Forward, backward, and weighted stochastic bridges

Peter D. Drummond
Physics Department, Swinburne University of Technology, Melbourne 3122, Victoria, Australia
(Received 1 May 2017; revised manuscript received 10 September 2017; published 12 October 2017)

We define stochastic bridges as conditional distributions of stochastic paths that leave a specified point in phase-space in the past and arrive at another one in the future. These can be defined relative to either forward or backward stochastic differential equations and with the inclusion of arbitrary path-dependent weights. The underlying stochastic equations are not the same except in linear cases. Accordingly, we generalize the theory of stochastic bridges to include time-reversed and weighted stochastic processes. We show that the resulting stochastic bridges are identical, whether derived from a forward or a backward time stochastic process. A numerical algorithm is obtained to sample these distributions. This technique, which uses partial stochastic equations, is robust and easily implemented. Examples are given, and comparisons are made to previous work. In stochastic equations without a gradient drift, our results confirm an earlier conjecture, while generalizing this to cases with path-dependent weights. An example of a two-dimensional stochastic equation with no potential solution is analyzed and numerically solved. We show how this method can treat unexpectedly large excursions occurring during a tunneling or escape event, in which a system escapes from one quasistable point to arrive at another one at a later time.

DOI: 10.1103/PhysRevE.96.042123

I. INTRODUCTION

What is the probability that a physical system will arrive near a certain destination, knowing its initial configuration? What is the probability that it started near a certain point in phase-space, given a final configuration? What happens at intermediate times if both end-points are known? The last question was first considered by Schrödinger [1], in the context of diffusion in a classical system. Such problems are known as "Brownian bridges." In physics, this terminology is often extended to the term "stochastic bridges." These problems are especially important in studying questions of large but rare fluctuations and are related to the concept of a prehistory of a fluctuation [2]: what are the most likely prior events that lead to a given subsequent observation?

The underlying trajectories of diffusion in a phase-space are often described with a stochastic differential equation (SDE). These are very common in physics, chemistry, biology, engineering, and finance [3]. Their transforms are the topic of extended analysis in the mathematics literature [4,5] and are related to diagrammatic methods in classical statistical physics [6]. They are equivalent to probability conserving Fokker-Planck equations (FPEs) [7,8]. However, there are many examples where the relevant SDEs are trajectories with weights, that is, path-dependent probabilities, as in the case of the backwards Kolmogorov equation [9], which describes the probability that an initial condition will reach a predetermined final state. These also occur when computing finite-temperature canonical ensembles [10,11] and dynamical simulations [12] in quantum many-body physics.

Here, we consider stochastic bridges with arbitrary numbers of components, as well as spatially varying weight and drift terms. This allows us to treat time-reversed as well as forward time stochastic bridge. Although software exists for such bridges in ecological models [13], the general, nonlinear stochastic bridge is nontrivial to sample efficiently. We show that such problems are generically transformable to an unweighted partial stochastic differential equation in an extra dimension. The advantage of the approach is that a numerical evaluation is straightforward, using standard partial stochastic differential equation (PSDE) algorithms with equally weighted samples, and it is readily extended to bridges of stochastic fields in arbitrary space dimension. This extends previously known algorithms [14,15] to include weighted and backwards time equations, as well as confirming an earlier conjecture [14] for stochastic bridges without potential conditions.

There are many known methods for obtaining numerical samples of *unweighted* SDEs [16–19]. When there is a weight factor as well, efficiency issues become important. Simple methods of direct sampling combined with weights encounter severe problems. As the spread of weights increases, most computed trajectories have a very low weight compared to the most probable trajectory. As a result, the effective number of equally weighted trajectories decreases with time, and the sampling error increases [20]. There are methods to treat this, including the use of genetic or branching algorithms [10,21]. However, even these relatively sophisticated methods do not appear useful to constrained stochastic bridge calculations, because most trajectories—apart from a subset of measure zero—do not satisfy the constraint.

The origin of the sampling problem that occurs with stochastic bridges, is that standard SDE solvers [18] do not allow a constrained end-point. The direct use of stochastic differential equations is inefficient when applied to stochastic bridges. Except for special cases [22], algorithms for these problems are not widely known, especially for high-dimensional dynamical variables and fields with arbitrary nonlinear terms and weights. As well as in physics, applications are known in probability theory [23], stochastic finance [24], queuing theory, ecological analysis [25], and computer science.

Stochastic bridges and related problems have been treated numerically in other ways as well. To give a brief account, Markov chain Monte Carlo methods of various types exist

[22] and could potentially be generalized to the problems studied here. Similar ideas have been used to study rare events and large fluctuations, using transition path sampling [26] and other methods [27]. The stochastic method obtained in this paper is a partial stochastic differential equation in an extra dimension, obtained by extending earlier work on path-integral solutions to Fokker-Planck equations [28–33]. This can be further improved through Metropolis-adjusted Langevin algorithms (MALA) [22,34].

The time-reversed stochastic bridge is closely related to the time-reversal question implicit in Schrödinger's original title: "Über die Umkehrung der Naturgesetze," or "On the time-reversal of natural laws" [1]. Such issues arise generally in many physical applications, including the question of time-reversed quantum mechanics [35] and quantum trajectories, as well as in stochastic control problems and inverse scattering. We show how to treat these problems in a unified way, using stochastic evolution in an additional dimension. This must converge to a steady state, so the time to equilibrate is important. Numerical examples indicate that such convergence issues are straightforward to control; they are also found in stochastic quantization [36].

As anticipated by Schrödinger, there is a general time-reversal symmetry in stochastic bridges. This is not found in one-directional stochastic differential equations. The symmetry is proved analytically, by utilizing both forwards and backwards form of the underlying diffusion equations. An explanation is given in terms of the use of time-symmetric path integral propagators. These results generalize and extend earlier results in the literature for evaluating stochastic bridges using stochastic partial differential equations. Analytic and numerical examples are given in simple cases, which show rapid convergence in the additional dimension.

II. WEIGHTED STOCHASTIC EQUATIONS

In this section, we present the definitions and notation underpinning the algorithm obtained later. While the results explained here are mostly well known, it is useful to summarize them in a unified form. For generality and definiteness, we are interested in stochastic time-evolution problems that either start from a specific configuration $\boldsymbol{x}_A$, or else terminate at a specific configuration $\boldsymbol{x}_B$. These are known as forward and backwards time problems, respectively.

A. The weighted Fokker-Planck equation

We will consider a general form of weighted FPE [37] in which there are zeroth-, first-, and second-order derivative terms, for a probability density $P(t,\boldsymbol{x})$ in an n-dimensional real vector space of coordinates $\boldsymbol{x} = (x^1, \ldots x^n)^T$. The relevant equation has the form

$$\frac{\partial}{\partial t} P(t,\boldsymbol{x}) = \mathcal{L}[\boldsymbol{x}] P(t,\boldsymbol{x}). \tag{1}$$

where $\mathcal{L}[\boldsymbol{x}]$ is a general second-order differential operator:

$$\mathcal{L}[\boldsymbol{x}] = U(\boldsymbol{x}) - \partial_i A^i(\boldsymbol{x}) + \tfrac{1}{2}\partial_i \partial_j \mathcal{D}^{ij}. \tag{2}$$

We use the Einstein summation convention, in which repeated indices are summed. Derivative operators are defined using the notation that $\partial_i \equiv \partial/\partial x^i$ and are assumed to act on all terms to the right. Here, U, $\boldsymbol{A}$, and $\boldsymbol{\mathcal{D}}$ are scalar, vector, and matrix quantities, respectively. We suppose that $\boldsymbol{\mathcal{D}}$ is positive definite and symmetric, and for the purposes of this paper, constant in space. Cases where partial integration may fail owing to the size of distribution tails or singular coefficients are excluded. These will not be investigated, which means that there are restrictions on the growth rate of the coefficients at large $|\boldsymbol{x}|$ [38].

