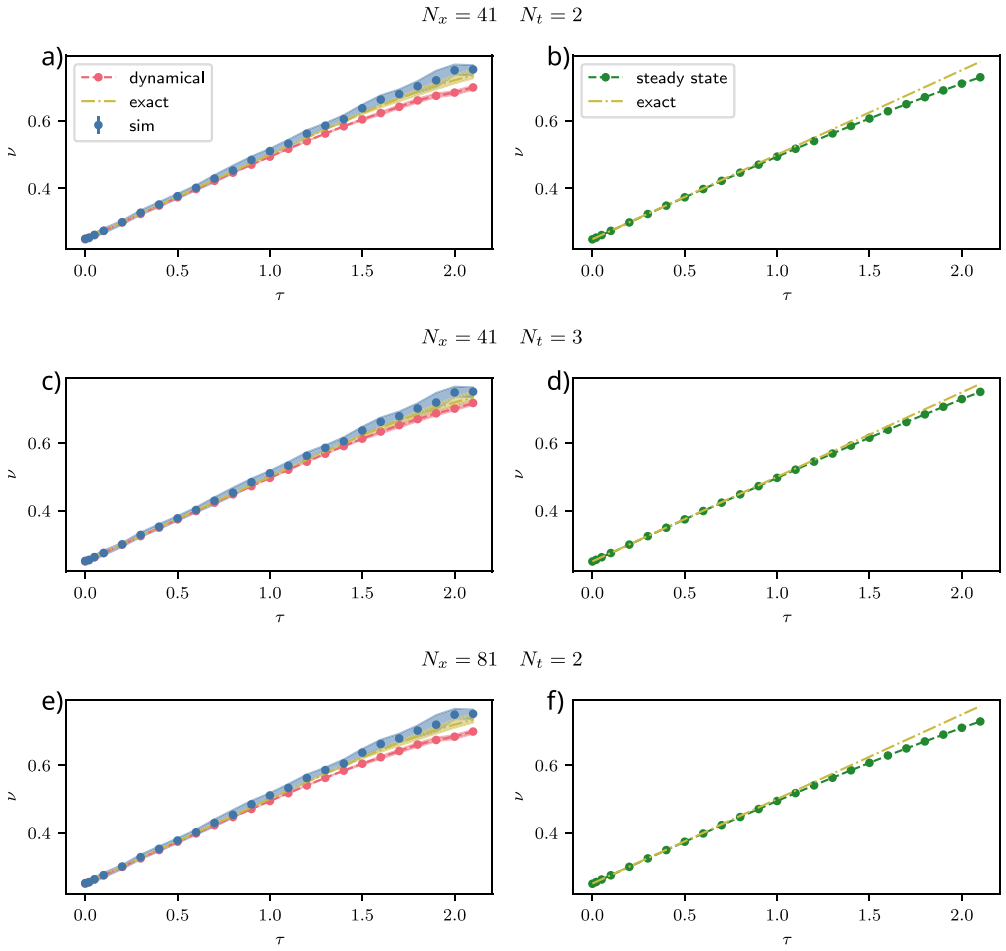

Figure 5. Variance ν for $x(t)$ in equation (15) as a function of delay time τ. In the left panels, the data points show ν calculated as the time average over the interval (7.5, 10) numerically (red), by BD simulation (blue), and by exact analytics (yellow). The colored areas mark the interval's minimal and maximal values of $\nu(t)$. The error bars in simulations were calculated from ten parallel runs and are smaller than the data points. In the right panels, we show stationary variance attained at long times calculated numerically (green), by exact analytics (yellow), and using the small delay approximation (light blue). The numerical solver uses space and time discretization N_t and N_x shown above the individual panels and $x_+ - x_- = 4\sqrt{\nu}$.

the normalization constant N in the stationary density for x in equation (B.5). In figure 6, we show the probability density for the coordinate at time $t = 6$ for delay time $\tau = 0.4$ and initial position $x_0 = 0$ occupied with unit probability at time 0. The numerical result for $N_t = 2$ and $N_x = 81$ (red) agrees within the error bars with the density obtained by BD simulations (dark blue). As expected, the linear approximation fails.

In figure 7, we show the coordinate variance ν after time $t = 4.5$ when the system has already reached a steady state. This is confirmed by the perfect agreement of results obtained by the dynamical (red) and steady-state algorithms (data not shown). The numerical results were obtained for time and space discretization $N_t = 2$ and $N_x = 81$. They nicely agree with

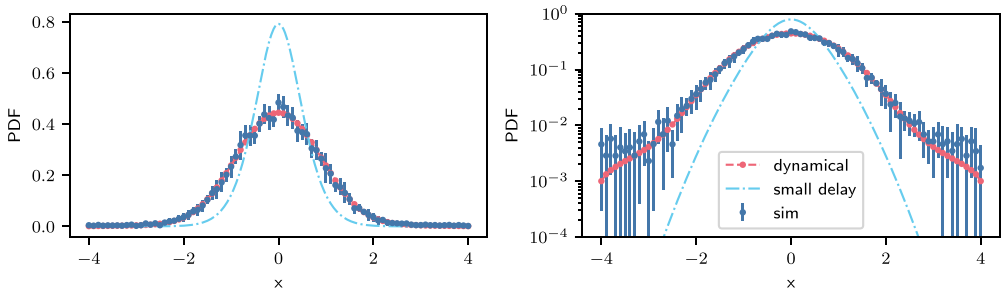

Figure 6. Probability density for $x(t)$ in SDDE (16) for $\tau = 0.4$, $t = 6$, and $x_0 = 0$ in normal (left) and logarithmic (right) scales. In numerics, we used $N_x = 81$ and $N_t = 2$. The errors of the BD simulation were calculated from ten parallel runs.

