Conditional Path Sampling of SDEs and the Langevin MCMC Method

Andrew M. Stuart* Jochen Voss† Petter Wiberg‡

Abstract

We introduce a stochastic PDE based approach to sampling paths of SDEs, conditional on observations. The SPDEs are derived by generalising the Langevin MCMC method to infinite dimensions. Various applications are described, including sampling paths subject to two end-point conditions (bridges) and nonlinear filter/smoothers.

*Mathematics Institute, University of Warwick, Coventry CV4 7AL, England.

†Mathematics Institute, University of Warwick, Coventry CV4 7AL, England; previously: University of Kaiserslautern, 67653 Kaiserslautern, Germany.

‡Mathematics Institute, University of Warwick, Coventry CV4 7AL, England; since May 2004: Goldman-Sachs, London.

1 Introduction

The Markov chain Monte Carlo (MCMC) method is an important tool for solving computational problems in Bayesian statistics [BS94, Liu01, RC99]. The main idea is to sample from the target density by constructing a Markov chain that has the target density as its stationary distribution. Assuming that the Markov chain is ergodic, statistical quantities can be computed by simulating the Markov chain and computing statistics aggregated over time.

An important basic concept in sampling is Langevin dynamics [RC99]. Suppose that $p \propto \exp(-V)$ is a target density on $\mathbb{R}^n$. The stochastic differential equation (SDE)

$$\frac{dx}{dt} = -\nabla V(x) + \sqrt{2}\,\frac{dW}{dt} \tag{1.1}$$

has p as its invariant density — it satisfies the stationary Fokker-Planck equation

$$0 = \nabla \cdot (p\nabla V + \nabla p)\,. \tag{1.2}$$

So, assuming that (1.1) is ergodic, $x(t)$ produces samples from the target density p as $t \to \infty$. The empirical measure generated by (1.1) converges to the invariant measure whose density p solves (1.2) and the ergodic theorem gives that

$$\int f(\zeta)p(\zeta)\,d\zeta = \lim_{T\to\infty} \frac{1}{T}\int_0^T f\big(x(t)\big)\,dt. \tag{1.3}$$

Thus averages over p can be computed by taking time averages over a trajectory of (1.1).

The Langevin method is important for three primary reasons. Firstly, in its own right as the basis for the construction of a variety of MCMC algorithms. For example (1.1) can be discretised in time by the Euler-Maruyama method [KP99] (a Metropolis step can be added to remove bias if the time step is such that the bias is significant). Secondly it is also important conceptually as it arises as the diffusion limit of a number of other MCMC algorithms and hence unifies a variety of sampling methods [RGG97]. Thirdly, there are also a number of variants of the method, which are important, such as: generalisations to second order dynamics in time (which is actually the equation referred to by the name Langevin in the molecular modelling context), whilst preserving the invariant density as p when marginalising $\frac{dx}{dt}$ out and retaining x; pre-conditioning by a positive-definite symmetric matrix A to obtain

$$\frac{dx}{dt} = -A\nabla V(x)\,dt + \sqrt{2}A^{1/2}\,\frac{dW}{dt}, \tag{1.4}$$

which also preserves the invariant density as p.

For these three reasons it is natural to study Langevin methods when attempting to sample in *infinite* dimensions. In this paper we study a number of infinite dimensional sampling problems, related to conditional sampling of paths of SDEs. We construct Langevin methods for their resolution. This leads naturally to some interesting stochastic PDEs (SPDEs).

1.1 Outline of the Paper

In section 2 we study the sampling of paths of SDEs, subject to end-point conditions on both ends of a finite interval (bridges). We state an SPDE which performs this sampling, and corresponds to an infinite dimensional Langevin technique.

In section 3 we study nonlinear filtering/smoothing by the Langevin technique, developing a methodology conceptually quite different from the particle filter methods currently used for the filtering problem, which are based on the Zakai equation. Again we state an SPDE which solves the sampling problem, and corresponds to an infinite dimensional Langevin technique. For off-line data, filter/smoothing may be more natural than pure filtering, and hence our method is of interest.