In the case where $U(\boldsymbol{x}) = 0$, probability is conserved, and one can probabilistically sample to obtain an equivalent Ito SDE for the sampled path trajectories. This is

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{\mathcal{B}}\boldsymbol{\zeta}(t). \tag{3}$$

As usual, the noise coefficient $\boldsymbol{\mathcal{B}}$ is an $n \times m$ matrix such that $\boldsymbol{\mathcal{D}} = \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^T$, and $\boldsymbol{\zeta} = (\zeta^1, \ldots \zeta^m)^T$ is an m-dimensional real Gaussian noise vector, defined such that

$$\langle \zeta^i(t)\zeta^j(t')\rangle = \delta^{ij}\delta(t-t'). \tag{4}$$

When convenient, we use the difference notation, that $\boldsymbol{W} = \int_0^t \boldsymbol{\zeta}(t')dt'$, and the equation becomes

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x})dt + \boldsymbol{\mathcal{B}} \cdot d\boldsymbol{W}. \tag{5}$$

Unlike the standard FPE, the weighted equation described by Eq. (1) does not generally result in a conserved probability if $U(\boldsymbol{x}) \neq 0$. One can transform the more general weighted FPE Eq. (1) into a stochastic form like Eq. (3), by weighting moments of the distribution with a factor e^{Ω}, where the additional weight Ω evolves in time along each trajectory according to

$$d\Omega = U(\boldsymbol{x})dt. \tag{6}$$

Given an initial condition $\boldsymbol{x} = \boldsymbol{x}_A$ at $t = t_A$, any subsequent moment of form $O(\boldsymbol{x})$ can be evaluated as a weighted average. This is obtained by considering a set of initial trajectories starting at $\boldsymbol{x} = \boldsymbol{x}_A$ with equal weights, then summing over N independent trajectories $(\boldsymbol{x}^{(1)}, \ldots \boldsymbol{x}^{(N)})$ with weights $(\Omega^{(1)}, \ldots \Omega^{(N)})$ at a subsequent time t, and taking the limit of $N \to \infty$, so that

$$\langle O\rangle_t = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} O(\boldsymbol{x}^{(i)}(t)) e^{\Omega^{(i)}(t)}. \tag{7}$$

Such weighted averaging is not efficient as a numerical algorithm when there is a large range of weights Ω [20]. The question of sampling efficiency is largely independent of the question of convergence as a function of step-size. However, it cannot be readily solved by just using more trajectories, as this only delays the time when the sampling error growth creates large errors. The issue of how to efficiently solve weighted equations without a large sampling error will be investigated here.

B. Backwards Kolmogorov equation

A common case of a weighted equation is the backward Kolmogorov equation (BKE) [9], which is widely used to treat problems such as first passage times [3]. To understand the terminology, we recall that the standard Fokker-Planck equation with $U(\boldsymbol{x}) = 0$ is also called the forward Kolmogorov

equation, since it predicts a relative probability forwards in time. If all stochastic trajectories commence at an initial time t_A and point $\boldsymbol{x}_A$, then the probability of moving forwards to a time t, reaching a neighborhood of volume $dV \equiv d^n\boldsymbol{x}$ around a point $\boldsymbol{x}_B$, is given by $P(t,\boldsymbol{x}_B)dV \equiv P(t,\boldsymbol{x}_B|t_A,\boldsymbol{x}_A)dV$. This is formally the solution to Eq. (1), given the initial condition at time t_A that $P(t_A,\boldsymbol{x}) \equiv \delta^N(\boldsymbol{x}-\boldsymbol{x}_A)$, and with $t > t_A$. The probability is normalized to unity, since all trajectories have to terminate somewhere, and the probability is defined relative to the current trajectory location at the latest time t.

One can also enquire about the conditional probability density of trajectories *starting* at a location $\boldsymbol{x}$, given a known *final* destination $\boldsymbol{x}_B$. This is obtained by first solving the backwards Kolmogorov equation subject to a fixed final condition that $P(t_B,\boldsymbol{x}) \equiv \delta^N(\boldsymbol{x}-\boldsymbol{x}_B)$ at time t_B, which is

$$-\frac{\partial}{\partial t}P(t_B,\boldsymbol{x}_B|t,\boldsymbol{x}) = \mathcal{L}^*[\boldsymbol{x}]P(t_B,\boldsymbol{x}_B|t,\boldsymbol{x}), \tag{8}$$

where $\mathcal{L}^*[\boldsymbol{x}]$ is an adjoint operator defined by

$$\mathcal{L}^*[\boldsymbol{x}] = U(\boldsymbol{x}) + A^i(\boldsymbol{x})\partial_i + \tfrac{1}{2}\partial_i\mathcal{D}^{ij}\partial_j. \tag{9}$$

This function is *not* normalized to unity when integrated over the initial location $\boldsymbol{x}$, unless the coefficients are independent of position. The reason for this is that the probability $P(t,\boldsymbol{x})$ occurring in the normal Fokker-Planck approach is defined as a final state probability density, so it is only guaranteed to be normalized when integrated over the final time.

To transform the BKE into the standard form used elsewhere, we introduce a frame change to a time-reversed coordinate, $t' = t_B - t$. On defining a BKE function of reverse time, $P_{\rm BKE}(t',\boldsymbol{x}) \equiv P(t_B,\boldsymbol{x}_B|t_B-t',\boldsymbol{x})$, one obtains

$$\frac{\partial}{\partial t'}P_{\rm BKE}(t',\boldsymbol{x}) = \mathcal{L}^*[\boldsymbol{x}]P_{\rm BKE}(t',\boldsymbol{x}). \tag{10}$$

Next, $\mathcal{L}^*[\boldsymbol{x}]$ can be re-expressed in the form of Eq. (2), with the result that

$$\mathcal{L}^*[\boldsymbol{x}] = U'(\boldsymbol{x}) + \partial_i A^i(\boldsymbol{x}) + \tfrac{1}{2}\partial_i\mathcal{D}^{ij}\partial_j, \tag{11}$$

where there is now a modified potential of $U'(\boldsymbol{x}) = U(\boldsymbol{x}) - J$.

Here, the term J is the trace of the Jacobian of the drift, i.e., $J = Tr(\boldsymbol{J})$, where $J^i_j = \partial_j A^i$. To understand the implications of this result, we note that in the forward time direction, if $U = 0$, the equivalent stochastic equation is

$$\frac{\partial\boldsymbol{x}}{\partial t} = \boldsymbol{A}(\boldsymbol{x}) + (\text{stochastic terms}), \tag{12}$$

while in the backward time direction, provided the diffusion is constant, there is a modified weight $U'(\boldsymbol{x})$, and the drift terms change sign:

$$\frac{\partial\boldsymbol{x}}{\partial t'} = -\boldsymbol{A}(\boldsymbol{x}) + (\text{weight and stochastic terms}). \tag{13}$$

The time-reversed deterministic behavior in the drift term $\boldsymbol{A}$ is as expected from the fact that there is a change in reference frame from t to $t' = -t$. However, when the drift coefficients are not constant the additional weighting terms have subtle effects. In summary, for nonlinear drift the BKE trajectories have changed relative probabilities as well as being time-reversed. This means that the usual stochastic techniques for sampling require additional weighting terms. This makes them very inefficient unless the algorithm is modified substantially.

III. TIME-REVERSED INFERENCE

For an illustration of the applications of the BKE, we now obtain the time-reversed probability of *leaving* from a small volume dV near location $\boldsymbol{x}$, given a known final location $\boldsymbol{x}_B$. In typical applications, given a final condition $\boldsymbol{x} = \boldsymbol{x}_B$ at $t = t_B$, we wish to evaluate earlier observable moment expectations. This is one of many conditional problems in probability theory.

A. Time-reversed diffusion

To obtain the required result, we apply Bayes' theorem [39]. The quantity given by the BKE is $P(t_B,\boldsymbol{x}_B|t,\boldsymbol{x})dV$, the conditional probability that a trajectory is at final location near $\boldsymbol{x}_B$, if initially near $\boldsymbol{x}$. We wish to obtain the complement of this, namely the conditional probability that a trajectory is initially near $\boldsymbol{x}$, given a known final location at $\boldsymbol{x}_B$.

Bayes' theorem gives the result that the *joint* probability of these two events at $\boldsymbol{x}$, $\boldsymbol{x}_B$ is

$$\begin{aligned} P(t_A,\boldsymbol{x};t_B,\boldsymbol{x}_B)dV^2 &= P(t_B,\boldsymbol{x}_B|t_A,\boldsymbol{x})p(t_A,\boldsymbol{x})dV^2 \\ &= P(t_A,\boldsymbol{x}|t_B,\boldsymbol{x}_B)p(t_B,\boldsymbol{x}_B)dV^2, \end{aligned} \tag{14}$$

where $P(t,\boldsymbol{x}|t_B,\boldsymbol{x}_B)$ is the conditional probability for $t < t_B$, that a trajectory comes from an initial location near $\boldsymbol{x}$, if it terminates near $\boldsymbol{x}_B$.

Here we have also introduced *a priori* probabilities $p(t_A,\boldsymbol{x})dV$, $p(t_B,\boldsymbol{x})dV$ of the two events at t_A and t_B. In the case of a constant probability C of trajectory starting points, $p(t_A,\boldsymbol{x})dV = C$. Noting that for conserved probabilities, $\int P(t_A,\boldsymbol{x}|t_B,\boldsymbol{x}_B)d^n\boldsymbol{x} = 1$, we obtain

$$p(t_B,\boldsymbol{x}_B) = C\int P(t_B,\boldsymbol{x}_B|t_A,\boldsymbol{x})d^n\boldsymbol{x}. \tag{15}$$

Hence, on canceling the constant factors of C, the time-reversed conditional probability is

$$P(t_A,\boldsymbol{x}|t_B,\boldsymbol{x}_B) = \frac{P(t_B,\boldsymbol{x}_B|t_A,\boldsymbol{x})}{\int P(t_B,\boldsymbol{x}_B|t_A,\boldsymbol{x})d^n\boldsymbol{x}}. \tag{16}$$

Algorithms for sampling the weighted Fokker-Planck equation allow one, among other things, to sample the BKE. This gives a useful strategy for generating a time-reversed stochastic process. In the next section, we show how to use these results to derive an algorithm for a more general problem known as the stochastic bridge.

