Annual Review of Condensed Matter Physics

A Primer on Stochastic Partial Differential Equations with Spatially Correlated Noise

Katherine A. Newhall

Department of Mathematics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; email: knewhall@unc.edu

Annu. Rev. Condens. Matter Phys. 2025. 16:195–208

First published as a Review in Advance on November 11, 2024

The *Annual Review of Condensed Matter Physics* is online at conmatphys.annualreviews.org

https://doi.org/10.1146/annurev-conmatphys-042624-033003

Keywords

stochastic heat equation, colored noise, canonical ensemble, magnetization

Abstract

With the growing number of microscale devices from computer memory to microelectromechanical systems, such as lab-on-a-chip biosensors, and the increased ability to experimentally measure at the micro- and nanoscale, modeling systems with stochastic processes is a growing need across science. In particular, stochastic partial differential equations (SPDEs) naturally arise from continuum models—for example, a pillar magnet's magnetization or an elastic membrane's mechanical deflection. In this review, I seek to acquaint the reader with SPDEs from the point of view of numerically simulating their finite-difference approximations, without the rigorous mathematical details of assigning probability measures to the random field solutions. I stress that these simulations with spatially uncorrelated noise may not converge as the grid size goes to zero in the way that one expects from deterministic convergence of numerical schemes in two or more spatial dimensions. I then present some models with spatially correlated noise that maintain sampling of the physically relevant equilibrium distribution. Numerical simulations are presented to demonstrate the dynamics; the code is publicly available on GitHub.

1. INTRODUCTION

Modeling systems with stochastic processes—temporal evolution that progresses randomly—is a growing need across science. The randomness often arises as a method to account for dynamics at a scale that is not resolved, like the thermal forcing of a heat bath. For some observed phenomena, this noise is secondary and not important to consider, for example, the amount of external field needed to reverse magnetization. For others, it is the driving force, as in the waiting time for a spontaneous magnetization reversal to occur.

Stochastic partial differential equations (SPDEs) arise as a natural extension of continuum models. Examples include stochastic Allen-Cahn-type equations for polarization in ferromagnetic materials to study domain wall motion (1–3) and the stochastic Landau–Lifshitz–Gilbert equation for the magnetization to study thermally inducted switching (4). These models use noise that is uncorrelated in both space and time, added in a way that respects a Boltzmann equilibrium.

From a mathematical point of view, many SPDEs with spatially uncorrelated noise in one spatial dimension have been shown to be well-posed (5, 6), whereas in two and higher spatial dimensions they rarely are. From a computational point of view, this means that standard numerical methods converge for one spatial dimensional equations, even if the rate of convergence differs from their deterministic counterparts. For example, the Euler scheme for the 1D stochastic heat equation converges, but in general, no convergence gains are obtained by using higher-order schemes (7). In two and higher spatial dimensions, solutions obtained through simulations may be dependent on the mesh size, with numerical algorithms not converging in the way they do for deterministic PDE (partial differential equation) simulations; see, for example, Reference 8. In fact, References 1 and 2 are aware of this issue and restrict the numerical study to meshes no smaller than the lattice spacing. Much less common on the modeling side is to see spatial correlations included in the noise. One example is Reference 9, which utilizes spatially correlated noise to model nematic liquid crystals with a stochastic Landau–de Gennes equation. Another is Reference 10, which utilizes spatially correlated noise to model fluctuations in the net reproductive rate or the competition term in reaction-diffusion population dynamics.

In this review, I seek to acquaint the reader with SPDEs from the point of view of numerically simulating them, without the rigorous mathematical details of assigning probability measures to the random field solutions. Using the stochastic heat equation as a concrete example, which can be solved exactly, I show that the norm of the solution is infinite in two or more spatial dimensions in line with the problem not being well-posed in these cases. Recall that for well-posed problems, a unique solution exists that depends continuously on input parameters; above was mentioned a case for which the solution did not exist in the set of functions usual for describing physics systems (often a Hilbert space). I then present methods for creating SPDEs with spatially correlated noise that still obey a fluctuation–dissipation relation to ensure the sampling of physically relevant equilibrium distributions. Numerical simulations are used to demonstrate some properties of the resulting SPDEs with correlated noise, motivating future studies that question the use of uncorrelated spatial noise in SPDE models and the consequences of using correlated noise.

In Section 2, I review the basics of stochastic differential equations (SDEs) before proceeding in Section 3 to discuss SPDEs. This section discusses two examples of overdamped Langevin-type equations, the stochastic heat equation and an overdamped version of the stochastic Landau–Lifshitz–Gilbert equation, and one example of an energy-conserving system with stochastic initial data rather than stochastic driving. Results from numerical simulations are presented in Section 4.

2. STOCHASTIC DIFFERENTIAL EQUATIONS

In this section, I describe SDEs and stochastic integrals in terms of finite discretizations in time. These depend on the Wiener processes, $W(t)$, the mathematical representation of Brownian motion. Its properties are (*a*) it begins at zero, $W(0) = 0$; (*b*) its increments are Gaussian random variables with mean zero, $\mathbb{E}[W(t) - W(s)] = 0$, and its standard deviation is the square root of the time increment, $\sqrt{\mathbb{E}[(W(t) - W(s))^2]} = \sqrt{|t - s|}$; (*c*) increments over nonoverlapping time increments are independent; and (*d*) it is continuous (almost surely). Here, t and s represent two points in time, and $\mathbb{E}$ is used to denote expectation over all realizations (an ensemble average).

The first three properties follow from defining $W(t)$ as the limit of a random walk, with $X^0 = 0$ and $X^n = \sum_{i=1}^{n} Z_i \to W(n\Delta t)$, with the Z_i being independent identical random variables: $Z_i = \sqrt{\Delta t}$ with probability 1/2 and $Z_i = -\sqrt{\Delta t}$ with probability 1/2. Property (*b*) follows from the central limit theorem, noting $\mathbb{E}[(X^n)^2] = n\mathbb{E}[Z_i^2] = n\Delta t = t$. Property (*c*) intuitively follows from considering $X^n - X^m$ and $X^s - X^t$ with $n > m \geq s > t$. The same Z_i does not appear in both quantities for any i and only would if the time increments overlapped. Because $X^n - X^m$ and $X^s - X^t$ contain independent random variables, they themselves are independent. Property (*d*) is harder to intuitively see; rigorous treatment comes from the Kolmogorov continuity theorem (see, e.g., 11).