For both the sampling problems in sections 2 and 3, the path we wish to sample is a random variable depending on both time and randomness in form of the the driving noise. The conditioning gives information about that noise and creates a highly non-trivial sampling problem. In some cases (see the discussion below) the SPDEs we write down can be proven to sample from the required distribution. However, there are problems for which proving this appears non-trivial and open problems remain. For this reason we present a robust, but currently non-rigorous, method for the derivation of these SPDEs, via discretisation. (The discretisation used to *derive* the SPDEs is not the discretisation technique that we use to *solve* the SPDEs in section 5.) Section 4 contains the heuristics which underly the derivation of the SPDEs in sections 2 and 3 by this discretisation approach. These heuristics rely on discretising the infinite dimensional problem, applying the finite dimensional Langevin algorithm (1.1), and passing to a continuum limit; numerical results are presented which resolve certain ambiguities in this limiting procedure.

In section 5 we describe some numerical experiments which illustrate the sampling methods we have introduced. All the experiments are performed by applying standard finite difference techniques to discretise the SPDEs which have been constructed to perform infinite dimensional sampling.

1.2 Conclusions and Future Directions

To the best of our knowledge the methodology introduced in this paper is new, in the generality considered here. The SPDEs are derived by the method of discretisation of the sampling problem and passage to the limit. This is a flexible tool applicable to many problems. We end this introduction by highlighting relationships between this work and existing literature, and by describing the many future directions in which this work can be taken.

(i) *Rigorous justification of sampling properties for bridge processes.* For the case of bridge sampling, discussed in section 2, the SPDE that we write down can be proven to perform the desired sampling. This is done by considering the bridge process as a Gibbs measure for stationary paths of an SDE defined on the whole real line, and is described in [BL03]. In that paper bridge processes are considered for SDEs whose drift is of gradient form. This idea is used to study the connection between invariant measures of SPDEs and bridge processes in [RVE04] and its use in the context of sampling is studied further in [HSVW]). Many open questions remain, particularly bridge processes in SDEs where the drift is not in gradient form.

(ii) *Rigorous justification of sampling properties for nonlinear filtering/smoothing.* For the Kalman case, found by taking linear drift vector fields in section 3, the SPDE which we write down can be proven to sample from the Kalman filter/smoother. This is demonstrated in [HSVW]. Many open problems remain, in particular for nonlinear drift fields.

(iii) *Passage to the limit in finite difference equations.* In section 4 we derive the SPDEs heuristically through passage to the limit in a system of (finite difference like) SDEs. This is a useful tool for the general problem of infinite dimensional sampling. Making this rigorous would be another useful route by which to justify the SPDE-based sampling method introduced here.

(iv) *Rates of convergence in the SPDEs.* The ergodicity and rate of convergence to stationarity in the SPDEs we have derived derived is central to understanding their efficiency. A rigorous analysis, using recently developed theory of ergodicity for SPDEs of reaction-diffusion type, would be of interest. For some problems the rate of convergence may be exponentially slow in the non-dimensional parameters of the problem (e.g. for phase transition problems such as (5.1), (5.2) which are notoriously difficult to sample). For such problems the basic SPDE methods proposed here will be inefficient (see also the discussion of transition path sampling below).

(v) *Generalisations of the basic Langevin idea.* As described at the start of this section, the Langevin method is the starting point for a variety of more sophisticated methods which are tailored to improve efficiency. For example, one can make nontrivial choices for A in (1.4) in the infinite dimensional context. This is initiated in [RSV], where we choose A to be a Green's operator. Further analysis of these issues, in the context of metropolising the time-stepping algorithms used for the SPDEs, will be of central importance in making the sampling methods presented here into efficient algorithms.

(vi) *Applications.* It is of interest to investigate use of the methods introduced here to solve conditional path sampling problems arising in applications such as signal processing, econometrics and finance.

(vii) *Comparative evaluations.* Having optimised the methods introduced here, and their variants, it will be of interest to compare them with other recently introduced methods for bridge sampling, such as [CPS, RS01, BR], and methods for nonlinear filtering, such as those outlined in [DdFG01]).

(viii) *Transition path sampling.* Finding transition pathways in chemically reacting systems is an applied problem of considerable importance ([DBG02, BCDG02, ERVE02, ERVE]), and one which can be cast similarly to the bridge sampling problems of section 2. It is likely that the SPDE based method proposed here would be ineffectual for such problems because of the large (relative to the size of the noise) energy barriers in path space which would need to be traversed to perform effective sampling; large deviation theory could be used to quantify such observations. However, it is conceivable that the methodology proposed here could , in combination with other methods, play a small part in tackling the hard infinite dimensional sampling problem of transition path sampling.