B. Time-reversed linear problems

As an example of these applications of the BKE, we consider linear stochastic equations, both forward and backward in time. The previous results show that one example of the general weighted Fokker-Planck equation is a backward Kolmogorov equation, since it has an equivalent weighted trajectory. However, in some cases there can also be an equivalent, *unweighted* set of trajectories.

This is trivially the case for a forward time, linear stochastic equation, of form

$$d\boldsymbol{x} = \boldsymbol{J}\boldsymbol{x}dt + \boldsymbol{B}d\mathbf{W}(t), \tag{17}$$

where $\boldsymbol{J}$ is a constant $n \times n$ matrix, and $\boldsymbol{B}$ is also constant in this example.

The fact that the time-reversed equation is also unweighted follows, since the results given above show that the weight term $U'(\boldsymbol{x}) = -J$ is the same for each trajectory. It therefore cancels on calculating the conditional time-reversed probability. Thus, the time-reversed stochastic trajectory in the linear case is simply

$$d\boldsymbol{x} = -\boldsymbol{J}\boldsymbol{x}dt' + \boldsymbol{B}d\mathbf{W}(t'). \tag{18}$$

In summary, with constant diffusion and linear drift the effect of time-reversal gives the known result that it reverses the drift sign, while keeping the diffusion term unchanged. This result is also known to be valid from an earlier analysis of linear stochastic time-reversal [40].

IV. STOCHASTIC BRIDGES

We defined a *forward time* stochastic process as the relative probability of trajectories that start definitely at $\boldsymbol{x}_A$, while ending in a neighborhood dV of $\boldsymbol{x}_B$. Similarly, we defined a *time-reversed* stochastic process as the relative probability of trajectories that end definitely at $\boldsymbol{x}_B$, while starting in a neighborhood dV of $\boldsymbol{x}_A$. Given the discussion above, these are generally different problems. From the probabilistic viewpoint, they are simply two different, abstract conditional probabilities, $P(B|A)$ and $P(A|B)$, which are not the same.

There are more general types of multitime conditional probabilities as well. In this section we will be interested in stochastic trajectories constrained at two end-points A, C, while being unconstrained at an intermediate point B. We therefore define a stochastic bridge as the probability $P(B|A,C)$, i.e., the probability of obtaining an intermediate point $\boldsymbol{x} = \boldsymbol{x}_B$, with constraints on initial and final points $\boldsymbol{x}_A, \boldsymbol{x}_C$.

Similar results were previously obtained for simpler cases, including scalar variable [15,41] and potential problems [14]. In these papers, the drift term $\boldsymbol{A}$ was obtained as the gradient of a potential V, and the diffusion was constant in space. As a consequence, these algorithms are unable to treat the general class of problems dealt with here, which include arbitrary drift *without* a potential, as well as weighted trajectories and time-reversed stochastic paths.

A. Path integral solution

To obtain results for stochastic bridges, we first obtain an efficient algorithm for the general weighted stochastic trajectories described by Eq (1). This includes the case of the BKE. It gives a technique for propagating time-reversed stochastic trajectories in nonlinear cases more general than in Eq (17), and a technique for treating weighted stochastic processes. To do this, we write the general solution in its path-integral form [42] and transform the results to a stochastic partial differential equation in a higher dimension.

Only the case where $\mathcal{D}^{ij}$ is invertible will be treated. Although this restriction is not essential, it simplifies the method, and reduces the algebraic complexity. In these cases, the path-integral result is straightforward. There have been a number of well-known path-integral formulations of diffusion processes, pioneered by Onsager and Machlup [43,44], although this early work was restricted to linear equations. Later developments due to Stratonovich [45] and others [30,31,46,47] extended and made this early work more rigorous. As known from this previous work, there is an analogy to a covariant or curved-space path integral, where $\boldsymbol{g} \equiv \boldsymbol{\mathcal{D}}^{-1}$ plays the role of a metric tensor. While a full covariant analysis is only needed when the diffusion is space-dependent, which will be treated elsewhere, we will utilize this definition of $\boldsymbol{g}$ here.

1. Green's function approach

The simplest derivation, corresponding to the earliest work of this type, used a convolution of Greens functions. The Green's function of the FPE for short times $\Delta t = \epsilon$ is the solution to an initial delta function condition of $\delta(\boldsymbol{x} - \boldsymbol{x}_A)$. Defining a relative velocity $\boldsymbol{v} = (\boldsymbol{x} - \boldsymbol{x}_A)/\epsilon - \boldsymbol{A}(\boldsymbol{x}_A)$ compared to the average drift velocity $\boldsymbol{A}$, the solution is [37]

$$G(\boldsymbol{x}|\boldsymbol{x}_A) = \frac{1}{\mathcal{N}(\epsilon,\boldsymbol{x}_A)} e^{\epsilon[-\boldsymbol{v}^T\boldsymbol{g}\boldsymbol{v}/2 + U(\boldsymbol{x})]}. \tag{19}$$

The normalization term is $\mathcal{N}(\epsilon,\boldsymbol{x}_A) = \sqrt{\det[2\pi\epsilon\boldsymbol{D}(\boldsymbol{x}_A)]}$, and since $\boldsymbol{D}$ is positive-definite, it has positive eigenvalues and a positive determinant. Noninvertible cases where some eigenvalues vanish are treated by taking a limit that turns the Gaussian into a δ function. An orthogonal local coordinate transformation can always be found such that a symmetric diffusion can be diagonalized, but we will not assume this.

The path integral is a convolution of successive Green's functions. Consider a discrete set of $N+1$ times $t_k = t_A + k\epsilon$ for $k = 1, N+1$ where $t_C = t_{N+1}$, and corresponding coordinates $\boldsymbol{x}_l$, with $\boldsymbol{x}_A = \boldsymbol{x}_0$, $\boldsymbol{x}_C = \boldsymbol{x}_N$. Defining $d[\boldsymbol{x}] \equiv d^n\boldsymbol{x}_1 \ldots d^n\boldsymbol{x}_N$, the solution over the finite interval $[t_A,t_C]$, given an initial distribution $P(t_A,\boldsymbol{x}) = \delta^n(\boldsymbol{x} - \boldsymbol{x}_A)$, is then

$$P(t_B,\boldsymbol{x}_C) = \lim_{\epsilon\to 0} \int .. \int d[\boldsymbol{x}] G(\boldsymbol{x}_C|\boldsymbol{x}_{N-1}) .. G(\boldsymbol{x}_1|\boldsymbol{x}_A). \tag{20}$$

This form of the path integral is equivalent to the well-known Ito-Euler or forward-time numerical algorithm [18] for a stochastic equation, but with additional weights now included from the potential $U(\boldsymbol{x})$. We note that the right-hand side explicitly includes $\boldsymbol{x}_C$, and it depends implicitly on t_C through the relation that $t_A = t_C - (N+1)\epsilon$.

The results for the case of constant diffusion with no drift correspond to a standard quantum mechanical Feynman path integral in imaginary time [48]. In more general cases, this form of path integral is *not* time-symmetric, nor is it invariant under coordinate transformations. This makes it unsuitable for the functional calculus methods we will use later, as Feynman also recognized [49].

2. Time-symmetric path-integral

A time-symmetric treatment of path integrals was originally obtained by Stratonovich for the case of constant diffusion [45]. To summarize this method for the constant diffusion case,

one may take the limit of $\epsilon \to 0$ with functions evaluated at the midpoints, not the starting points of each interval [33,50]. This allows the path integral to be treated as a normal functional integral.