The mathematical form of an SDE,

$$\mathrm{d}X = A(X,t)\mathrm{d}t + B(X,t)\mathrm{d}W, \tag{1.}$$

is written in terms of the Wiener increment $\mathrm{d}W$. Because the mean-squared displacement of the Wiener process grows linearly in time, $\mathbb{E}[|W(t + \Delta t) - W(t)|^2] \propto \Delta t$, the increment $\mathrm{d}W$ scales like $\sqrt{\Delta t}$ as $\Delta t \to 0$. Thus, the numerical approximation to Equation 1 given by the Euler–Maruyama method is

$$X(t + \Delta t) = X(t) + A(X(t), t)\Delta t + B(X(t), t)\sqrt{\Delta t}g,$$

where g is a standard Gaussian random variable (independent, mean zero, variance one). The physics form of Equation 1 is often written in terms of white noise, $\xi(t)$, as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = A(X(t), t) + B(X(t), t)\xi(t) \tag{2.}$$

with $\mathbb{E}[\xi(t)\xi(s)] = \delta(t - s)$. White noise, $\xi(t)$, would be notated $\mathrm{d}W/\mathrm{d}t$, but the limit,

$$\lim_{\Delta t \to 0} \frac{W(t + \Delta t) - W(t)}{\Delta t},$$

does not exist mathematically, as the Wiener increment $W(t + \Delta t) - W(t)$ scales like $\sqrt{\Delta t}$ causing the above limit to grow like $1/\sqrt{\Delta t}$. The delta correlation of white noise is consistent with the property that two Wiener increments over time periods that do not overlap, $W(t + \Delta t) - W(t)$ and $W(s + \Delta t) - W(s)$ with $t + \Delta t \leq s$, are independent and so are the discrete-time white noises, $[W(t + \Delta t) - W(t)]/\Delta t$ and $[W(s + \Delta t) - W(s)]/\Delta t$. In the infinitesimal limit of $\Delta t \to 0$, they only overlap when they are identically equal, $t = s$, hence, the delta-function correlation (as the variance has gone to infinity when they overlap).

Given that the derivative of the Wiener process doesn't exist, it is reasonable to worry that Equations 1 or 2 might not have a unique interpretation. In particular, it is necessary to define the stochastic integral written in terms of either white noise $\xi(t)$ or the Wiener increment $\mathrm{d}W$,

$$\int_0^t B(X(s), s)\xi(s)\mathrm{d}s \leftrightarrow \int_0^t B(X(s), s)\mathrm{d}W. \tag{3.}$$

For deterministic integrals, taking the limit of different Riemann sum approximations (left endpoint versus right endpoint, for example) results in identical answers. This is not the case for stochastic integrals; the following example shows dependence on the Riemann sum approximation used. Consider the expectation of the integral

$$\mathbb{E}\left[\int_0^t W(s)\mathrm{d}W\right] = \lim_{\Delta t\to 0}\sum_j \mathbb{E}\Big[\big[(1-\alpha)W(t_j)+\alpha W(t_j+\Delta t)\big]\big[W(t_j+\Delta t)-W(t_j)\big]\Big], \quad 4.$$

where the parameter α controls where the integrand is evaluated: $\alpha = 0$ corresponds to the left endpoint rule and $\alpha = 1$ corresponds to the right endpoint rule. Using the property that $\mathbb{E}[W(s)W(t)] = \min(s,t)$, the right-hand side reduces to

$$\begin{aligned}\mathbb{E}\left[\int_0^t W(s)dW\right] &= \lim_{\Delta t\to 0}\sum_j (1-\alpha)t_j + \alpha(t_j+\Delta t) - (1-\alpha)t_j - \alpha t_j \\ &= \lim_{\Delta t\to 0}\sum_j \alpha\Delta t = \alpha t.\end{aligned} \quad 5.$$

The value of the expectation of the integral varies from 0 to t depending on where the integrand is evaluated in the Riemann sum approximation; therefore, we must specify which discretization is to be used. Two standard discretizations are used in practice: the Ito stochastic integral (left endpoint rule),

$$\int_0^t B(X(s),s)\mathrm{d}W \approx \sum_i B(X(s_i),s_i)(W(s_{i+1})-W(s_i)), \quad 6.$$

where $s_{i+1} = s_i + \Delta t$, and the Stratonovich stochastic integral (trapezoid rule),

$$\int_0^t B(X(s),s)\circ \mathrm{d}W \approx \sum_i \frac{B(X(s_i),s_i)+B(X(s_{i+1}),s_{i+1})}{2}(W(s_{i+1})-W(s_i)). \quad 7.$$

The Ito left endpoint rule is more natural in systems that are slow to respond to the noise, but its use requires corrections to deterministic calculus rules of differentiation and integration. In contrast, the Stratonovich trapezoid rule does not require these corrections. If the noise is additive, meaning the noise term $B(X, t)$ in Equation 2 is independent of X, these two interpretations lead to identical solutions. As discussed in Section 3.3 for the spin ensemble with multiplicative noise, an additional term appears in the Ito SDE that doesn't come from a physical force in the system but rather corrects for noise effect on the magnitudes of the spins. This highlights a case in which it is important to either choose the Stratonovich interpretation or else carefully determine the terms needed to write an equivalent Ito SDE for systems with multiplicative noise.

3. STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

This section starts by building an SPDE from a finite discretization of the heat equation as an example. The convergence issues in more than one spatial dimension are discussed as well as why from a physics point of view correlated noise might be a desired option. I then discuss models that are driven with correlated noise to overcome the mathematical issues while maintaining a fluctuation–dissipation relation to sample the canonical ensemble. These include a correlated noise version of an overdamped Langevin equation and models built from using Metropolis–Hastings (MH) dynamics. The latter are especially useful for spin systems that include a geometric constraint. The section ends with a brief note about energy-conserving systems in the continuum limit and its equivalence to the continuum Langevin dynamics.

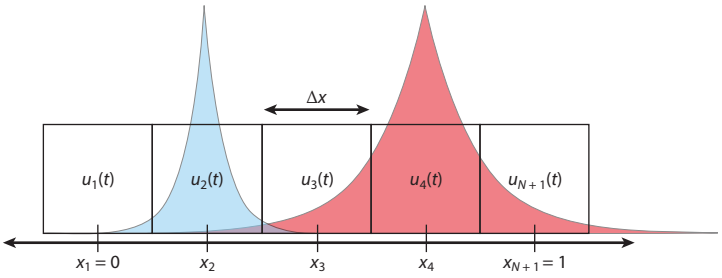

Figure 1

A depiction of the discretization setup for the case of $N = 4$ showing the x location of the boxes, along with the $u_j(t)$ representing the temperature within a box. Two possible covariance functions are depicted for the dynamics being modeled by Gaussian noise. For the blue covariance function centered around x_2, the length scale of the correlations is smaller than the box size Δx, indicating that independent noise driving each box is a reasonable assumption at this length scale of discretization. For the red covariance function centered around x_4, independent noise driving each box is not a reasonable assumption.