2 Bridge Paths of SDEs

Suppose that we want to sample paths from the stochastic differential equation

$$\frac{dX}{du} = f(X) + \sigma \frac{dW}{du} \tag{2.1a}$$

subject to the end-point conditions

$$X(0) = x^-, \quad X(1) = x^+. \tag{2.1b}$$

These are bridge paths. Here W is a standard Brownian motion. Such problems arise naturally in parameter estimation problems in financial time series analysis (for example in [RS01]) and in econometrics (see for example [CPS]).

Langevin sampling on the path space provides a sampler for such constrained problems. For (2.1) the infinite dimensional Langevin equation is the SPDE

$$\frac{\partial x}{\partial t} = \frac{1}{\sigma^2}\left(\frac{\partial^2 x}{\partial u^2} - f(x)f'(x) - \frac{\sigma^2}{2}f''(x)\right) + \sqrt{2}\frac{\partial w}{\partial t} \qquad \forall (u,t) \in [0,1]\times(0,\infty), \tag{2.2a}$$

$$x(0,t) = x^-, \quad x(1,t) = x^+ \qquad \forall t \in (0,\infty), \tag{2.2b}$$

$$x(u,0) = x_0(u) \qquad \forall u \in [0,1]. \tag{2.2c}$$

Here $\frac{\partial w}{\partial t}$ is a white noise in (u,t).

The derivation of the SPDE (2.2) to sample from (2.1) is given in section 4. The SPDE as written is purely formal. To make sense of solutions we should apply the variation of constants formula based on the heat operator $\left(\frac{\partial}{\partial t} - \frac{1}{\sigma^2}\frac{\partial^2}{\partial u^2}\right)$, as described in [DPZ92]. Although stated for scalar X the methodology generalises to vector valued X.

To avoid confusion we adopt the following convention: we use upper case X for solutions of the SDEs from which we sample paths and denote the corresponding time by u; we denote the (finite or infinite dimensional) Langevin process by x with algorithmic time t.

If the end-point conditions (2.1b) are replaced by leaving the path free at $u = 0, 1$ then we obtain the boundary conditions

$$\frac{\partial x}{\partial u}(0,t) = f\big(x(0,t)\big), \quad \frac{\partial x}{\partial u}(1,t) = f\big(x(1,t)\big) \qquad \forall t \in (0,\infty) \tag{2.3}$$

in place of (2.2b). Furthermore, the boundary conditions from (2.2b) and (2.3) can be combined to specify X at $u = 0$ and leave it free at $u = 1$. This leads to the boundary conditions

$$x(0,t) = x^-, \quad \frac{\partial x}{\partial u}(1,t) = f\big(x(1,t)\big) \qquad \forall t \in (0,\infty). \tag{2.4}$$

Note, however, that in the case of a free right end-point condition with X specified at $u = 0$, there are much simpler methods available for sampling [KP99]. Of course we do not recommend use of the infinite dimensional Langevin method for such problems.

3 Nonlinear Filtering and Smoothing

Consider the system of stochastic differential equations

$$\begin{aligned} &\text{signal:} && \frac{dX}{du} = f(X) + \sigma\,\frac{dW}{du}, && X(0) \sim \mathcal{N}(a,\delta^2) \\ &\text{observation:} && \frac{dY}{du} = g(X) + \gamma\,\frac{dV}{du}, && Y(0) = 0, \end{aligned} \tag{3.1}$$

where W and V are standard Brownian motions. We think about this pair of equations as describing a noisily observed signal. The first equation gives the stochastic evolution of the signal. The second equation describes a noisy observation of this state.