In the case where $\mathcal{D}$ is constant, one may replace $\boldsymbol{A}(\boldsymbol{x})$ by $\boldsymbol{A}(\bar{\boldsymbol{x}})$, and $U(\boldsymbol{x})$ by

$$\bar{U} = U(\bar{\boldsymbol{x}}) - \tfrac{1}{2}J, \tag{21}$$

where $\bar{\boldsymbol{x}} = (\boldsymbol{x} + \boldsymbol{x}')/2$ and $J = \partial_i A^i(\bar{\boldsymbol{x}})$. This gives the same formal path integral result as Eq. (20) in the limit of $\epsilon \to 0$, except for the additional spatial derivative term in the potential. On introducing $\bar{\boldsymbol{v}} = (\boldsymbol{x}' - \boldsymbol{x})/\epsilon - \boldsymbol{A}(\bar{\boldsymbol{x}})$, and after some manipulation [33], one obtains a corrected propagator:

$$\bar{G}(\boldsymbol{x}'|\boldsymbol{x}) = \frac{1}{\mathcal{N}(\epsilon)} e^{\epsilon[-\bar{\boldsymbol{v}}^T \boldsymbol{g}\bar{\boldsymbol{v}}/2 + \bar{U}(\boldsymbol{x})]}. \tag{22}$$

The final path integral is written—after the limit $\epsilon \to 0$ is taken—as

$$P(t_B,\boldsymbol{x}_B) = \int d\mu[\boldsymbol{x}(t)] \exp\left\{-\int_{t_A}^{t_C} L(\boldsymbol{x},\dot{\boldsymbol{x}})dt\right\}, \tag{23}$$

where the functional measure is defined as

$$d\mu[\boldsymbol{x}(t)] = \lim_{\epsilon\to 0} \frac{d[\boldsymbol{x}]}{\mathcal{N}(\epsilon)^N}, \tag{24}$$

the end-points are constrained so that $\boldsymbol{x}(t_A) = \boldsymbol{x}_A$ and $\boldsymbol{x}(t_C) = \boldsymbol{x}_C$, and $\dot{\boldsymbol{x}} \equiv \partial\boldsymbol{x}/\partial t$. The effective path-integral Lagrangian in the time-integral can now be written as

$$L(\boldsymbol{x},\dot{\boldsymbol{x}}) = \tfrac{1}{2}(\dot{\boldsymbol{x}} - \boldsymbol{A}(\boldsymbol{x}))^T \boldsymbol{g}[\dot{\boldsymbol{x}} - \boldsymbol{A}(\boldsymbol{x})] - \bar{U}(\boldsymbol{x}). \tag{25}$$

Here the assumption is made that all functions are smooth, so that in the limit of $\epsilon \to 0$, one has $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}(\bar{\boldsymbol{x}})$. However, the potential is modified since $\bar{U}(\boldsymbol{x}) \neq U(\boldsymbol{x})$, owing to the Jacobian correction term in Eq. (21).

3. *Time-reversed path-integrals*

In the time reversed or BKE case, we must replace $\boldsymbol{A}$ by $-\boldsymbol{A}$, and U by $U' = U - J$. Taking into account the derivation explained above and applying it in the reverse time direction, this would then give a Lagrangian in which the final potential term $\bar{U}' = \bar{U}$ is unchanged, giving

$$L' = \frac{1}{2}\left(\frac{\partial x^i}{\partial t'} + A^i\right) g_{ij} \left(\frac{\partial x^j}{\partial t'} + A^j\right) - \bar{U}. \tag{26}$$

If we rewrite the path integral and its derivatives as a function of forward time, so $t' \to t \equiv -t'$, then this leads to a form for the Lagrangian, which is the same as for the FPE, namely, Eq. (25).

Thus, the Lagrangian is unchanged whether it is derived from the FPE or the BKE, provided the final integrals are expressed in a compatible form, with the same time coordinate. To explain this, the FPE gives the probability density of arriving near $\boldsymbol{x}_B$, having started at a fixed point, $\boldsymbol{x}_A$. Conversely, the BKE allows us to investigate probability densities of starting near $\boldsymbol{x}_A$, conditional on arriving at a fixed point, $\boldsymbol{x}_B$. Since it is the final position that is known, it is normally written as an equation in reverse time, t'. If we write the path integral formulation of these probability densities so they are expressed in the same time-coordinate, then the final Lagrangians are identical.

B. Path integrals and stochastic bridges

Path integral trajectories can have constrained intermediate and/or end points. Suppose that we consider a trajectory in the forward time direction, starting at $\boldsymbol{x}_A$, passing through $\boldsymbol{x}_B$ and ending at $\boldsymbol{x}_C$, thus giving a Dirichlet type fixed boundary condition. In this case we simply have to evaluate the probability $P(B)$ that the trajectory passes through $\boldsymbol{x}_B$.

We therefore wish to sample the path integral as a functional of coordinate space and time, i.e., to sample paths $[\boldsymbol{x}] = [\boldsymbol{x}(t)]$, with a constrained functional probability that is given by

$$\mathcal{P}([\boldsymbol{x}]) = \exp\left\{-\int_{t_A}^{t_C} L(\boldsymbol{x},\dot{\boldsymbol{x}})dt\right\}, \tag{27}$$

under the constraint that $\boldsymbol{x}(t_A) = \boldsymbol{x}_A$, and $\boldsymbol{x}(t_C) = \boldsymbol{x}_C$.

The goal is to obtain a set of equally weighted stochastic trajectories for sampling purposes. If $U = 0$, then a SDE algorithm with independent variable t will achieve this, although the end points cannot be treated efficiently: only a set of measure zero will satisfy the constraints. If $U \neq 0$, then even a set that meets these constraints will have different weights, in general, make sampling inefficient.

One method to proceed when there are constraints, is to define a new stochastic process in a virtual "time" coordinate τ, such that its steady-state solution for $\tau \to \infty$ is $\mathcal{P}([\boldsymbol{x}])$. This is achieved using a technique similar to that employed in stochastic quantization [36]. We can define a new probability functional $\mathcal{P}([\boldsymbol{x}],\tau)$ in an extra dimension τ, with the steady-state requirement that

$$\lim_{\tau\to\infty} \mathcal{P}([\boldsymbol{x}],\tau) = \mathcal{P}([\boldsymbol{x}]). \tag{28}$$

The steady-state solution at $\tau \to \infty$, is required to have a potential form as a functional integral over paths $\boldsymbol{x}'(t)$ that have their end-points constrained so that $\boldsymbol{x}(t_A) = \boldsymbol{x}_A$ and $\boldsymbol{x}(t_B) = \boldsymbol{x}_B$:

$$\mathcal{P}([\boldsymbol{x}(t)],\tau \to \infty) \propto \exp\left\{\int^{\boldsymbol{x}(t)} \boldsymbol{\alpha}([\boldsymbol{x}'])d[\boldsymbol{x}']\right\}. \tag{29}$$

From direct functional differentiation of Eq. (29), and applying the condition that

$$\frac{\partial}{\partial\tau}\mathcal{P}([\boldsymbol{x}],\tau \to \infty) = 0, \tag{30}$$

this is achieved provided $\mathcal{P}([\boldsymbol{x}],\tau)$ satisfies a functional Fokker-Planck equation in the form

$$\frac{\partial}{\partial\tau}\mathcal{P}([\boldsymbol{x}],\tau) = \frac{\delta}{\delta\boldsymbol{x}}\left[-\boldsymbol{\alpha}[\boldsymbol{x}] + \frac{\delta}{\delta\boldsymbol{x}}\right]\mathcal{P}([\boldsymbol{x}],\tau). \tag{31}$$

Since the solution of the new functional FPE must agree with the required path-integral in Eq. (27), it follows directly that

$$\boldsymbol{\alpha}([\boldsymbol{x}]) = -\frac{\delta}{\delta\boldsymbol{x}}\int_{t_A}^{t_C} L(\boldsymbol{x},\dot{\boldsymbol{x}})dt. \tag{32}$$

The full applicability of methods like this are under active investigation in the mathematical and probability literature [14]. To obtain the partial stochastic differential equation coefficient $\boldsymbol{\alpha}[\boldsymbol{x}]$, it follows from Eq. (31) that the functional $\boldsymbol{\alpha}[\boldsymbol{x}]$ is given by

$$\begin{aligned}\alpha_k[\boldsymbol{x}(t)] &= -\frac{\delta}{\delta x^k(t)}\int L(\boldsymbol{x},\dot{\boldsymbol{x}})dt \\ &= -\frac{\partial L}{\partial x^k(t)} + \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_k(t)}.\end{aligned} \tag{33}$$

Substituting into Eq. (25), and dropping the time arguments for brevity, one obtains that

$$\alpha_k[\boldsymbol{x}] = \boldsymbol{v}^T\boldsymbol{g}\frac{\partial \boldsymbol{A}}{\partial x^k} + \frac{\partial \bar{U}}{\partial x^k} + g_{kl}\dot{v}_l, \tag{34}$$

where $\boldsymbol{v} = [\dot{\boldsymbol{x}} - \boldsymbol{A}(\boldsymbol{x})]$. We now re-express the coefficients of Eq. (34) in a simpler form. Introducing second derivatives, $\ddot{\boldsymbol{x}} \equiv \partial^2\boldsymbol{x}/\partial t^2$, one then obtains a compact expression for $\boldsymbol{\alpha}$ in terms of time-derivatives of the trajectory location, namely,

$$\boldsymbol{\alpha} = \boldsymbol{g}\ddot{\boldsymbol{x}} + \boldsymbol{\mathcal{C}}\dot{\boldsymbol{x}} + \boldsymbol{U}. \tag{35}$$

Here the circulation matrix,

$$\boldsymbol{\mathcal{C}} = \boldsymbol{J}^T\boldsymbol{g} - \boldsymbol{g}\boldsymbol{J}, \tag{36}$$

describes the extent to which the drift $\boldsymbol{A}$ violates potential conditions. To be more explicit, we can write the circulation term with its components as

$$\mathcal{C}_{kj} = \frac{\partial A^i}{\partial x^k}g_{ij} - g_{kl}\frac{\partial A^l}{\partial x^j}. \tag{37}$$