3.1. Motivation and Discretization

In this section, I build SPDEs starting from a finite discretization of space using the heat equation as an example system. This provides the basis for discussing and deriving correlated noise–driven SPDEs.

The unit interval in x is discretized into $N + 1$ boxes, each of size $\Delta x = \frac{1}{N}$. The functions $u_j(t)$ represent the temperature in the jth box. This is depicted in **Figure 1** for the case $N = 4$. The set of ordinary differential equations describing the $u_j(t)$ are

$$\frac{\mathrm{d}u_j(t)}{\mathrm{d}t}\Delta x = -c\frac{u_j(t) - u_{j-1}}{\Delta x} + c\frac{u_{j+1}(t) - u_j}{\Delta x}, \qquad 8.$$

for $j = 2 \ldots N$ representing the flow of heat out of box j down into box $j - 1$ and the flow into box j from the above box $j + 1$. The value of u_1 and u_{N+1} are given by the boundary conditions. The parameter c represents the thermal diffusivity. (Inspecting the Hamiltonian below, c could also be viewed as a coupling coefficient or spring constant, penalizing neighboring values of u_j from differing.)

The typical way to add noise to this equation is to assume that the length scale of correlations in the part of the dynamics not being explicitly modeled that generates the noise is shorter than the length of a box; therefore, each box gets perturbed by an independent noise source. This case is depicted in **Figure 1** by the covariance function centered around x_2. The length scale of the correlations is smaller than the box size Δx. Dividing Equation 8 by Δx and then adding noise scaled by $1/\sqrt{\Delta x}$ results in a set of SDEs,

$$\frac{\mathrm{d}u_j(t)}{\mathrm{d}t} = c\frac{u_{j+1} - 2u_j(t) + u_{j-1}}{\Delta x^2} + \sqrt{\frac{2\beta^{-1}}{\Delta x}}\xi_j(t), \qquad 9.$$

that sample the canonical distribution,

$$\mu = Z^{-1}e^{-\beta H}, \qquad 10.$$

with the Hamiltonian given by

$$H = \Delta x \sum_{j=1}^{N} \frac{c}{2}\left(\frac{u_{j+1} - u_j}{\Delta x}\right)^2. \qquad 11.$$

Intuitively, scaling the noise in Equation 9 by $1/\sqrt{\Delta x}$ aligns with the idea that a smaller box would have more fluctuations because fewer particles are being averaged over; the law of large numbers says such fluctuations scale like $1/\sqrt{n}$ as the number of observations of a random variable $n \to \infty$. Note the deterministic part of Equation 9 is $\frac{\partial}{\partial u_j}\left(\frac{H}{\Delta x}\right)$ as the Δx from the Hamiltonian appears in the noise term of Equation 9. Comparing with the canonical distribution in Equation 10, $\beta \Delta x$ has gone to the noise term, and the remaining part of the Hamiltonian has gone to the deterministic term of Equation 9 to maintain the fluctuation–dissipation relationship.

Taking $N \to \infty$ ($\Delta x \to 0$) in Equation 9, the set of $\xi_j(t)/\sqrt{\Delta x}$ converge to $\eta(x, t)$, i.e., space-time white noise. Notice the similar scaling of $1/\sqrt{\Delta x}$ to the $1/\sqrt{\Delta t}$ scaling of white-in-time noise from the previous section. The resulting SPDE is

$$\frac{\partial u(x,t)}{\partial t} = c\frac{\partial^2 u(x,t)}{\partial x^2} + \sqrt{2\beta^{-1}}\eta(x,t), \qquad 12.$$

with the Hamiltonian converging to

$$H = \int_0^1 \frac{c}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2 \mathrm{d}x. \qquad 13.$$

Although Equation 12 is perfectly reasonable in one dimension, in higher spatial dimensions the $N \to \infty$ limit produces rougher and rougher solutions with no real notion of convergence. This can be seen just by looking at the expected norm of the solution. Equation 12 can be analytically solved, even in higher spatial dimensions, as in Fourier space it is statistically equivalent to

$$\mathrm{d}\hat{u}_{\vec{k}}(t) = -|\vec{k}|^2 c\hat{u}_{\vec{k}}(t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}W_{\vec{k}}(s), \qquad 14.$$

a set of decoupled SDEs, each an Ornstein–Uhlenbeck process, with solution given by

$$\hat{u}_{\vec{k}}(t) = \sqrt{2\beta^{-1}}\int_0^t e^{-|\vec{k}|^2(t-s)}\mathrm{d}W_{\vec{k}}(s).$$

The expectation of the norm is then

$$\mathbb{E}[\|u\|^2] = \sum_k \mathbb{E}[(\hat{u}_{\vec{k}})^2] = \sum_{\vec{k}} \frac{\beta^{-1}}{|\vec{k}|^2}(1 - e^{-2|\vec{k}|^2 t}),$$

which utilized the known second moment of the Ornstein–Uhlenbeck process. The convergence of this sum is dominated by the $1/|\vec{k}|^2$ term. In one spatial dimension, $\sum_k k^{-2}$ converges, but in two spatial dimensions, $\sum_{j,k}(j^2 + k^2)^{-1}$ does not. In fact, this is the borderline case with all higher dimensions diverging as well; see Reference 8 for more details. What this means (and examples of this are shown in Reference 8) is that numerical solutions to the 2D case depend drastically on the value of N used in the discretization, with deterministic evolution getting lost in highly oscillatory noise. Physically, the original assumption that boxes receive independent fluctuations because the correlation length scale of the dynamics generating the noise is shorter than a box size is eventually violated as the box size shrinks to zero. Such a case is depicted in **Figure 1** by the covariance function centered around x_4. The length scale of the correlations is larger than the box size Δx.