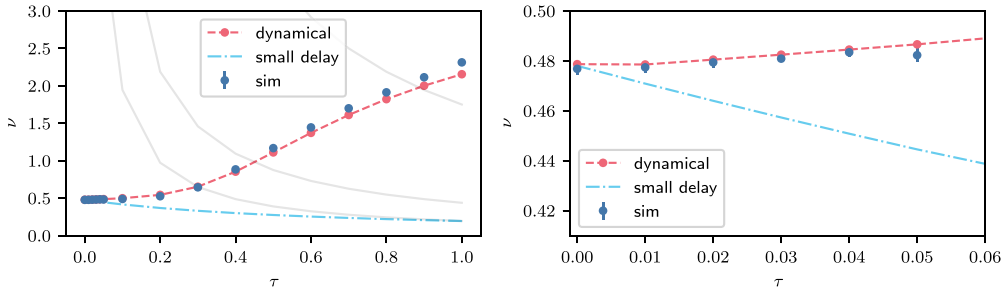

Figure 7. Stationary variance for $x(t)$ in equation (16) with absorbing boundaries at $x = \pm 4$ as function of the delay time τ. The numerical results (dark red) were computed using $N_x = 81$ and $N_t = 2$. The error bars for BD simulation data (blue) were obtained from ten parallel runs using averaging over the time interval $(4.5, 6)$. The grey lines in the left panel indicate when $3\sqrt{\nu}$, $2\sqrt{\nu}$, and $\sqrt{\nu}$ equal the divergence threshold $\sqrt{3/(4\tau)}$. The right panel magnifies the initial part of the left panel to highlight that the small delay approximation (light blue) fails to capture the system dynamics completely, except for $\tau = 0$.

dark blue symbols calculated using BD simulations up to $\tau \approx 0.5$. We have tested that a better agreement is reached upon increasing the spatial discretization N_x. Increasing only the time discretization N_t has a negligible effect. The right panel shows the initial part of the plot in the left panel to illustrate the complete failure of the small delay approximation in light blue. The grey lines in the left panel hint how many trajectories reach the unstable distance from the origin $\sqrt{3/(4\tau)}$ for the given delay. Specifically, when ν is larger than the nth grey line from the left, the unstable point lies within $(4-n)\sqrt{\nu}$ interval of the non-Gaussian density for x.

3.3. *Transition rates for delayed cusp potential*

As the last numerical example, we consider the problem of transitions between the two minima of delayed double-well cusp potential $U(x) = \frac{1}{2}(x - \operatorname{sign} x)^2$, where $\operatorname{sign} x = |x|/x$ denotes the signum function. By choosing appropriate length and time units and neglecting the δ-function obtained by taking the derivative of the signum function, the general SDDE for the dynamics in a delayed cusp potential reads:

$$\dot{x}(t) = -\left[x(t-\tau) - \operatorname{sign} x(t-\tau)\right] + \sigma\xi(t)\,. \tag{17}$$

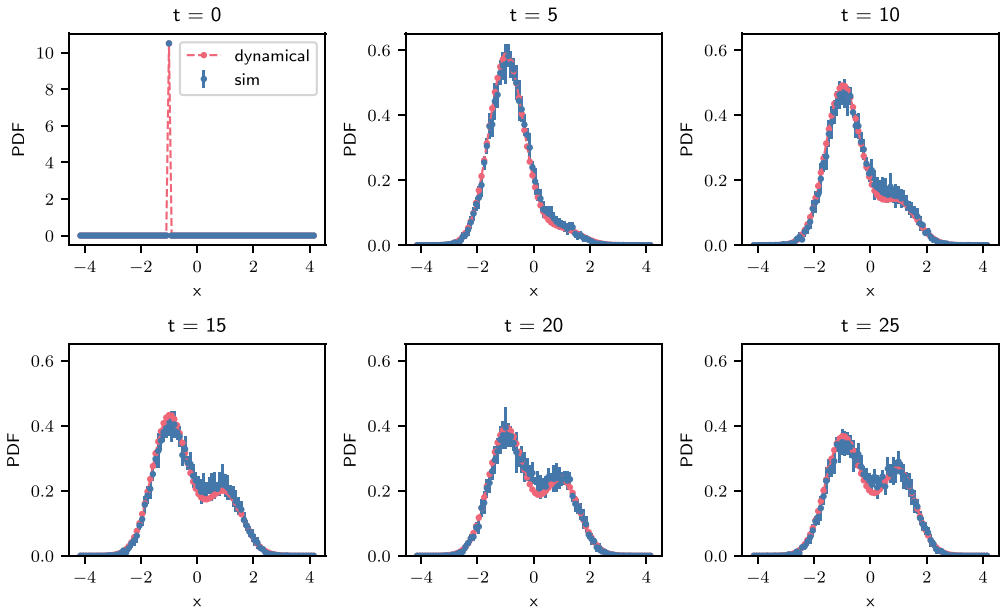

Figure 8. Probability density for $x(t)$ in SDDE (17) with $\tau = 1$ and $\sigma = 0.5$ for times $t = 0, 5 \ldots 25$. The numerical results (red) obtained with $N_x = 88$ and $N_t = 2$ and absorbing boundaries at $x_{\mp} = \mp 4.14$ at all times nicely agree with probability densities calculated from BD simulations with natural boundaries (blue). In the simulations, the error bars were obtained from ten parallel runs.

Due to the signum function on the right-hand side, this SDDE is nonlinear, and we are unaware of any exact solutions for the corresponding dynamics or transition rates. However, relatively good estimates for transition rates were obtained for small and intermediate delays using a time-dependent Fokker–Planck equation valid for linear SDDEs describing the dynamics in the individual wells [29]. In appendix C, we derive the small delay approximation and review approximations for transition rates based on the exact stationary solution of the linear delayed dynamics in one of the wells.