We also introduce the vector $\boldsymbol{U}$, which is an effective drift rate in the virtual time coordinate, where

$$U_k = \frac{\partial}{\partial x^k}\left[\bar{U} - \frac{1}{2}\boldsymbol{A}^T\boldsymbol{g}\boldsymbol{A}\right]. \tag{38}$$

C. Stochastic equation in virtual time

The distribution of Eq. (29) solves the problem of the stochastic bridge in the form treated here. We now wish to sample this probability distribution efficiently, with equally weighted samples. This is achieved by noting that the functional Fokker-Planck Eq. (31) is exactly equivalent to a partial stochastic differential equation (PSDE) in virtual time,

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \boldsymbol{\alpha}([\boldsymbol{x}]) + \sqrt{2}\boldsymbol{\xi}(t,\tau). \tag{39}$$

It has noise correlations given by

$$\langle \xi^i(t,\tau)\xi^j(t',\tau')\rangle = \delta^{ij}\delta(t-t')\delta(\tau-\tau'),$$

and transverse Dirichlet boundary conditions at t_A, such that $\boldsymbol{x} = \boldsymbol{x}_A$, and at t_C such that $\boldsymbol{x} = \boldsymbol{x}_C$.

Writing the PSDE in its explicit form, one has

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \boldsymbol{g}\ddot{\boldsymbol{x}} + \boldsymbol{\mathcal{C}}\dot{\boldsymbol{x}} + \boldsymbol{U} + \sqrt{2}\boldsymbol{\xi}(t,\tau). \tag{40}$$

The path-integral is then solved by taking the steady-state or $\tau \to$ limit of the distribution of this stochastic process, using a numerical method for a PSDE [51] with Dirichlet boundary conditions. To compare this with other expressions in the literature, we notice that previous analyses [14] have considered cases without weights ($U = 0$) and with a drift potential V such that

$$g_{ji}A^i = -\frac{\partial V}{\partial x^j}. \tag{41}$$

This places an immediate constraint on the drift A^i, known as the drift potential condition. Differentiating Eq. (41) with respect to x^k, and using the equality of second derivatives under change of order of differentiation, leads to

$$g_{ji}J_{i,k} = g_{ki}J_{i,j}. \tag{42}$$

This is the condition that the circulation term vanishes, so that $\boldsymbol{\mathcal{C}} = 0$. Thus, the circulation term $\boldsymbol{\mathcal{C}}$ contains all the nonpotential contribution to the drift. Under the restriction that $U = \boldsymbol{\mathcal{C}} = 0$, our results are identical to the previous ones. The case that $U = 0$ but $\boldsymbol{\mathcal{C}} \neq 0$ was conjectured to give the result of Eq. (40) in earlier literature [14], although using an abstract notation. Our result gives a derivation for this conjecture, makes the notation explicit, and extends it to weighted stochastic processes.

D. Time-reversed stochastic bridges

The time-reversed path originates from constraining the future time phase-space position, resulting in the backwards Kolmogorov equation, which has an additional weight term. We can use the path integral method to constrain the earlier time as a time-reversed stochastic bridge, reversing the procedures in the previous subsection. We wish to demonstrate that the time-ordering of imposing the constraints doesn't matter, which was the motivation of Schrödinger's original work on "Umkehrung" or time-reversal.

Because $\bar{U}' = \bar{U}$, we see from Eq (11) that the virtual stochastic equation for the time-reversed path $\boldsymbol{x}'(t')$ is:

$$\frac{\partial \boldsymbol{x}'}{\partial \tau} = \boldsymbol{g}\frac{\partial^2\boldsymbol{x}'}{\partial t'^2} + \boldsymbol{\mathcal{C}}'\frac{\partial \boldsymbol{x}'}{\partial t'} + \boldsymbol{U} + \sqrt{2}\boldsymbol{\xi}(t',\tau). \tag{43}$$

It is instructive to write this using a variable change to the forward time coordinate, with $\boldsymbol{x}(t) = \boldsymbol{x}'(t') = \boldsymbol{x}'(t_B - t)$, and $\partial\boldsymbol{x}/\partial t = -\partial\boldsymbol{x}'/\partial t'$. This is the same coordinate system as the usual FPE, and one might expect to therefore get the same bridge equations. After all, once written in the same coordinate frame, each describes an identical overall constraint on the bridge. We note that $dt' = -dt$ in the time-reversed frame of reference, and $A^i = -A'^i$ when using the BKE, so that the result of the variable change is to only change the sign of the circulation term. Since $\boldsymbol{C}' = -\boldsymbol{C}$, this has no effect: all terms match once the sign change in the time-coordinate is taken into account. We obtain a partial stochastic equation identical to the previous form. In summary, these equations are independent of whether the forward or backwards Kolmogorov equation chosen initially, provided the same time coordinate is utilized in the final equation.

This is as one might expect: neither end-point constraint is conditional on the other. The result given here satisfies Schrödinger's requirements with regard to time-reversibility, without having the product form that Schrödinger considered [1]. The agreement of the forward and reverse time results gives further evidence that this weighted, virtual-time stochastic equation is correct.

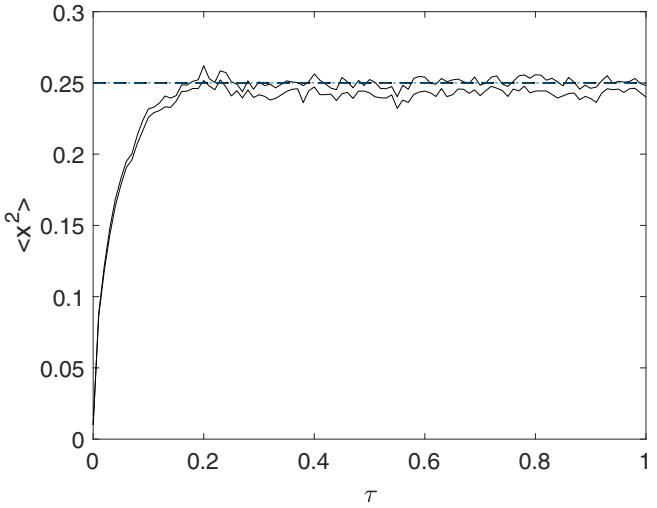

FIG. 1. Convergence of Brownian bridge variance (solid lines) to analytic result (dotted lines) with virtual time τ, at fixed real time of $t = 0$. Upper and lower lines indicate sampling errors of $\pm 1\sigma$.

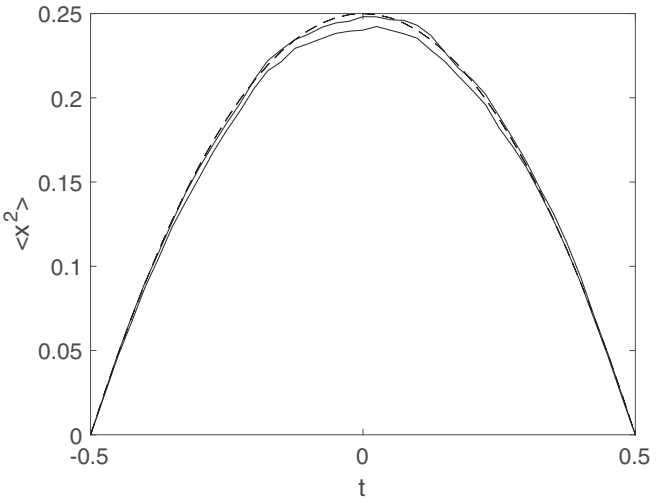

FIG. 2. Convergence of Brownian bridge variance (solid lines) to analytic result (dotted lines) with real time t, at a fixed virtual time $\tau = 1$. Upper and lower lines indicate sampling errors of $\pm 1\sigma$.

V. EXAMPLES

We now consider a number of examples of this approach, both to give a comparison with previous analyses and to show where there are new results.

A. The scalar Brownian bridge

The first example we treat is the classic, scalar Brownian bridge, in which the unconstrained stochastic process is

$$dx = dW. \tag{44}$$

It is known [52] that the constrained problem of $x^1 = x^2 = 0$ is solved by the random trajectory

$$x(t) = W(t) - \frac{t}{T} W(T), \tag{45}$$

where $x(0) = x(T) = 0$, and $W(t) = \int_0^t dW(t)$.

In this case one can sample the stochastic trajectory directly. The forward and backward bridges are identical, and one can easily compute moments at intermediate times, including the variance,

$$\langle x^2(t) \rangle = \frac{t(T-t)}{T}. \tag{46}$$

Since there is neither drift nor potential terms, the partial stochastic equation method gives

$$\frac{\partial x}{\partial \tau} = \frac{\partial^2 x}{\partial t^2} + \sqrt{2}\xi. \tag{47}$$

B. Scalar bridge numerics

Solving this PSDE with fixed or Dirichlet boundary conditions at $t = \pm 0.5$ gives the results for $\langle x^2(t) \rangle$ shown in Figs. 1 and 2. This demonstrates rapid convergence with virtual time τ, and excellent agreement with the known result for the Brownian bridge variance at the midpoint, $t = 0$.