Both the mathematical convergence issue and violation of the physical assumption issue can be fixed by putting the correlations back into the noise. A covariance matrix C_{ij} encodes the covariance between the noises arriving at box i and box j and generally decays with the distance between the two boxes. (A possible way to construct C_{ij} is discussed in Section 4.1.) Using this to correlate the white-in-time noises in Equation 9 results in

$$\frac{\mathrm{d}u_j(t)}{\mathrm{d}t} = c\frac{u_{j+1} - 2u_j(t) + u_{j-1}}{\Delta x^2} + \sqrt{2\beta^{-1}}\sum_i C_{ji}^{1/2}\xi_i(t), \qquad 15.$$

where the $1/\sqrt{\Delta x}$ has been absorbed into C_{ij} so that in the white noise case $C_{ii} = N$ and zero otherwise, converging to the delta function $\delta(x-y)$ as $N \to \infty$. The problem with just adding C to the noise term in Equation 15 is that it no longer obeys the fluctuation–dissipation relation. That is, Equation 9 was specially crafted so that the dissipation of energy from the deterministic term was balanced with fluctuations of the right size in the noise term to sample the canonical distribution (Equation 10). By only changing the fluctuations in Equation 15, this balance is disrupted and it no longer samples the canonical distribution. By inspection of the Fokker–Planck equation, it is perhaps possible to guess (and then verify next) that

$$\frac{\mathrm{d}\vec{u}(t)}{\mathrm{d}t} = -\frac{C}{N}\nabla\left(\frac{H}{\Delta x}\right) + \sqrt{2\beta^{-1}}C^{1/2}\vec{\xi}(t) \qquad 16.$$

retains both correlations in the noise and samples the canonical distribution (Equation 10). In the limit as $N \to \infty$, this equation becomes the nonlocal SPDE,

$$\frac{\partial u(x,t)}{\partial t} = -\int_0^1 C(x-y)\frac{\partial G}{\partial u}\mathrm{d}y + \sqrt{2\beta^{-1}}\eta^C(x,t), \qquad 17.$$

where $\eta^C(x,t)$ is white in time and correlated in space, i.e., $\mathbb{E}[\eta^C(x,t)\eta^C(y,s)] = \delta(t-s)C(x-y)$, and the Hamiltonian is defined as $H = \int_0^1 G(u(x))\mathrm{d}x$.

I briefly validate that the equilibrium distribution in Equation 10 is the steady-state distribution of Equation 16. The Fokker–Planck equation for the evolution of the distribution $\rho(\vec{u},t)$ is

$$\frac{\partial \rho(\vec{u},t)}{\partial t} = \nabla\cdot\left[\frac{C}{N}\nabla\frac{H}{\Delta x}\rho(\vec{u},t)\right] + \beta^{-1}C\nabla^2\rho(\vec{u},t). \qquad 18.$$

Plugging in $\rho(\vec{u},t) = \exp(-\beta H)$ from Equation 10 into the right-hand side of the Fokker–Planck equation, we obtain

$$\nabla^2\rho(\vec{u},t) = -\beta\nabla\cdot\nabla H e^{-\beta H} + \beta^2\nabla H\cdot\nabla H e^{-\beta H},$$
$$\nabla\cdot[\nabla H\rho(\vec{u},t)] = \nabla\cdot\nabla H e^{-\beta H} - \beta\nabla H\cdot\nabla H e^{-\beta H}.$$

Because $N\Delta x = 1$, the first and second terms on the right-hand side of the Fokker–Planck equation cancel, verifying that $\partial\rho/\partial t = 0$ and, thus, $\exp(-\beta H)$ is the steady-state distribution (the normalizing factor does not affect the derivatives).

3.2. Metropolis–Hastings for Langevin Dynamics

A more algorithmic approach to utilizing colored noise and sampling the desired canonical distribution in Equation 10 comes from the limiting behavior of MH dynamics. The algorithm proposes a new microstate from some generic distribution that is easy to sample (i.e., a Gaussian distribution), and then accepts or rejects this new microstate based on the desired canonical distribution to sample. Sequential application of this algorithm produces a sequence of microstates that sample the desired canonical distribution (in the limit as the number of sampled microstates goes to infinity).

Specifically for the purposes here, the proposal of a new microstate is based on a random walk with the desired correlation in the noise,

$$\tilde{u}_j = u_j^n + \epsilon\sum_i C_{ji}^{1/2}g_i, \qquad 19.$$

where ϵ controls the size of the spatial step and the g_i are independent, identically distributed, standard Gaussian random variables. This proposed new microstate is accepted with probability

α based on sampling the canonical distribution (Equation 10),

$$\vec{u}^{n+1} = \begin{cases} \tilde{\vec{u}} & \text{with prob. } \alpha, \\ \vec{u}^n & \text{with prob. } 1-\alpha, \end{cases} \tag{20.}$$

$$\alpha = \min(1, e^{-\beta(H(\tilde{\vec{u}})-H(\vec{u}^n))}). \tag{21.}$$

In this way, proposed microstates with lower energy are always accepted, whereas proposed microstates with higher energy are sometimes accepted. By Taylor expanding for small ϵ, we have, for the difference in Hamiltonians,

$$H(\tilde{\vec{u}}) - H(\vec{u}^n) \approx [H(\vec{u}^n) - \epsilon C^{1/2}\vec{g}\cdot\nabla H] - H(\vec{u}^n) = \epsilon C^{1/2}\vec{g}\cdot\nabla H; \tag{22.}$$

the drift becomes

$$\mathbb{E}[\vec{u}^{n+1} - \vec{u}^n] \approx \mathbb{E}[\epsilon C^{1/2}\vec{g}\min(1, e^{-\beta\epsilon C^{1/2}\vec{g}\cdot\nabla H})] = -\frac{\beta\epsilon^2}{2}C\nabla H, \tag{23.}$$

and the diffusion becomes

$$\vec{u}^{n+1} - \vec{u}^n - \mathbb{E}[\vec{u}^{n+1} - \vec{u}^n] \approx \epsilon C^{1/2}\vec{g}. \tag{24.}$$

Note that the error terms in these Taylor expansions are themselves random variables and could be arbitrarily large on any given realization; rather, convergence is shown in probability (see 12). Defining the time step to be $\Delta t = \beta\epsilon^2/2$, the drift and diffusion together form the stochastic Euler step,

$$\vec{u}^{n+1} = \vec{u}^n - C\nabla H\Delta t + \sqrt{2\beta^{-1}\Delta t}C^{1/2}\vec{g}, \tag{25.}$$

for the SDE in Equation 16 above. Notice in Equation 16 that the product $N\Delta x = 1$ was added and split between the terms C/N and $H/\Delta x$ to aid in seeing the convergence to the SPDE in Equation 17.