Before focusing on the transition rates, we study the dynamics of the probability density. In figure 8, we initialize $x(t)$ with unit probability in the left well ($x(0) = -1$) with history $x(t) = -1$ for $t < 0$ and depict the densities for $\tau = 1$ and $\sigma = 0.5$ obtained numerically (red) and by BD simulations (blue) for times $t = 0, 5, 10, 15, 20$, and 25. In the numerics, we imposed absorbing boundaries at $x_{\mp} = \mp 4.14$ and used the space and time discretization with $N_x = 88$ and $N_t = 2$. The corresponding numerical results nicely match the densities from simulations performed with natural boundary conditions at all the times.

For a Markovian diffusion in a bistable potential, the transition rates can be either calculated from the first passage time for reaching the top of the barrier between the two wells or as the number of jumps between the wells per unit time. Since these two approaches are not guaranteed to yield the same results for non-Markovian dynamics, we calculate the rates using both.

Let us assume that we initialize the system with unit probability at the bottom of the left well, as in figure 8. In the case of absorbing boundary at $x = 0$, the transition rate κ can then be defined as $\dot{p}_l(t) = -\kappa p_l(t)$, which leads to

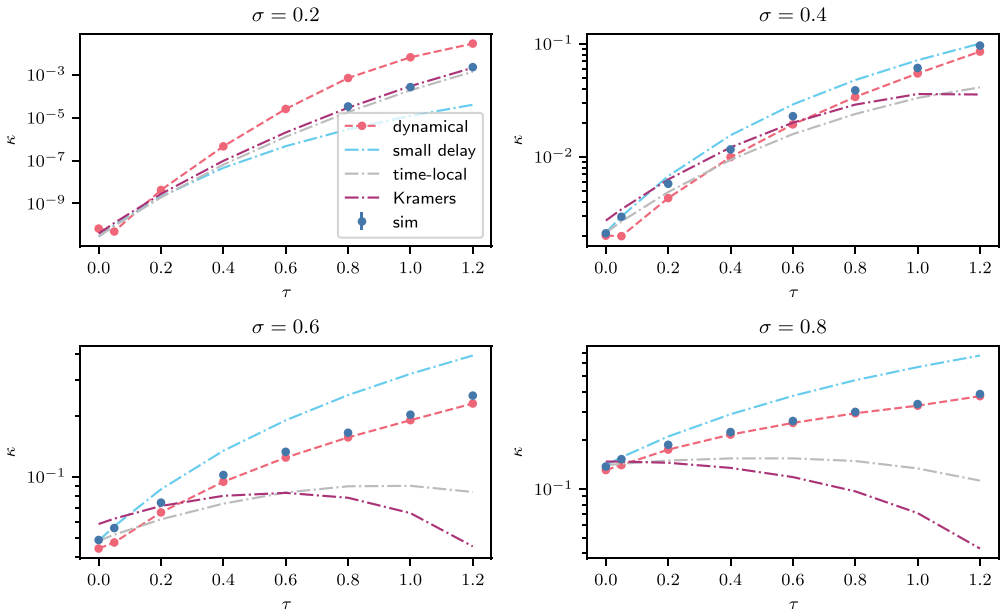

Figure 9. The transition rate (18) for the cusp potential calculated from SDDE (17) with absorbing boundary at the origin as a function of delay time τ. The error bars in simulation data (blue) were computed from ten parallel runs and are always smaller than the symbols. The dynamical numerical algorithm used to produce the red data points was implemented with $\Delta_t = \tau/2$, $N_x = 44$, $x_- = -4.14$, and $x_+ = -0.0487$. The strategies used to obtain the small delay (light blue), time-local (grey), and Kramers (dark red) predictions for the rates are described in appendix C.

$$\kappa = -\frac{\mathrm{d}}{\mathrm{d}t}\log p_l(t)\,, \tag{18}$$

where $p_l(t) = \int_{-\infty}^{0} \mathrm{d}x\rho(x,t)$ is the probability of finding the system in the left well at time t.

For an overdamped Markovian dynamics, the particle reaching the top of the barrier will end up in either well with probability 1/2. Therefore, the rate computed using an absorbing boundary at the top of the barrier (18) is twice the real transition rate between the two wells obtained from the number of jumps between them. For an underdamped Markovian dynamics, the particle reaching the top of the barrier from the left almost surely ends up in the right well, and the two rates equal. For non-Markovian dynamics such as the one in SDDE (17), no such a relation is known and the only safe way to calculate the true rate of transition between the two wells of the cusp potential is thus to consider the dynamics without the absorbing boundary at $x = 0$ shown in figure 8, and define the rate κ as $\dot{p}_l(t) = -\kappa[p_l(t) - p_r(t)]$, where $p_r(t) = \int_0^{\infty} \mathrm{d}x\rho(x,t) = 1 - p_l(t)$. The transition rate thus reads

$$\kappa = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\log\left(2p_l(t) - 1\right). \tag{19}$$

In figures 9 and 10 we show the rates (18) and (19) calculated numerically (red), by BD simulations (blue), small delay approximation (light blue), a Kramers theory (dark red), and a time-local approximation (C.5) to the dynamical equation for the probability density (2), which represents a numerical refinement of the Kramers theory (grey). The details of the last three