Our aim is to sample paths of the signal X conditioned on a given observation path Y; more precisely, to find $X(v)$ given $(Y(u))_{u\in[0,1]}$, for all $v \in [0,1]$. The Langevin strategy can be applied to this situation, too. The resulting SPDE is

$$\begin{aligned} \frac{\partial x}{\partial t} = \frac{1}{\sigma^2}\left(\frac{\partial^2 x}{\partial u^2} - f(x)f'(x) - \frac{\sigma^2}{2}f''(x)\right) \\ + \frac{g'(x)}{\gamma^2}\left(\frac{dY}{du} - g(x)\right) + \sqrt{2}\frac{\partial w}{\partial t} \qquad \forall (u,t) \in [0,1]\times(0,\infty), \end{aligned} \tag{3.2a}$$

$$\frac{\partial x}{\partial u}(0,t) = f\big(x(0,t)\big) + \frac{\sigma^2}{\delta^2}\big(x(0,t) - a\big), \quad \frac{\partial x}{\partial u}(1,t) = f\big(x(1,t)\big) \qquad \forall t \in (0,\infty), \tag{3.2b}$$

$$x(u,0) = x_0(u) \qquad \forall u \in [0,1]. \tag{3.2c}$$

As in the previous section, $\frac{\partial w}{\partial t}$ is a white noise in (u,t).

Notice that if f,g are linear then we are in a setup which enables us to obtain the Kalman filter/smoother, namely

$$\hat{X}(v) = \mathbb{E}\big(X(v) \bigm| (Y(u))_{u\in[0,1]}\big) \qquad \forall v \in [0,1].$$

To be precise, $\hat{X}$ can be found as

$$\hat{X}(v) = \lim_{T\to\infty} \frac{1}{T}\int_0^t x(v,t)\,dt.$$

Moreover, the empirical measure generated by $(x(\cdot\,,t))_{t\geq 0}$ enables us to compute other statistics associated with the Kalman filter/smoother such as covariances. However the primary interest in (3.2) stems from the case where f,g are nonlinear and we obtain a simple method for finding the nonlinear filter/smoother. For off-line situations this may be competitive with particle filters (see [DdFG01]).

The derivation of the SPDE (3.2) to sample from X given Y is given in the next section. As in the previous section the SPDE (3.2) is purely formal and should be given a rigorous interpretation through the variation of constants formula as in [DPZ92]. Again, generalisations to vector valued X and Y are possible.

4 Derivation and Numerical Justification

The purpose of this section is to derive the SPDEs written down in the previous two sections. As mentioned in the introduction, there are a number of special cases of the SPDEs introduced here which can be justified cleanly and directly on a case by case basis. However, discretisation and passage to the limit, whilst somewhat unwieldy, appears to be a useful and flexible tool which applies to a wide range of problems in a unified fashion. Hence we describe it.

It is important to appreciate that what we describe in this section is a method for *deriving* the SPDEs stated in the previous two sections, and *is not our recommended method for solving these SPDEs*. We say a few words about how the SPDEs are solved numerically in the next section. The reader interested only in seeing statements of the SPDEs used for conditional path sampling, and then numerical examples illustrating their use, can easily ignore this section without losing any understanding.

Our approach to derive the SPDEs is to discretise the sampling problem, apply the finite dimensional sampling method (1.1), and pass to a continuum limit. Euler's method applied to (2.1a) gives

$$X_n = X_{n-1} + f(X_{n-1})\,\Delta u + \sigma\,\Delta W_n \qquad \forall n = 1, \ldots, N, \tag{4.1}$$

where $\Delta u = 1/N$ and the ΔW_n are i.i.d. $\mathcal{N}(0, \Delta u)$ distributed random variables. Using this i.i.d. property, the density of the distribution of $X = (X_1, \ldots, X_N)$ on $\mathbb{R}^N$ is readily shown to be

$$\varphi(x) = C \exp\bigl(-\Delta u\,\mathcal{V}(x)\bigr) \tag{4.2}$$

with the potential

$$\mathcal{V}(x) = \sum_{n=1}^{N} \frac{1}{2\sigma^2}\Bigl(\frac{x_n - x_{n-1}}{\Delta u} - f(x_{n-1})\Bigr)^2. \tag{4.3}$$

Applying the Langevin method (1.4) with $A = I/\Delta u$ gives

$$\frac{dx}{dt} = -\nabla\mathcal{V}(x) + \sqrt{\frac{2}{\Delta u}}\,\frac{dW}{dt}. \tag{4.4}$$

We use the notational conventions described at the end of section 2.