3.3. Metropolis–Hastings for Spin System

The advantage of studying the convergence of the MH method is that it eliminates the need for a creative guess at how to fix the fluctuation–dissipation relation and, thus, might prove useful for more complex systems, like in the Landau–Lifshitz–Gilbert model for spins. The derivation developed in References 13 and 12 is summarized here.

Models for magnetization evolve a set of spins, σ_j, each of which has a conserved length but can orient in any direction. Here, I focus on the case of 3D spin vectors. Another way to envision this is that the spin vector represents a point moving about on the surface of a sphere. In this case, the MH proposal for each spin would be a small step along the surface of the sphere (mathematically, along a geodesic). This small step can be approximated by a move in the tangent plane, which is then projected back onto the surface of the sphere, namely

$$\tilde{\sigma}_j = \frac{\sigma_j^n + \epsilon P_\perp(\sum_i C_{ji}^{1/2} g_i)}{|\sigma_j^n + \epsilon P_\perp(\sum_i C_{ji}^{1/2} g_i)|}, \tag{26.}$$

where $P_\perp$ is a projection perpendicular to σ_j^n. As above, ϵ controls the size of the step, and the g_i are independent, identically distributed, standard Gaussian random variables.

The problem that arises is the accept/reject probability in Equation 21 only guarantees sampling Equation 10 if the proposal is symmetric. That is, the probability of making the move

to the proposal is the same as the probability of making the move back to the original point, $\mathbb{P}(\sigma \to \tilde{\sigma}) = \mathbb{P}(\tilde{\sigma} \to \sigma)$. Note this is slightly different from reversibility, which equates the flux of probability back and forth between two microstates, $\mathbb{P}(\sigma)\mathbb{P}(\sigma \to \tilde{\sigma}) = \mathbb{P}(\tilde{\sigma})\mathbb{P}(\tilde{\sigma} \to \sigma)$. For white noise ($C_{ii} = N$ and zero otherwise), the symmetry is preserved for the sphere, but likely not for any arbitrary curved surface as the changing curvature would consolidate or fan out nearby points that were moved. In the case of correlated noise, the symmetry is broken even for the sphere as the projection and the correlation do not commute: Correlating projected noise is not the same as projecting correlated noise. However, in the case of the projection being $\sigma_j^n \times (\cdot)$, these asymmetries appear to arise in higher orders of ϵ, and the machinery in Section 3.2 results in a set of SDEs that sample the desired distribution (as verified via the Fokker–Planck equation).

Following the machinery in Section 3.2, with the projection being $\sigma_j^n \times (\cdot)$ in Equation 26, the set of SDEs for a system of N spins located along a 1D lattice in x that sample the canonical distribution is given by

$$\frac{\mathrm{d}\vec{s}}{\mathrm{d}t} = P\frac{C}{N}P^T\Delta_N\vec{s} - 2\beta^{-1}\frac{\mathrm{Tr}(C)}{N}\vec{s} + \sqrt{2\beta^{-1}}PC^{1/2}\vec{\xi}(t), \tag{27.}$$

to be interpreted in the Ito sense. Here, $\vec{s}$ is the stacked vector of the three components for each of the N spins, $\vec{s} = [\sigma_{1,x}, \sigma_{2,x} \ldots \sigma_{N,x}, \sigma_{1,y}, \ldots \sigma_{N,y}, \sigma_{1,z}, \ldots \sigma_{N,z}]$, and $\Delta_N\vec{s}$ is the discrete Laplacian of this stacked vector. In terms of the spin vectors themselves, for $\mathrm{d}\sigma_{j,q}/\mathrm{d}t$ they would return $(\sigma_{j+1,q} - 2\sigma_{j,q} + 2\sigma_{j-1,q})/\Delta x^2$ for each $j = 2 \ldots N-1$ and $q \in \{x, y, z\}$ corresponding to the Hamiltonian

$$H = \Delta x \sum_{j=1}^{N} \tfrac{1}{2}\left(\frac{\|\sigma_{j+1} - \sigma_j\|}{\Delta x}\right)^2,$$

with $\sigma_{N+1} = \sigma_1$ for periodic boundary conditions. The covariance matrix C is now a block diagonal matrix with three repetitions of the covariance matrix along the diagonal, and P is a block matrix describing the cross-product projection,

$$P = \begin{pmatrix} 0 & -Z & Y \\ Z & 0 & X \\ -Y & X & 0 \end{pmatrix}, \tag{28.}$$

with each block being diagonal with those components of the spins,

$$Q = \begin{pmatrix} \sigma_{1,q} & 0 & \ldots & 0 \\ 0 & \sigma_{2,q} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_{N,q} \end{pmatrix}, \tag{29.}$$

for $Q \in \{X, Y, Z\}$ and corresponding $q \in \{x, y, z\}$.

The first term on the right-hand side of Equation 27 is the gradient descent of the Hamiltonian, with the addition of the projection matrix P and its transpose to account for the geometric constraint of conserving the length of each σ_j. The covariance matrix C appears in this term to balance the fluctuations from the correlated noise in the same manner as in the heat equation. Because the projection matrix depends on the σ_j, the third term in Equation 27 is multiplicative noise; the machinery has automatically provided the second term, the Ito correction term $-2\beta^{-1}\frac{\mathrm{Tr}(C)}{N}\vec{s}$, that would not be present if the multiplicative noise was interpreted in the Stratonovich sense. The second term corrects for the fact that the multiplicative noise does not conserve the length of each σ_j when using the Ito interpretation.

In the limit as $N \to \infty$, Equation 27 limits to the nonlocal SPDE,

$$\partial_t \sigma(x,t) = -\sigma(x,t) \times \int_{\mathbb{T}^d} C(x-y)(\sigma \times \Delta\sigma)(y,t)\mathrm{d}y + \sqrt{2\beta^{-1}}\sigma(x,t) \times \eta^C(x,t), \qquad 30.$$

written with the Stratonovich interpretation of the multiplicative noise.

3.4. Stochastic Energy-Conserving Systems

The above-described methods constructed overdamped Langevin-like parabolic SDEs in which the stochasticity comes from the noise driving the system. These overdamped systems have no momentum while maintaining a fluctuation–dissipation relationship consistent with the canonical distribution. Finite dimensional energy-conserving systems sample the microcanonical distribution, yet it is possible to construct energy-conserving continuum (infinite dimensional) systems that sample the canonical distribution. Here, the stochasticity comes from the initial conditions with no further added noise driving the system and, thus, no fluctuation–dissipation relationship.