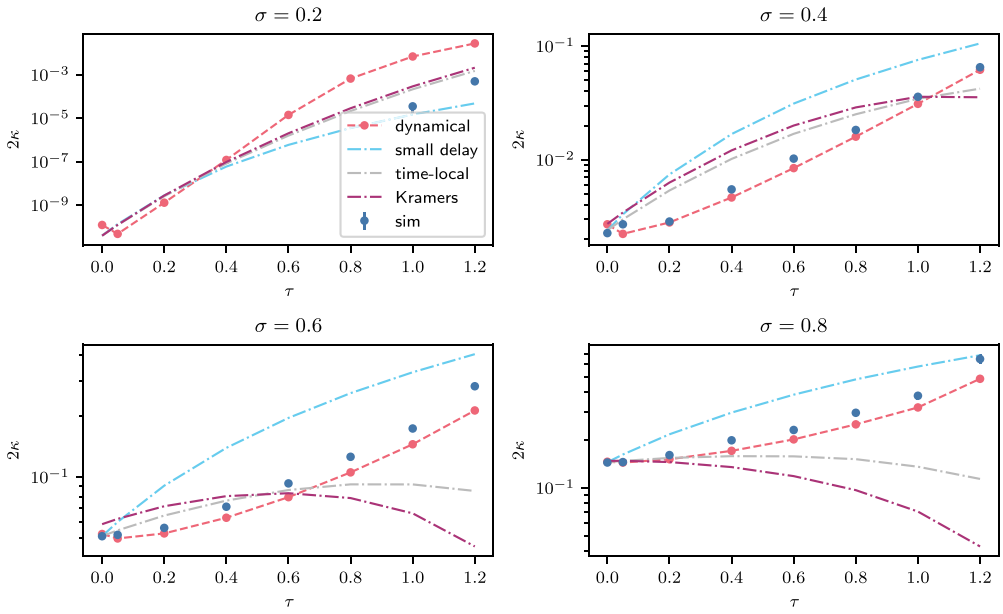

Figure 10. Twice the transition rate (19) for the cusp potential calculated from SDDE (17) with natural boundaries as a function of delay time τ. The error bars in simulation data (blue) were computed from ten parallel runs and are always smaller than the symbols. The dynamical numerical algorithm used to produce the red data points was implemented with $\Delta_t = \tau/2$, $N_x = 88$, $x_\mp = \mp 4.14$. The strategies used to obtain the small delay (light blue) and time-local (grey) are described in appendix C. The shown lines for the Kramers rates times two are the same as in figure 9.

procedures, which we implemented to compare their performance to the numerics, are given in appendix C. In figure 11, we further directly compare transition rates (18) obtained using the absorbing boundary and the true rates (19) for simulation, numerics, and the time-local theory. In all these figures, the transition rates are calculated in steady-state, which we identify as a state where the variance of the probability density for $x(t)$ reaches 75% of its stationary value in a single infinite well (cf figure 4). To avoid numerical overflows, we excluded all data points where arguments in the logarithms were too small.

The numerical algorithm predicts the rates from BD simulations obtained using the two definitions very well for all investigated parameter regimes except for the low noise regime with $\sigma = 0.2$. There, its failure is caused by strong localization of the coordinate distribution near the minimums of the potential. To predict the transition rates accurately in this regime, one needs to increase the spatial discretization from $\Delta_x = 0.08$ employed in our calculations. The missing data points for BD simulation for $\sigma = 0.2$ and $\tau < 0.8$ demonstrate that we have not observed a single transition in our 20 000 runs. While this can be improved by running more and longer simulation runs, it shows the power of the other techniques to determine the transition rates for a weak noise.

For the transition rates calculated using the absorbing boundary in figure 9 for $\sigma = 0.2$, the Kramers approximation and its more precise variant, denoted as time-local, work exceptionally well. None of the applied methods except the numerics yield satisfactory agreement with the simulation for any larger noise value.

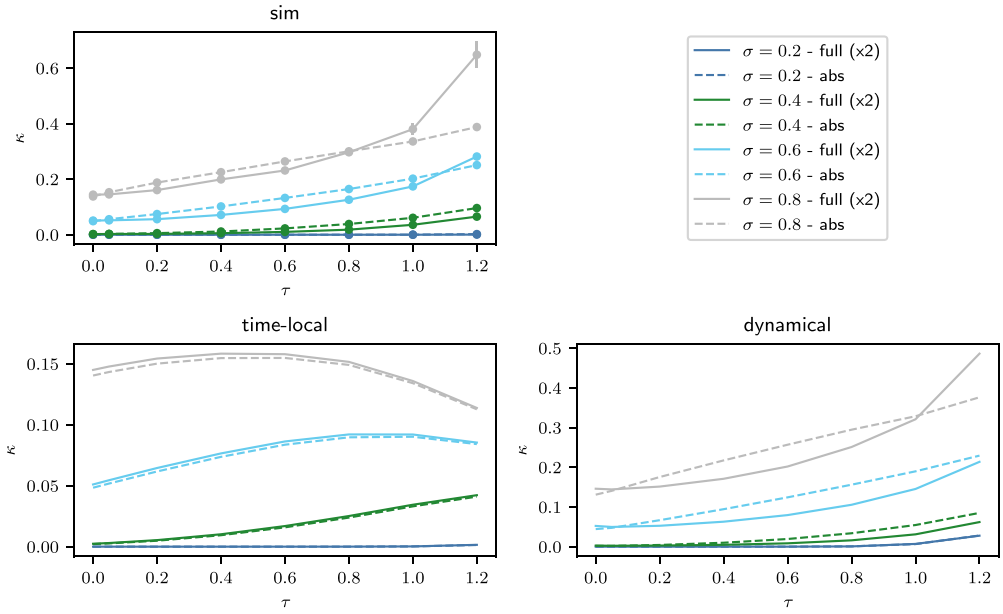

Figure 11. Comparrison between rates (18) calculated using absorbing boundary (abs) at the top of the barrier and twice the true rates (19) calculated from diffusion in the full delayed potential (full). The shown data from BD simulations (sim), time-local approach (time-local), and numerics (dynamical) are the same as in figures 9 and 10.