The results in Fig. 2 show that this agreement is maintained, within sampling error, over the full range of real time. Figure 3 shows a three-dimensional representation of the results.

These results used a public domain PSDE solver [53], using 2000 time steps in τ, 31 time steps in real time t, and 10^4 stochastic trajectories. The algorithm was an iterated semi-implicit or midpoint technique [17,51]. Central finite differences were used for the transverse derivatives with Dirichlet boundary condition. This algorithm can be accelerated using spectral methods [54]. To check convergence, the following procedure was used:

Sampling Error: A subensemble averaging technique was used, in which the stochastic trajectories were dividing into subensembles. After subensembles were averaged, the resulting means were averaged, and error-bars were estimated using standard deviations of the mean. These are shown as upper and lower lines in two-dimensional graphs.

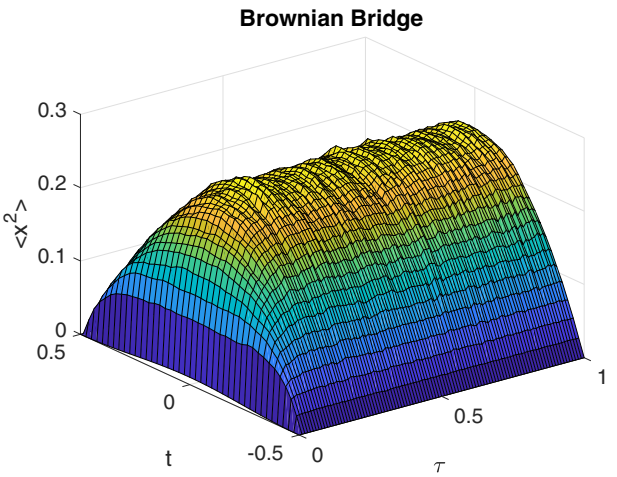

FIG. 3. Three-dimensional Brownian bridge variance surface evolving with real and virtual time.

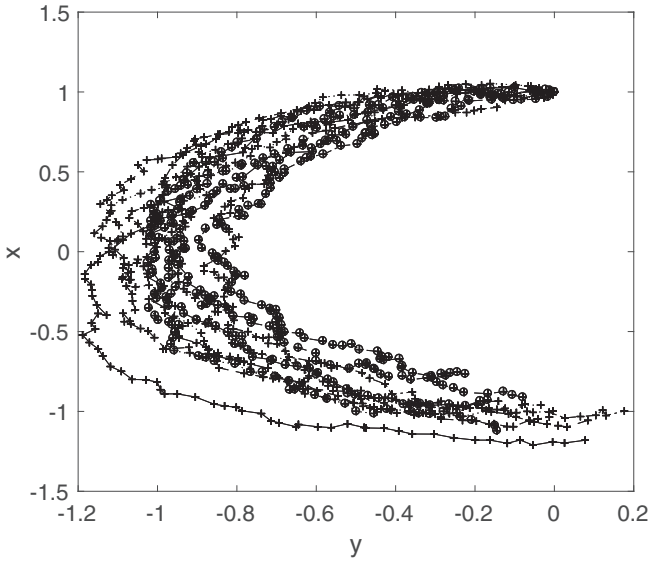

FIG. 4. Parametric plot of x versus y values for 10 random unconstrained stochastic trajectories for a linear stochastic differential equation with $\epsilon = 0.01$, from time $t = 0$ to $t = \pi$.

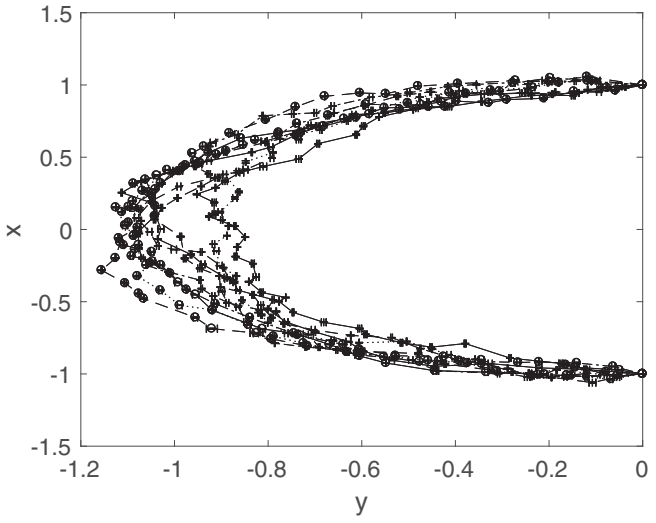

FIG. 5. Parametric plot of final x versus y values for 10 random trajectories at $\tau = 0.4$, for a linear stochastic bridge with $\epsilon = 0.01$, from time $t = 0$ to $t = \pi$.

Virtual time-step error: Calculations were repeated with twice the number of virtual time-steps to check virtual-time convergence. In all cases, error-bars were less than $\pm 1\%$.

Virtual time equilibration: Calculations were repeated with double the maximum virtual time window, to ensure convergence.

Real time-step error: Calculations were repeated with larger numbers of real time-steps to check real-time convergence in step-size.

C. Linear case

The next example we treat is a general linear FPE. To simplify the resulting algebra, we assume that $\boldsymbol{D} = \boldsymbol{g} = \epsilon \boldsymbol{\mathcal{I}}$, where $\boldsymbol{\mathcal{I}}$ is an identity matrix. This has an exact solution for the time-reversed process in which only the drift is time-reversed [40], so it is a useful test case. For this problem, we have a forward SDE of

$$d\boldsymbol{x} = \boldsymbol{J}\boldsymbol{x}dt + \sqrt{\epsilon}d\boldsymbol{W}, \tag{48}$$

with $\boldsymbol{A} = \boldsymbol{J}\boldsymbol{x}$, and $\boldsymbol{J}$ is a constant matrix.

This can be treated using the stochastic equation given above, although this will not include constraints. One may also sample it using the constrained path-integral method, which gives an equivalent partial stochastic differential equation of

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \frac{1}{\epsilon}[\ddot{\boldsymbol{x}} + (\boldsymbol{J}^T - \boldsymbol{J})\dot{\boldsymbol{x}} - \boldsymbol{J}^T\boldsymbol{J}\boldsymbol{x}] + \sqrt{2}\boldsymbol{\xi}. \tag{49}$$

In the reverse time direction, the corresponding BKE has $\tilde{\boldsymbol{A}} = -\boldsymbol{A}$, and a potential of $\tilde{U} = -\mathrm{Tr}[\boldsymbol{J}] = -J$. Since this is constant, one has $\boldsymbol{U}' = 0$, just as with the forward time case. The path integral expressions for the bridge are identical in the forward and backward directions, which illustrates the general time-reversal symmetry in the path-integral algorithm.

As a numerical example of linear evolution in higher dimensions without a drift potential, consider the case

$$\boldsymbol{J} = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \tag{50}$$

giving the equivalent stochastic differential equation of

$$\dot{x} = y + \sqrt{\epsilon}\zeta_1, \;\; \dot{y} = -x + \sqrt{\epsilon}\zeta_2. \tag{51}$$

This describes circular motion with no equivalent gradient, combined with random diffusion in the x-y plane. The bridge is from $\boldsymbol{x} = [1,0]$ to $\boldsymbol{x} = [-1,0]$, with $\epsilon = 0.01$, over a time interval of π. The equivalent virtual time bridge equation is

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \frac{1}{\epsilon}[\ddot{\boldsymbol{x}} - 2\boldsymbol{J}\dot{\boldsymbol{x}} - \boldsymbol{x}] + \sqrt{2}\boldsymbol{\xi}(t,\tau). \tag{52}$$

In the limit of small ϵ, the bridge solution has a deterministic equation that is similar to the motion of a charged particle in a magnetic field in the z direction, resulting in circular orbits in the x-y plane. This results in characteristic quasicircular bridges, if the initial and final constraints are chosen to occur on a circle with a fixed radius from the origin. Here, Fig. 4 gives a parametric plot for ten random stochastic trajectories starting at $\boldsymbol{x} = [1,0]$, while Fig. 5 graphs the constrained x, y values in the bridge. This shows the expected circular motion, which only occurs if the circulation term of $\boldsymbol{C} = -2\boldsymbol{J}$ is present.

In these figures, there were 1200 time steps using a semi-implicit midpoint method [17] for the stochastic equation, with negligible step-size errors. The bridge equations were solved with a fourth-order Runge-Kutta finite-difference method with 4×10^4 virtual time-steps and 40 real-time steps. The accompanying figures give solutions of the PSDE at the final value of virtual time, in this case $\tau = 0.4$. There was little or no difference found by using $\tau = 0.2$, indicating a steady state.