In Reference 14, the SPDE

$$\frac{\partial^2 u(x,t)}{\partial t^2} = c\frac{\partial^2 u(x,t)}{\partial x^2} + V'(u) = 0, \qquad 31.$$

with initial conditions drawn from the infinite-energy canonical distribution (Equation 10) with

$$H = \int_0^1 \frac{1}{2}p^2 + \frac{c}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2 + V(u)\mathrm{d}x, \qquad 32.$$

where $p = \partial u/\partial t$, was shown to describe the limiting dynamics of the finite discretized energy-conserving system

$$\frac{\mathrm{d}u_j}{\mathrm{d}t} = p_j, \qquad \frac{\mathrm{d}p_j}{\mathrm{d}t} = c\frac{u_{j-1} - 2u_j + u_{j+1}}{\Delta x^2} - V'(u_j), \qquad 33.$$

with initial conditions taken from the microcanonical ensemble with energy $\frac{N}{\beta}$,

$$\mu_m = Z_m^{-1}\delta\left(H - \frac{N}{\beta}\right). \qquad 34.$$

Furthermore, it was shown in Reference 14 that in the limit of $N \to \infty$, both the canonical and microcanonical ensembles are statistically equivalent; the characteristic functions for the set of u_j and p_j are equivalent, and thus the probability distributions are equivalent. The intuition for why this happens is because, for the canonical ensemble, the probability gets concentrated in a thin shell near the average energy, N/β, as N grows. As $N \to \infty$, this thin shell becomes infinitely thin and thereby identical to the microcanonical ensemble confined to this (now infinite energy) shell. (The infinite energy comes from the fact that $\mathrm{d}u/\mathrm{d}x$ in H is infinite, as the solutions are rough and not differentiable in space.)

4. SIMULATIONS OF SPECIFIC EXAMPLES

Using one spatial dimension in which the white noise–driven dynamics are well defined, simulation results are presented to show the effects of the correlation added to the above-derived models. In particular, time-averaged and ensemble statistics do not change, but the temporal evolution is affected. The MATLAB code used to generate these figures can be accessed on GitHub (**https://github.com/knewhall/SPDE_examples**).

4.1. Correlation Matrix

Before simulating the SPDEs, the exact form of the correlation matrix needs to be chosen. One method to construct a correlation matrix is to generate it in Fourier space, with the set of basis functions consistent with the desired boundary conditions of the SPDE. White noise has equal power in all frequencies, whereas colored noises have frequency-dependent power. Specifically, scaling like $1/f^{\kappa}$ for $0 < \kappa$ produces spatially decaying correlations. The general form of the covariance matrix used here is

$$C = V\Lambda^{-2\kappa}V^{T}, \qquad 35.$$

in which the matrix V contains the Fourier modes as column vectors, and the diagonal matrix Λ contains the corresponding frequencies. The correlated noise vector is then $V\Lambda^{-\kappa}\vec{g}$, where the elements of $\vec{g}$ are independent Gaussian random variables and κ controls the decay rate of the frequencies in this decomposition. The eigenvalues of the correlation matrix decay with 2κ. The detailed matrices for Dirichlet and periodic boundary conditions are presented next, consistent with the definition of C used in either Equation 16 or 27 with C having N on the diagonal for white noise.

For the case of Dirichlet boundary conditions $u_1(t) = u_N(t) = 0$, the Fourier frequencies are $\lambda_j = j\pi$ for $j = 1 \ldots N$, and the Fourier modes are $\vec{v}_j = \sqrt{2}\sin(\lambda_j \vec{x})$, where $x_i = \frac{i-1}{N-1}$. An example is shown in **Figure 2*a*** with $N = 65$ and $\kappa = 0.75$. When $\kappa = 0$ is chosen, the matrix C reduces to N on the diagonal, which is consistent with the uncorrelated SDE in Equation 9. As $\kappa \to \infty$, the noise becomes 100% correlated, and C has the same entry everywhere. Because the boundary points have zero variance, C becomes zero everywhere, trivially reducing Equation 16 to $\mathrm{d}\vec{u}/\mathrm{d}t = 0$.

For the case of periodic boundary conditions, both sine and cosine functions are required, with $\lambda = [1, 2\pi, 4\pi, 6\pi, \ldots N\pi, 2\pi, 4\pi, 6\pi, \ldots (N-2)\pi]$ for even N and $\vec{v}_1 = 1$, $\vec{v}_j = \sqrt{2}\cos(\lambda_j \vec{x})$ for $j = 2 \ldots N/2$, $\vec{v}_j = \cos(\lambda_j \vec{x})$ for $j = N/2 + 1$ (note the change in the prefactor needed to ensure this discrete eigenvector has length N), and $\vec{v}_j = \sqrt{2}\sin(\lambda_j \vec{x})$ for $j = N/2 + 2 \ldots N$, taking $x_i = \frac{i-1}{N}$. An example is shown in **Figure 2*b*** with $N = 64$ and $\kappa = 0.75$. When $\kappa = 0$ is chosen, the matrix C reduces to N on the diagonal, consistent with the uncorrelated SDE in Equation 9. As $\kappa \to \infty$ the noise becomes 100% correlated, and C has the same entry everywhere. This value is 1; the SDE in Equation 16 reduces to $\mathrm{d}\vec{u}/\mathrm{d}t = -\nabla H + \sqrt{2\beta^{-1}}\xi_1 \vec{1}$ where $\vec{1}$ is a vector of all ones so that every component of $\vec{u}$ is driven by the same noise ξ_1.

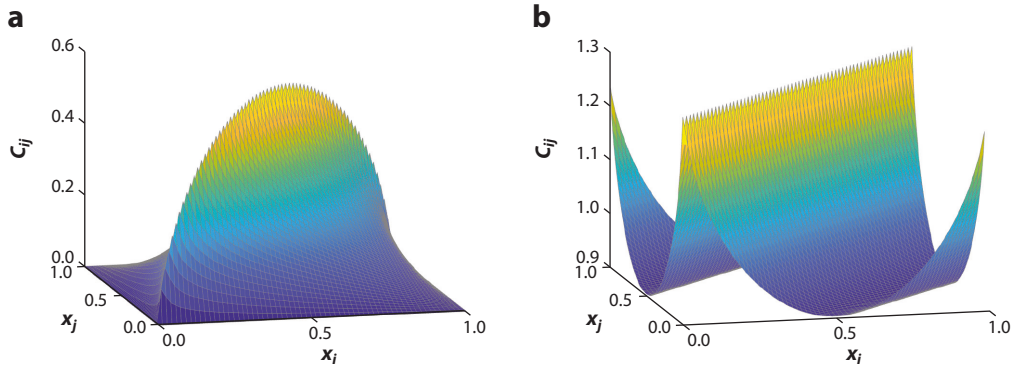

Figure 2

Covariance matrix for (*a*) Dirichlet boundary conditions and (*b*) periodic boundary conditions, both with $\kappa = 0.75$. Recall that κ controls the decay rate of the eigenvalues in the correlation matrix for the noise.