The transition rates calculated from hopping between the two wells (19), shown in figure 10, are predicted well only using the numerics (except for the lowest noise, as discussed above). We attribute this failure of standard techniques for rate evaluation to the non-Markovianity of the underlying dynamics [29, 39]. Most of the commonly used methods for the calculation of transition rates, except for BD simulations and the present numerics, infer the true transition rate (19) from the rate (18) obtained using absorbing boundary. And the non-Markovianity breaks this correspondence.

In figure 11, we show as dashed lines the transition rates calculated with absorbing boundary (18) and as solid lines twice the real transition rates (19). For the time-local algorithm, when both rates are computed from an overdamped Markovin Fokker–Planck equation, both lines overlap as expected. The disagreement between the two lines for numerics and BD simulations increases with noise. These results highlight the importance of calculating the rates for non-Markovian dynamics from the correct definition (19). On the downside, assessing the rate by numerically resolving the dynamics in the complete double well potential necessitates significantly more bins and, consequently, more computation time than calculating the rate using an absorbing boundary.

4. Conclusion

We presented a novel numerical algorithm that directly calculates probability density for stochastic coordinates described by SDDEs. We employed Markov embedding to derive a closed Fokker–Planck equation for the probability density over an enlarged state space. Its

microstates represent possible histories of the system dynamics over one delay interval in discrete time. By further discretizing space, we have written the dynamical equation for occupation probabilities in the history space in matrix form. This allows us to evolve the probability density in the history space by matrix multiplication or to calculate the stationary distribution directly from the corresponding eigenvalue problem. The probability density for the coordinate can be obtained by marginalizing the probability in the history space. By employing different marginalizations, the method can also be used for computing correlation functions or multi-time probability densities. The numerical method can be easily implemented with various boundary conditions, as we demonstrated by computing stationary densities and transition rates.

The computational demands of the numerics are proportional to $N_x^{N_t+2}$, where N_t+1 is the number of discrete times per one delay interval and N_x is the number of discrete points in space. Fortunately, for most numerical examples we investigated, precise results have already been obtained for $N_t = 2$. The only exception where increasing N_t to three improved the precision was the case of overdamped diffusion in a combination of delayed and time local harmonic potential. The main strength of the numerical method is that it provides precise results in tasks such as calculating transition rates or probability densities, where one has to generate many stochastic trajectories in simulations to get sufficient statistics.

The method is now in its infancy, and we expect it to be sped up compared to our implementations, available online [38]. Besides, the algorithm can be generalized to yield characteristic functions for various functionals over the stochastic process, as suggested for overdamped Markov processes in [36]. This generalization would make it a valuable tool in assessing stochastic thermodynamics of stochastic processes with time delay [30].

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.10406336.

Acknowledgments

We thank Klaus Kroy for inspiring discussions. V H gratefully acknowledges funding by Charles University through Project PRIMUS/22/SCI/009.

Appendix A. Linear stochastic differential eqautions

Linear stochastic differential equations of the form

$$\dot{x}(t) = -ax(t) - bx(t-\tau) + \sigma\xi(t) \tag{A.1}$$

with nonnegative parameters a, b, σ and τ and Gaussian white noise $\xi(t)$ are most easily formally solved by the method of Laplace Transform [29, 40]. The formal solution can then be used to calculate the moments $\langle x\rangle$ and $\langle x^2\rangle$ characterizing the Gaussian probability density for x.

We provide below the solution for the special initial history $x(t) = 0$ for $t < 0$. The Laplace transform of equation (A.1) reads

$$s\tilde{x}(s) - x(0) = -a\tilde{x}(s) - be^{-s\tau}\tilde{x}(s) + \sigma\tilde{\xi}(s), \tag{A.2}$$

where s denotes the Laplace variable and $\tilde{f}(s)$ denotes the transformed function $f(x)$. Solving this equation for $\tilde{x}(s)$ yields

$$\tilde{x}(s) = \frac{x(0) + \sigma\tilde{\xi}(s)}{s + a + b\mathrm{e}^{-s\tau}} = \sum_{l=0}^{\infty} \left(x(0) + \sigma\tilde{\xi}(s)\right) \frac{(-b)^l}{(s+a)^{l+1}} \mathrm{e}^{-\tau s l}. \tag{A.3}$$

Taking the inverse Laplace transform, one obtains

$$x(t) = (x(0)\,\delta(t) + \sigma\xi(t)) * \left[\sum_{l=0}^{\infty} \frac{(-b)^l}{l!} (t - l\tau)^l \mathrm{e}^{-a(t-l\tau)} \Theta(t - l\tau)\right], \tag{A.4}$$

where $*$ denotes a convolution. Defining the Green's function

$$\lambda(t) = \left[\sum_{l=0}^{\infty} \frac{(-b)^l}{l!} (t - l\tau)^l \mathrm{e}^{-a(t-l\tau)} \Theta(t - l\tau)\right] \tag{A.5}$$

allows to write the formal solution as

$$x(t) = \lambda(t)x(0) + \int_0^t \mathrm{d}t' \lambda(t - t')\,\sigma\xi(t'). \tag{A.6}$$

The formal solution, together with the noise autocorrelation $\langle \xi(t)\xi(t')\rangle = \delta(t - t')$ yields the mean and the variance

$$\mu(t) = \langle x(t)\rangle = x(0)\,\lambda(t), \tag{A.7}$$

$$\nu(t) = \langle x(t)^2\rangle - \langle x(t)\rangle^2 = \sigma^2 \int_0^t \mathrm{d}t' \lambda(t')^2. \tag{A.8}$$