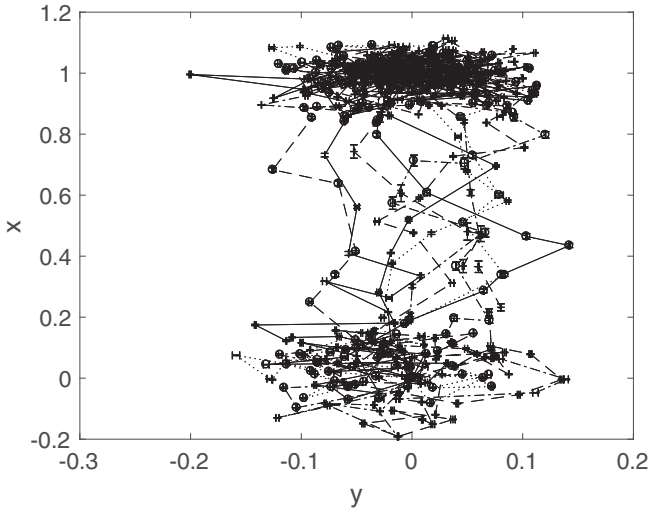

FIG. 6. Ten typical stochastic bridge trajectories, showing evolution of x,y during the constrained transition from a stable to a metastable point for the Maier-Stein equations with $\alpha = 1$, $\mu = 1$, and $\epsilon = 0.01$. Here, $\tau = 1$ and $t = 20$.

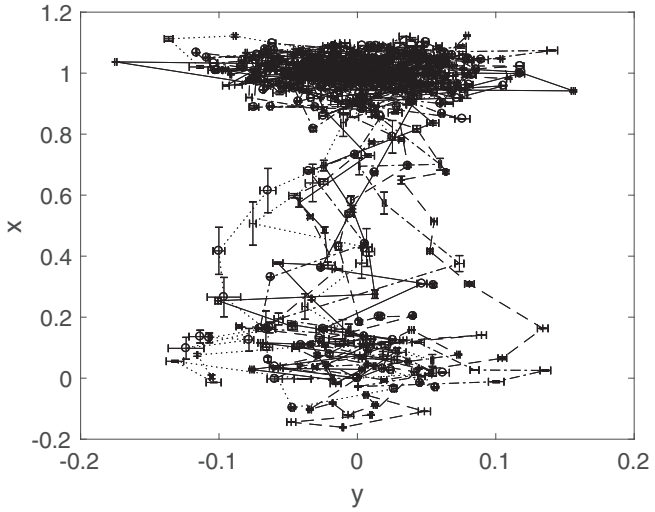

FIG. 7. Parametric plots of x versus y for ten independent stochastic bridge trajectories, with $\alpha = -10$, $\mu = 1$, and $\epsilon = 0.01$. Initial condition is the stable point at $\boldsymbol{x} = [1,0]$; total integration time is $\tau = 2$ and $t = 20$.

D. Nonpotential bridges

The last case we treat is a nonlinear FPE, also with a circulation term and no potential solution. There are many examples of this type, occurring in chemical reaction networks, thermal ratchets, and communications networks. These processes can lead to large rare fluctuations. The behavior of escape paths and transition events has been investigated previously. Due to the lack of a drift potential, the circulation term strongly modifies the bridge results.

A relevant example is a two-dimensional system proposed by Maier and Stein [55,56] and analyzed in greater detail by some later workers [26]. This describes two-dimensional over-damped stochastic motion in a force field that generally has no potential, where

$$\dot{x} = xp(x,y) + \sqrt{\epsilon}\zeta_1, \;\; \dot{y} = yq(x,y) + \sqrt{\epsilon}\zeta_2. \tag{53}$$

Here, $p(x,y) \equiv 1 - x^2 - \alpha y^2$ and $q(x) \equiv -\mu(1 + x^2)$. If $\alpha \neq \mu$, then this is not a gradient system. In the special case of $\alpha = \mu$, this has a drift potential. One may sample the corresponding bridge with an equivalent partial stochastic differential equation of

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \frac{1}{\epsilon}[\ddot{\boldsymbol{x}} + \boldsymbol{c}\dot{\boldsymbol{x}} + \boldsymbol{u}] + \sqrt{2}\boldsymbol{\xi}(t,\tau), \tag{54}$$

where the circulation term is an antisymmetric matrix:

$$\boldsymbol{\mathcal{C}} = \frac{\boldsymbol{c}}{\epsilon} = \frac{2}{\epsilon}(\alpha - \mu)xy\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{55}$$

and the virtual drift vector $\boldsymbol{U}$ is given by $\boldsymbol{U} = \boldsymbol{u}/\epsilon$, where

$$\boldsymbol{u} = \begin{bmatrix} x((3+\mu)\epsilon - p^2 + 2x^2 p + 2q\mu y^2) \\ y(\epsilon\alpha + 2p\alpha x^2 - q^2) \end{bmatrix}. \tag{56}$$

E. Nonpotential bridge numerical examples

The Maier and Stein system has a stable double-well potential, provided that $\alpha = \mu > 0$. If $\mu > 0$, then there are local attractors at $\boldsymbol{x} = (\pm 1,0)$ where the drift terms vanish, as well as an unstable equilibrium point at $\boldsymbol{x} = (0,0)$. In the cases where $\alpha \neq \mu$, these local attractors are not changed, but the off-axis flow with $y \neq 0$ *is* changed. This leads to large excursion events occurring during the escape trajectories, as illustrated by the solutions shown in Fig. 7. These are well-suited to investigation using stochastic bridge techniques. Here we focus on a stochastic bridge that connects a stable point to the saddle point, starting at $(-1,0)$ and ending at $(0,0)$. We suppose that $\mu = 1$ and analyze different values of α. This allows one to investigate the probability of different escape paths that connect the two stable points, conditioned on a predetermined transition time between the paths. An approximate analysis of the stochastic differential equation shows that, near the stable points, the effective equation in y has the form

$$\dot{y} \approx -2\mu y + \sqrt{\epsilon}\zeta_2.$$

As a result, for small enough noise coefficients ϵ, the y distribution will stabilize in a short time to give $\langle y^2 \rangle \approx \epsilon/4\mu$. However, paths with $y \neq 0$ result in x being attracted toward a new quasistable point, since on substituting $y^2 \approx \langle y^2 \rangle \approx \epsilon/4\mu$, the effective equation in x becomes

$$\dot{x} \approx -x(1 - x^2 - \epsilon\alpha/4\mu) + \sqrt{\epsilon}\zeta_2.$$

The x distribution will re-equilibrate to give an approximate Gaussian centered around $x = \sqrt{1 - \epsilon\alpha/4\mu}$. These analytic calculations predict that the early-time behavior of the x trajectories is to move rapidly to a new quasistable point that depends on α, prior to undergoing an escape from the local potential well.

In the potential case, with $\alpha = \mu = 1$, there is no circulation term, and bridge techniques that use gradient methods are able to calculate the stochastic bridge distributions. Typical trajectories are shown in Fig. 6.

With $\alpha < \mu$, there is no potential for the drift, although there is an effective metastable point as before. Paths with $y \neq 0$ result in large excursions in x, with x being amplified

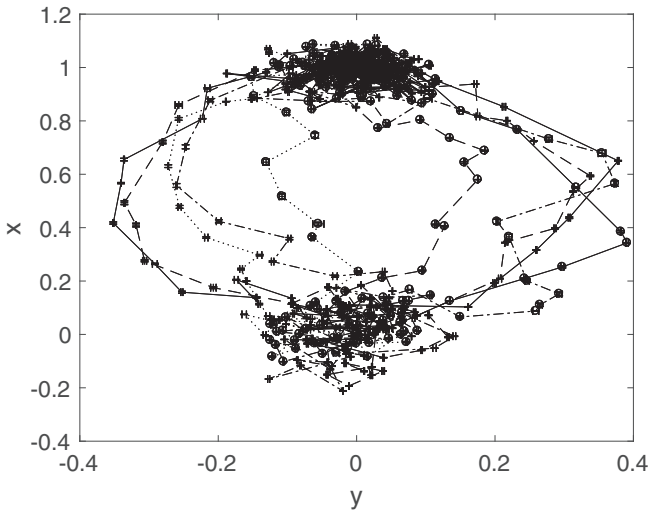

FIG. 8. Ten typical stochastic bridge trajectories, showing evolution of x, y during the constrained transition from a stable to a metastable point for the Maier-Stein equations with $\alpha = 6.67$, $\mu = 1$, and $\epsilon = 0.01$. Here, $\tau = 1$ and $t = 20$.