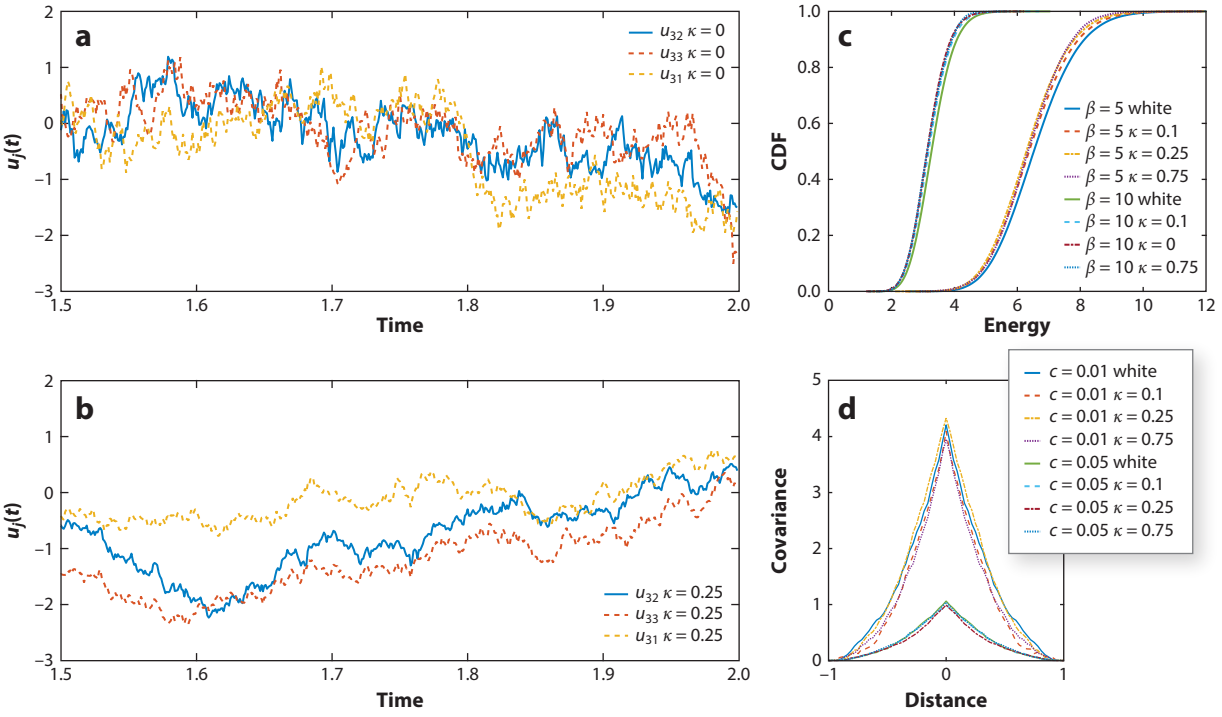

Figure 3

Temporal evolution of the point at the center of the domain, $u_{32} \approx u(\frac{1}{2})$, and two neighboring points $u_{31} \approx u(\frac{1}{2} - \Delta x)$ and $u_{33} \approx u(\frac{1}{2} + \Delta x)$ of Equation 16 for (*a*) white noise with $\kappa = 0$ and (*b*) colored noise with $\kappa = 0.25$. Recall that κ controls the decay rate of the eigenvalues in the correlation matrix for the noise. (*c*) Verification that the equilibrium distribution of the energy is independent of correlations in the driving noise. The CDF of the energy is shown at two different inverse temperatures β. (*d*) The covariance $\mathrm{Cov}(u(x), u(1/2))$ as a function of distance from the center of the domain, $x - 1/2$, are also independent of correlations in the driving noise; they depend solely on the coupling coefficient (thermal diffusivity) c. Abbreviation: CDF, cumulative distribution function.

4.2. Sampling

Figure 3 contains the results of numerically simulating the discrete stochastic heat equation, Equation 16 using the Hamiltonian in Equation 11, with $N = 65$ points. Dirichlet boundary conditions are imposed; therefore, the correlation matrix is formed as in **Figure 2*a***. If not specified in the legend, parameters are set to $c = 0.01$ for the thermal diffusivity, $\beta = 5$ for the inverse temperature, and $\Delta t = 0.001$ for the time step.

In **Figure 3*a*,*b***, the temporal evolution of three center values of u is shown for the case of white noise (**Figure 3*a***, $\kappa = 0$) and colored noise (**Figure 3*b***, $\kappa = 0.25$). The difference in temporal evolution is noticeable, with white noise displaying more rapid change and larger fluctuations than the colored noise case. In both cases, the trajectories of the center point and its two neighboring points are similar because the Hamiltonian is an energy penalty for differences between neighboring points.

Despite noticeable differences in the temporal evolution, regardless of correlation parameter κ, these simulations sample the same equilibrium distribution, as shown in **Figure 3*c***. The cumulative distribution function of the energy H, sampled in time, is shown. It depends only on the value of the inverse temperature β, with larger β corresponding to lower energies. At fixed β, the same canonical distribution is sampled in time, regardless of the amount of correlations added to the system, controlled by κ.

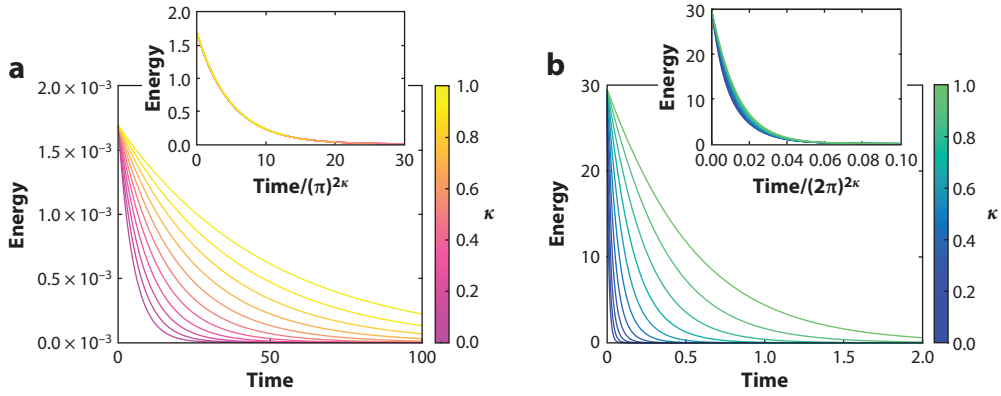

Figure 4

Decay of the energy to equilibrium for the deterministic dynamics (temperature $\beta^{-1} = 0$) of (*a*) Equation 16 with the inset showing the same evolution with time rescaled by $\pi^{2\kappa}$ and (*b*) Equation 27 with the inset showing the same evolution with time rescaled by $(2\pi)^{2\kappa}$. The color bars indicate the values of κ used; κ controls the decay rate of the eigenvalues in the correlation matrix for the noise.