If the limit $\lim_{t\to\infty} \nu(t)$ is finite, the probability density for x converges to a stationary Gaussian with zero mean and the variance [22]

$$\nu_{\mathrm{st}} = \frac{\sigma^2}{2} \left[1 + \frac{b \sin\left(\sqrt{b^2 - a^2}\tau\right)}{\sqrt{b^2 - a^2}}\right] \frac{1}{a + b\cos\left(\sqrt{b^2 - a^2}\tau\right)}. \tag{A.9}$$

For $a = 0$ (only delayed coordinate on the right hand side of (A.1)), the stationary variance simplifies to

$$\nu_{\mathrm{st}} = \frac{\sigma^2}{2} [1 + \sin(b\tau)] \frac{1}{b\cos(b\tau)}. \tag{A.10}$$

This expression diverges for $\tau = \pi/(2b)$. For $\tau \geq \pi/(2b)$, the steady state does not exist, and the variance increases with time forever. For $a < b$ (delayed component stronger than the non-delayed one), ν_{st} in (A.9) diverges for $\tau = \arccos(-a/b)/\sqrt{b^2 - a^2}$. For the special case $a = b$ when delayed and non-delayed components have equal strength, one obtains

$$\nu_{\mathrm{st}} = \frac{\sigma^2}{4} \left[\frac{1}{b} + \tau\right], \tag{A.11}$$

which is finite for any b and τ, and thus the system always reaches a steady state. Finally, when $a > b$ (delayed component weaker than the non-delayed one),

$$\nu_{\rm st} = \frac{\sigma^2}{2}\left[1 + \frac{b\sinh\left(\sqrt{a^2-b^2}\tau\right)}{\sqrt{a^2-b^2}}\right]\frac{1}{a + b\cosh\left(\sqrt{a^2-b^2}\tau\right)}, \tag{A.12}$$

which is also always finite.

Let us close this section with a summary of results for linear SDDEs obtained using small delay approximation [28], one of the most commonly applied techniques for treating stochastic delayed differential equations [30]. Either expanding the complete solutions given above to the leading order in the delay, or solving the SDDE (A.1) after using the approximation $x(t-\tau) \approx x(t) - \dot{x}(t)\tau$, one obtains

$$\mu(t) \approx x_0 \exp\left((a+b)(1+\tau b)t\right) \tag{A.13}$$

$$\nu(t) \approx \frac{\sigma^2(1+\tau b)}{2(a+b)}\left(1 - \exp\left(2(a+b)(1+\tau b)t\right)\right). \tag{A.14}$$

The corresponding stationary variance thus reads

$$\nu_{\rm st} \approx \frac{\sigma^2(1+\tau b)}{2(a+b)}. \tag{A.15}$$

Appendix B. Cubic stochastic differential equation

Consider the nonlinear SDDE

$$\dot{x}(t) = -x(t-\tau)^3 + \xi(t), \tag{B.1}$$

with initial condition $x(t) = 0$ for $t < 0$ and $x(0) = x_0$.

First, we investigate the stability of solutions to the deterministic part of this equation for given delay τ and x_0. This can be done by numerically integrating equation (B.1) with $\xi(t) = 0$ and tracking if the solution for given parameters diverges. Alternatively, one can estimate the region of stable parameters by integrating the delay differential equation analytically over the first few delay intervals.

Integrating the deterministic part of equation (B.1) with initial condition $x(t) = 0$ for $t < 0$ and $x(0) = x_0$ during the first three delay intervals, we find

$$x(0 < t \leqslant \tau) = x_0 \tag{B.2}$$

$$x(\tau < t \leqslant 2\tau) = x_0 - x_0^3 {\rm d}_\tau \tag{B.3}$$

$$x(2\tau < t \leqslant 3\tau) = x_0 - \tau x_0^3 - x_0^3 {\rm d}_{2\tau} + \frac{3}{2}x_0^5 {\rm d}_{2\tau}^2 - x_0^7 {\rm d}_{2\tau}^3 + \frac{1}{4}x_0^9 {\rm d}_{2\tau}^4, \tag{B.4}$$

with ${\rm d}_{n\tau} = t - n\tau$. For large enough τ, the coordinate changes sign after one delay interval, suggesting as a suitable condition for instability that the maximum absolute value of the coordinate within the third interval exceeds the absolute value of the initial coordinate, $|\max_{t\in[2\tau,3\tau]} x(t)| \geq |x_0|$. This leads to the condition $\frac{3}{4}x_0 - \tau x_0^3 > x_0$, which suggests that for initial coordinates $|x_0| > \sqrt{3/(4\tau)}$ equation (B.1) yields diverging trajectories. In figure B1,

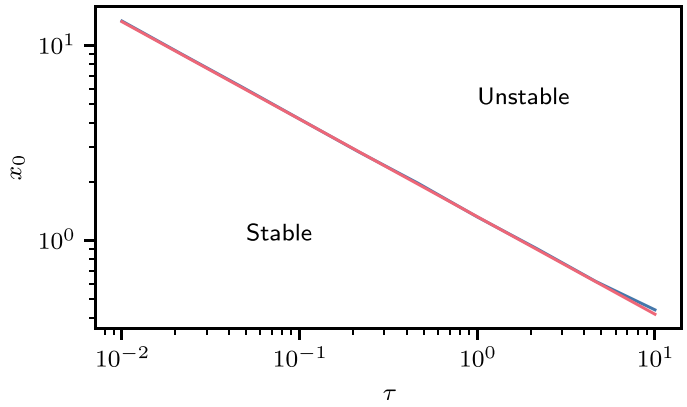

Figure B1. Initial conditions x_0 leading to stable (bounded) and unstable (diverging) solutions of the delay differential equation $\dot{x}(t) = -x(t-\tau)^3$ for different values of the delay τ. The theoretical prediction for the phase boundary (red) almost perfectly overlaps with the result from numerical analysis (blue).

we compare this analytical estimate to results from numerical analysis, finding an excellent agreement.