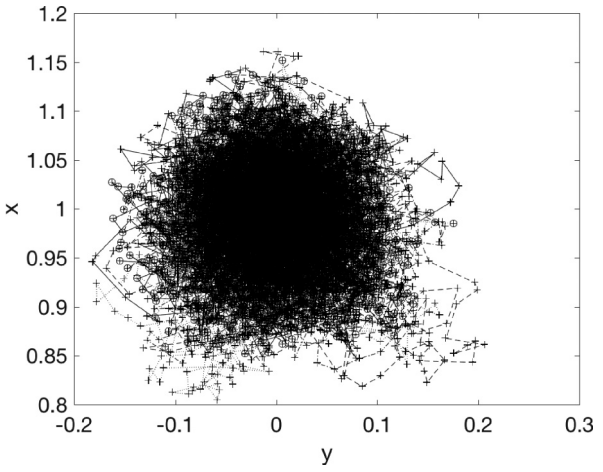

FIG. 9. Ten typical stochastic differential equation trajectories, showing evolution of x, y for $t = 20$, with the same parameters as in Fig. 8. Over this time scale, no escapes occur, and all trajectories are trapped near the stable region at $\boldsymbol{x} = [1,0]$.

away from the transition path, toward $\pm\infty$. There these are conditionally excluded in the bridge. In order for the system to travel stably between the stable points, y must remain smaller than its local equilibrium value. This gives transition paths that must thread the needle of small y values to make a transition. Typical trajectories with suppressed y values near $x = 0.8$, as they exit the stable region, are shown in Fig. 7. These have $\alpha = -10$.

The most interesting case is $\alpha \gg \mu$. Again, there is no potential solution. Paths with $y \neq 0$ will rapidly reduce x toward the origin, away from the stable point. As a result, there is a bifurcation during tunneling with characteristic curved transition paths that connect the stable points [26,55], and the opposite behavior to those with $\alpha \ll \mu$. A set of trajectories for the stochastic bridge with $\alpha = 6.67$ are shown in Fig. 8.

Finally, in Fig. 9, we show ten stochastic trajectories starting from the stable point, but *without* constraints and the same time duration. No escapes occur on this time scale. This shows the great utility of stochastic bridges for studying rare events. Simply taking the original equations and waiting for an escape to occur will generally take much longer than solving the bridge equations, with a total time duration that would become exponentially long as the noise is reduced.

In these numerical calculations, the same RK4 techniques with finite differences and convergence checks were used as in Fig. 1. All Maier-Stein bridge figures used 60 time steps and 1.5×10^4 virtual time steps. The stochastic differential equation solutions used 5000 time steps, with no significant discretization errors.

VI. SUMMARY

A general definition of stochastic bridges was given for arbitrary weighted stochastic equations. A path integral was obtained for the corresponding probability density. We show that there is a difference in stochastic equations in forward and backward time directions, owing to the different type of Bayesian conditioning involved. However, the final stochastic bridge results are completely identical regardless of the time-direction of the underlying equation, which directly relates to Schrödinger's original motivation for studying this problem. A numerical algorithm was obtained for the results given here, using a higher-dimensional stochastic partial differential equation. An example of the well-known exactly soluble Brownian bridge was evaluated numerically, agreeing with this previously known result.

Our results generalize previous ones and substantiate an earlier conjecture on stochastic bridges for equations with nonpotential flows. Several numerical examples are given of how large excursions away from the direct escape path may be analyzed using these methods.

ACKNOWLEDGMENTS

This work was performed in part at Aspen Center for Physics, which is supported by National Science Foundation Grant No. PHY-1607611.

[1] E. Schrödinger, Sitzber. Preuss. Akad. Wiss. Physik. Math. **K1**, 144 (1931).
[2] M. I. Dykman, P. V. E. McClintock, V. N. Smelyanski, N. D. Stein, and N. G. Stocks, Phys. Rev. Lett. **68**, 2718 (1992).
[3] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer-Verlag, Berlin, 1985), p. 442.
[4] J. L. Doob, Bull. Soc. Math. France **85**, 431 (1957).
[5] I. V. Girsanov, Theory Probab. Appl. **5**, 285 (1960).

[6] P. C. Martin, E. Siggia, and H. Rose, Phys. Rev. A **8**, 423 (1973).
[7] A. D. Fokker, Ann. Phys. **43**, 810 (1914).
[8] M. Planck, Sitzber Preuss. Akad. Wiss. Physik. Math. **K1**, 324 (1917).
[9] A. N. Kolmogorov, Usp. Mat. Nauk **5**, 5 (1938).
[10] P. D. Drummond, P. Deuar, and K. V. Kheruntsyan, Phys. Rev. Lett. **92**, 040405 (2004).
[11] T. Aimi and M. Imada, J. Phys. Soc. Jpn. **76**, 084709 (2007).
[12] P. Deuar and P. D. Drummond, Phys. Rev. A **66**, 033812 (2002).
[13] C. Calenge, Ecol. Model. **197**, 516 (2006).
[14] M. Hairer, A. M. Stuart, and J. Voss, Ann. Appl. Probab. **17**, 1657 (2007).
[15] H. Orland, J. Chem. Phys. **134**, 174114 (2011).
[16] G. N. Mil'shtein, Theor. Probab. Appl. **19**, 557 (1975).
[17] P. D. Drummond and I. K. Mortimer, J. Comput. Phys. **93**, 144 (1991).
[18] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1992).
[19] K. Burrage and T. Tian, New Zeal. J. Math. **29**, 115 (2000).
[20] M. R. Dowling, M. J. Davis, P. D. Drummond, and J. F. Corney, J. Comput. Phys. **220**, 549 (2007).
[21] N. Trivedi and D. M. Ceperley, Phys. Rev. B **41**, 4552 (1990).
[22] A. Beskos, G. Roberts, A. Stuart, and J. Voss, Stochast. Dynam. **08**, 319 (2008).
[23] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Vol. 293 (Springer Science & Business Media, Berlin, 2013).
[24] P. Glasserman, *Monte Carlo Methods in Financial Engineering*, reprint ed. (Springer, Berlin, 2010), p. 596.
[25] J. S. Horne, E. O. Garton, S. M. Krone, and J. S. Lewis, Ecology **88**, 2354 (2007).
[26] G. E. Crooks and D. Chandler, Phys. Rev. E **64**, 026109 (2001).
[27] R. J. Allen, D. Frenkel, and P. R. ten Wolde, J. Chem. Phys. **124**, 024102 (2006).
[28] R. Graham, Z. Phys. B **26**, 397 (1977).
[29] R. Graham, Phys. Rev. Lett. **38**, 51 (1977).
[30] R. Graham, Z. Phys. B **26**, 281 (1977).
[31] H. Dekker, Phys. Rev. A **24**, 3182 (1981).
[32] H. Dekker, Phys. Lett. A **65**, 388 (1978).
[33] U. Deininghaus and R. Graham, Z. Phys. B **34**, 211 (1979).
[34] C. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer Texts in Statistics (Springer, New York, 2013).
[35] D. T. Pegg, S. M. Barnett, and J. Jeffers, Phys. Rev. A **66**, 022106 (2002).
[36] P. H. Damgaard and H. Hüffel, Phys. Rep. **152**, 227 (1987).
[37] H. Risken, *The Fokker-Planck Equation*, 2nd ed. (Springer-Verlag, Berlin, 1996).
[38] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, reprint ed. (Folens Publishers, Dublin, 1992), p. 228.
[39] T. Bayes, R. Price, and J. Canton, *An Essay Towards Solving a Problem in the Doctrine of Chances* (C. Davis, London, 1763).
[40] B. D. Anderson, Stochast. Process. Appl. **12**, 313 (1982).
[41] M. G. Reznikoff and E. Vanden-Eijnden, Comptes Rendus Math. **340**, 305 (2005).
[42] N. Wiener, Acta Math. **55**, 117 (1930).
[43] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
[44] S. Machlup and L. Onsager, Phys. Rev. **91**, 1512 (1953).
[45] R. L. Stratonovich, SIAM J. Control **4**, 362 (1966).
[46] B. S. DeWitt, Rev. Mod. Phys. **29**, 377 (1957).
[47] W. Horsthemke and A. Bach, Z. Phys. B **22**, 189 (1975).
[48] R. P. Feynman, A. R. Hibbs, and D. F. Styer, *Quantum Mechanics and Path Integrals* (Courier Corporation, North Chelmsford, MA, 2010).
[49] R. P. Feynman, Phys. Rev. **84**, 108 (1951).
[50] K. L. Hunt and J. Ross, J. Chem. Phys. **75**, 976 (1981).
[51] M. J. Werner and P. D. Drummond, J. Comput. Phys. **132**, 312 (1997).
[52] R. Mansuy and M. Yor, *Aspects of Brownian Motion* (Springer Science & Business Media, Berlin, 2008).
[53] S. Kiesewetter, R. Polkinghorne, B. Opanchuk, and P. D. Drummond, SoftwareX **5**, 12 (2016).
[54] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, New York, NY, 2007).
[55] R. S. Maier and D. L. Stein, Phys. Rev. Lett. **71**, 1783 (1993).
[56] R. S. Maier and D. L. Stein, Phys. Rev. Lett. **69**, 3691 (1992).