One might wonder if adding correlations to the noise has caused neighboring values of u_j to be more correlated. Confirmation that no additional correlations are being added to the values of u_j due to the correlated noise is shown in **Figure 3*d***. Here, the covariance between the center value $u(1/2) \approx u_{32}$ and other values $u(x)$ are displayed as a function of separating distance $x - 1/2$. This covariance is independent of the amount of correlations added to the system, controlled by κ, depending solely on the value of the thermal diffusivity c. As mentioned above, the coefficient c acts like a coupling coefficient in the Hamiltonian and, thus, correlates neighboring values of u_j. The shorter broader peak (relative to its height) is for the larger value of $c = 0.05$. This stronger coupling has dampened the effect of the noise, causing less variance in the values of u_j indicated by the lower value of the peak.

The difference in temporal dynamics seen between panels *a* and *b* of **Figure 3** is a result of the change to the deterministic part of the dynamics to adjust for the addition of colored noise while maintaining sampling of the canonical distribution. To further illustrate this change, how the deterministic dynamics evolve toward equilibrium is shown in **Figure 4** for both the heat equation, Equation 16 using the Hamiltonian in Equation 11 with $\beta^{-1} = 0$, and the deterministic spin system, Equation 27 with $\beta^{-1} = 0$. The system is started out of equilibrium, and the energy is tracked as it evolves toward equilibrium. In both cases, the equilibrium is all variables identical; this microstate has zero energy. For Equation 16, the initial configuration is a parabola, whereas for Equation 27 it is a figure eight. As the correlations increase with κ, the decay rate of the energy decreases. This is consistent with the first eigenvalue of the covariance matrix, which is less than one, dominating this scaling. This is shown in the insets of both panels in **Figure 4**. For the heat equation, this scaling is $\pi^{2\kappa}$ corresponding to the covariance matrix with Dirichlet boundary conditions. It appears to exactly collapse the data. For the spin system, this scaling is $(2\pi)^{2\kappa}$ corresponding to the covariance matrix with periodic boundary conditions. Note this matrix has an eigenvalue equal to one due to the translational degree of freedom. The $(2\pi)^{2\kappa}$ scaling appears to only approximately collapse the data, indicating further dependence on the interaction with the confining geometry.

5. SUMMARY AND OUTLOOK

Finite-difference approximations of SPDEs are a natural modeling technique for spatially extended systems. At a given finite mesh size, the underlying assumption of spatially uncorrelated

noise is justified but is violated in the limit of the mesh size to zero. Thus, space-time white noise–driven SPDEs may not be the best continuum model both for the violation of this assumption and for the fact that many are not mathematically well posed in two or more spatial dimensions.

I have shown two examples of using spatially correlated noise in a way that is consistent with a fluctuation–dissipation relation to maintain sampling of a physically relevant equilibrium distribution. Although equilibrium distributions and statistics taken over them are independent of correlation by design, there are noticeable differences in the temporal evolution. What effects will show up when studying noise-induced phenomena? For example, will there be an effect on expected waiting times for thermally induced magnetization reversal in ferromagnetic devices? Transition state theory computes the expected waiting time for such a transition from the equilibrium distribution; thus one might expect no change. The transition path itself is everywhere parallel to the gradient of the Hamiltonian in white noise–driven systems. Such a path would still be parallel to the covariance matrix times the gradient, but is this still the transition path? Will there be effects of the correlated noise in these leading-order approximations or only as corrections?

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I would like to thank Mastawal Tirfe for his help checking the simulations for errors. This work is partially supported by the National Science Foundation under grant number DMS-2307297.

LITERATURE CITED

1. Indergand R, Vidyasagar A, Nadkarni N, Kochmann DM. 2020. *J. Mech. Phys. Solids* 144:104098
2. Bauer N, Neumayer SM, Maksymovych P, Lavrentovich MO. 2022. *Phys. Rev. Mater.* 6(12):124401
3. Khajehtourian R, Frazier MJ, Kochmann DM. 2022. *Extreme Mech. Lett.* 50:101527
4. Bañas u, Brzeźniak Z, Prohl A. 2013. *SIAM J. Sci. Comput.* 35(1):B62–81
5. Walsh JB. 1986. In *École d'Été de Probabilités de Saint Flour XIV, 1984*, ed. R Carmona, H Kesten, JB Walsh, PL Hennequin, pp. 265–439. Berlin/Heidelberg: Springer
6. Da Prato G, Zabczyk J. 1992. *Stochastic Equations in Infinite Dimensions*, Vol. 45: *Encyclopedia of Mathematics and Its Applications*. Cambridge, UK: Cambridge Univ. Press
7. Davie AM, Gaines JG. 2001. *Math. Comput.* 70(233):121–34
8. Ryser MD, Nigam N, Tupper PF. 2012. *J. Comput. Phys.* 231(6):2537–50
9. Dalby JL, Majumdar A, Wu Y, Kisan Dond A. 2024. *Liq. Cryst.* 51(2):276–96
10. Méndez V, Llopis I, Campos D, Horsthemke W. 2010. *Theor. Popul. Biol.* 77(4):250–56
11. Stroock DW, Varadhan SRS. 2007. *Multidimensional Diffusion Processes*. Berlin/Heidelberg: Springer
12. Gao Y, Kirkpatrick K, Marzuola J, Mattingly J, Newhall KA. 2021. *Commun. Math. Sci.* 19(2):453–94
13. Gao Y, Marzuola JL, Mattingly JC, Newhall KA. 2020. *Phys. Rev. E* 102(5):052112
14. Newhall KA, Vanden-Eijnden E. 2017. *J. Nonlinear Sci.* 27(3):1007–42