Let us now turn to the small delay approximation of solutions to equation (B.1) [28]. Expanding the nonlinear delay term on the right-hand side as $x^3(t-\tau) \approx x^3(t) - 3x^2(t)\dot{x}(t)\tau$, and rearranging the resulting stochastic differential equation as an equation for $\dot{x}(t)$, one finds a Langevin equation with drift $f(x) = -x(t)^3(1+3\tau x(t)^2)$ and diffusion $D(x) = (1+3\tau x(t)^2)^2/2$ [41]. The corresponding stationary distribution for the coordinate reads

$$\rho_{\rm st}(x) = \frac{N}{2D(x)} \exp\left(\int \mathrm{d}x \frac{f(x)}{D(x)}\right) = \frac{N}{2D(x)} \exp\left(-\frac{3\tau x^2 - \log\left(3\tau x^2+1\right)}{9\tau^2}\right), \tag{B.5}$$

with a normalization constant N.

Appendix C. Diffusion in delayed cusp potential

Consider the nonlinear SDDE

$$\dot{x}(t) = -\left[x(t-\tau) - \mathrm{sign}\, x(t-\tau)\right] + \sigma\xi(t). \tag{C.1}$$

First, we derive the corresponding small delay approximation used to calculate the rates denoted as *small delay* in the main text. Expanding the right-hand side of (C.1) up to first order in τ and neglecting the contribution proportional to the δ-function, one obtains

$$\dot{x}(t) = -(1+\tau)\left[x(t) - \mathrm{sign}\, x(t)\right] + (1+\tau)\sigma\xi(t). \tag{C.2}$$

The corresponding Fokker–Planck equation reads

$$\partial_t \rho(x,t) = (1+\tau)\,\partial_x\left[x - \mathrm{sign}\, x\right]\rho(x,t) + (1+\tau)^2 \frac{\sigma^2}{2}\partial_x^2 \rho(x,t). \tag{C.3}$$

From this equation, we calculated transition rates, denoted as *small delay* in figures 9 and 10, numerically [29] using equations (18) and (19).

Now, we explain how we calculated the transition rates denoted in the main text as *time-local*. Let us assume that the two wells of the potential are so deep that, for a given noise intensity σ, the coordinate x equilibrates in a given well before any transition occurs. The stationary variance (A.10) for x diffusing in a single well reads $\nu = \frac{\sigma^2}{2}\frac{1+\sin\tau}{\cos\tau}$. For a Markovian diffusion in a harmonic potential, the variance is given by the ratio ω/D where $D = \sigma^2/2$ is the diffusion coefficient and ω is the product of the friction coefficient and the trap stiffness. From equation (2) and the discussion in [28] it follows that the delay does not affect the diffusion coefficient. Therefore, the stationary distribution for diffusion in a delayed single well obeys the Fokker–Planck equation with drift $\partial_x\omega(x\pm 1)\rho(x,t)$,

$$\omega = \frac{\cos(\tau)}{1+\sin(\tau)}, \tag{C.4}$$

and diffusion coefficient D. For deep cusp potentials, it seems reasonable to approximate the true dynamics by connecting two wells with the above described dynamics at $x = 0$. This leads to the Fokker–Planck equation

$$\partial_t\rho(x,t) = \frac{\cos(\tau)}{1+\sin(\tau)}\partial_x(x-\mathrm{sign}x)\,\rho(x,t) + \frac{\sigma^2}{2}\partial_x^2\rho(x,t). \tag{C.5}$$

From this equation, we calculated transition rates, denoted as *time-local* in figures 9 and 10, numerically [29] using equations (18) and (19). Since the Fokker–Planck equation describes overdamped diffusion in a cusp potential, the rate (18) equals twice the true rate (19).

An approximate analytical expression for transition rate in the overdamped setup (C.5) can be derived using Kramer's theory [42]. It reads [29]:

$$\kappa_k = \frac{\omega^{3/2}}{\sqrt{\pi\sigma^2}}\exp\left(-\frac{\omega}{\sigma^2}\right). \tag{C.6}$$

In figures 9–11, we denote the corresponding data as *Kramers*. These rates are analytical approximations of the *time-local* rates.

ORCID iDs

Nils Antary https://orcid.org/0009-0004-7593-2141
Viktor Holubec https://orcid.org/0000-0002-6576-1316

References

[1] Bechhoefer J 2005 Feedback for physicists: a tutorial essay on control *Rev. Mod. Phys.* **77** 783
[2] Khadka U, Holubec V, Yang H and Cichos F 2018 Active particles bound by information flows *Nat. Commun.* **9** 1–9
[3] Fernandez-Rodriguez M A *et al* 2020 Feedback-controlled active Brownian colloids with space-dependent rotational dynamics *Nat. Commun.* **11** 1–10
[4] Muiños-Landin S, Fischer A, Holubec V and Cichos F 2021 Reinforcement learning with artificial microswimmers *Sci. Robot.* **6** eabd9285
[5] Wang X, Chen P-C, Kroy K, Holubec V and Cichos F 2023 Spontaneous vortex formation by microswimmers with retarded attractions *Nat. Commun.* **14** 56
[6] Virágh C *et al* 2014 Flocking algorithm for autonomous flying robots *Bioinspir. Biomim.* **9** 025012
[7] Mijalkov M, McDaniel A, Wehr J and Volpe G 2016 Engineering sensorial delay to control phototaxis and emergent collective behaviors *Phys. Rev.* X **6** 011008

[8] Vásárhelyi G *et al* 2018 Optimized flocking of autonomous drones in confined environments *Sci. Robot.* **3** eaat3536
[9] Mackey M C and Glass L 1977 Oscillation and chaos in physiological control systems *Science* **197** 287–9
[10] Vasilakos K and Beuter A 1993 Effects of noise on a delayed visual feedback system *J. Theor. Biol.* **165** 389–407
[11] Frank T D and Beek P J 2001 Stationary solutions of linear stochastic delay differential equations: applications to biological systems *Phys. Rev.* E **64** 021917
[12] Tian T, Burrage K, Burrage P M and Carletti M 2007 Stochastic delay differential equations for genetic regulatory networks *J. Comput. Appl. Math.* **205** 696–707
[13] Milton J *et al* 2009 The time-delayed inverted pendulum: implications for human balance control *Chaos* **19** 026110
[14] García-Ojalvo J and Roy R 1996 Noise amplification in a stochastic Ikeda model *Phys. Lett.* A **224** 51–56
[15] Milton J G 2015 Time delays and the control of biological systems: an overview *IFAC-PapersOnLine* **48** 87–92
[16] Keane A, Krauskopf B and Postlethwaite C M 2017 Climate models with delay differential equations *Chaos* **27** 114309
[17] Shen Y, Meng Q and Shi P 2014 Maximum principle for mean-field jump-diffusion stochastic delay differential equations and its application to finance *Automatica* **50** 1565–79
[18] Bando M, Hasebe K, Nakanishi K and Nakayama A 1998 Analysis of optimal velocity model with explicit delay *Phys. Rev.* E **58** 5429–35
[19] Forgoston E and Schwartz I B 2008 Delay-induced instabilities in self-propelling swarms *Phys. Rev.* E **77** 035203
[20] Tarama S, Egelhaaf S U and Löwen H 2019 Traveling band formation in feedback-driven colloids *Phys. Rev.* E **100** 022609
[21] Piwowarczyk R, Selin M, Ihle T and Volpe G 2019 Influence of sensorial delay on clustering and swarming *Phys. Rev.* E **100** 012607
[22] Holubec V, Geiss D, Loos S A M, Kroy K and Cichos F 2021 Finite-size scaling at the edge of disorder in a time-delay Vicsek model *Phys. Rev. Lett.* **127** 258001
[23] Haskovec J 2021 Cucker-smale model with finite speed of information propagation: well-posedness, flocking and mean-field limit *Relat. Models* **16** 394–422
[24] Geiß D, Kroy K and Holubec V 2022 Signal propagation and linear response in the delay vicsek model *Phys. Rev.* E **106** 054612
[25] Kopp R A and Klapp S H 2023 Persistent motion of a brownian particle subject to repulsive feedback with time delay *Phys. Rev.* E **107** 024611
[26] Kopp R A and Klapp S H L 2023 Spontaneous velocity alignment of brownian particles with feedback-induced propulsion(a) *Europhys. Lett.* **143** 17002
[27] Chen P-C, Kroy K, Cichos F, Wang X and Holubec V 2023 Active particles with delayed attractions form quaking crystallites(a) *Europhys. Lett.* **142** 67003
[28] Guillouzic S, L'Heureux I and Longtin A 1999 Small delay approximation of stochastic delay differential equations *Phys. Rev.* E **59** 3970–82
[29] Geiss D, Kroy K and Holubec V 2019 Brownian molecules formed by delayed harmonic interactions *New J. Phys.* **21** 093014
[30] Loos S 2021 *Stochastic Systems With Time Delay: Probabilistic and Thermodynamic Descriptions of non-Markovian Processes far From Equilibrium* (*Springer Theses*) (Springer International Publishing)
[31] Holubec V, Ryabov A, Loos S A M and Kroy K 2022 Equilibrium stochastic delay processes *New J. Phys.* **24** 023021
[32] Buckwar E 2000 Introduction to the numerical analysis of stochastic delay differential equations *J. Comput. Appl. Math.* **125** 297–307
[33] Burrage K, Burrage P M and Tian T 2004 Numerical methods for strong solutions of stochastic differential equations: an overview *Proc. R. Soc.* A **460** 373–402
[34] Cao W, Zhang Z and Karniadakis G E 2015 Numerical methods for stochastic delay differential equations via the wong–zakai approximation *SIAM J. Sci. Comput.* **37** A295–318
[35] Sun X and Yang F 2022 Time evolution of probability density in stochastic dynamical systems with time delays: the governing equation and its numerical solution *Chaos* **32** 123124

[36] Holubec V, Kroy K and Steffenoni S 2019 Physically consistent numerical solver for time-dependent Fokker-Planck equations *Phys. Rev.* E **99** 032117
[37] Virtanen P *et al* 2020 SciPy 1.0: fundamental algorithms for scientific computing in python *Nat. Methods* **17** 261–72
[38] Antary N 2023 Matrix numerical method evaluating probability densities for stochastic delay differential equations (available at: https://doi.org/10.5281/zenodo.10406336)
[39] Bullerjahn J T, Sturm S and Kroy K 2020 Non-markov bond model for dynamic force spectroscopy *J. Chem. Phys.* **152** 064104
[40] McKetterick T J and Giuggioli L 2014 Exact dynamics of stochastic linear delayed systems: application to spatiotemporal coordination of comoving agents *Phys. Rev.* E **90** 042135
[41] Risken H and Frank T 1996 *The Fokker-Planck Equation: Methods of Solution and Applications* (*Springer Series in Synergetics*) (Springer)
[42] Hänggi P, Talkner P and Borkovec M 1990 Reaction-rate theory: fifty years after kramers *Rev. Mod. Phys.* **62** 251–341