4.1 Bridge Paths of SDEs

Suppose that the left and right boundary points are fixed in (4.1). Then $X_0 = x^-$ and $X_N = x^+$ are known and for $X_1, \ldots, X_{N-1}$ we use the conditional density which is obtained by fixing $x_N = x^+$ and choosing an appropriate normalisation constant in (4.2). The dynamics from (4.4) are determined by

$$\frac{\partial\mathcal{V}}{\partial x_n} = -\frac{1}{\sigma^2}\left(\frac{x_{n+1} - 2x_n + x_{n-1}}{\Delta u^2} - f(x_n)f'(x_n) - \frac{f(x_n) - f(x_{n-1})}{\Delta u} + f'(x_n)\frac{x_{n+1} - x_n}{\Delta u}\right)$$

for $1 \leq n \leq N-1$.

We consider the limit $N \to \infty$. In taking this limit we assume that $x_n = x(n\Delta u)$ samples a smooth function $x(\,\cdot\,)$, except at the end of the argument when we let $x(\,\cdot\,)$ have non-zero quadratic variation. For the first term on the right hand side we get

$$\frac{x_{n+1} - 2x_n + x_{n-1}}{\Delta u^2} \approx \frac{\partial^2 x}{\partial u^2}.$$

The second term needs no interpretation. For the final two terms we use Taylor expansion of f around x_{n-1} to give

$$\begin{aligned}
&-\frac{f(x_n) - f(x_{n-1})}{\Delta u} + f'(x_n)\frac{x_{n+1} - x_n}{\Delta u} \\
&\quad= \frac{f'(x_n) - f'(x_{n-1})}{\Delta u}(x_n - x_{n-1}) + f'(x_n)\frac{x_{n+1} - 2x_n + x_{n-1}}{\Delta u} \\
&\qquad\quad - \frac{1}{2}\bigl(f''(x_{n-1}) + o(1)\bigr)\frac{(x_n - x_{n-1})^2}{\Delta u}.
\end{aligned}$$

We neglect all but the last term, into which we substitute the quadratic variation for $x_n = X(n\Delta u)$ with $X(\,\cdot\,)$ solving (2.1), the equation from which we wish to sample paths. The quadratic variation is unaffected by the choice of boundary conditions, and so we get $\mathbb{E}(x_n - x_{n-1})^2 \approx \sigma^2 \Delta u$. Thus we take

$$-\frac{f(x_n) - f(x_{n-1})}{\Delta u} + f'(x_n)\frac{x_{n+1} - x_n}{\Delta u} \approx -\frac{\sigma^2}{2} f''(x).$$

Noting that $\frac{dW}{dt}$ in (4.4) is standard Brownian motion in $\mathbb{R}^{N-1}$, we see that the term $\sqrt{\frac{2}{\Delta u}}\,\frac{dW}{dt}$ is formally an approximation to space-time white noise. By collecting the terms in (4.4) and passing to the limit we arrive at the SPDE (2.2).

4.2 Free End-Points

If the right end-point is free and we have the boundary conditions (2.4), then a Langevin equation must be derived for x_N. The gradient of the potential at that point is

$$\frac{\partial \mathcal{V}}{\partial x_N} = -\frac{1}{\sigma^2 \Delta u}\left(f(x_{N-1}) - \frac{x_N - x_{N-1}}{\Delta u} \right),$$

which leads to the Langevin SDE

$$\frac{dx_N}{dt} = -\frac{1}{\sigma^2 \Delta u}\Big(f(x_{N-1}) - \frac{x_N - x_{N-1}}{\Delta u} \Big) + \sqrt{\frac{2}{\Delta u}}\,\frac{dW_N}{dt}.$$

Multiplying this equation by Δu and letting $\Delta u \to 0$ formally leads to the boundary condition (2.4). A similar argument gives the boundary condition (2.3) if the left end-point $u = 0$ is allowed to vary.

4.3 A Note of Caution Concerning Derivation of the SPDE

When taking the limit $N \to \infty$ for the drift in the Langevin equation (4.4) we used Taylor approximation around the point x_{n-1} to obtain the f'' term in (2.2a). Surprisingly we get a different result if we use Taylor expansion around x_n instead of x_{n-1}. In this case we get

$$-\frac{f(x_n) - f(x_{n-1})}{\Delta u} + f'(x_n)\frac{x_{n+1} - x_n}{\Delta u} \approx +\frac{\sigma^2}{2} f''(x),$$

which leads to the opposite sign for the f'' term in (2.2a)! The next section presents arguments which justify the derivation from section 4.1.

4.4 Resolving the Ambiguity

For the bridge path problem (2.1) the choice of sign in the f'' term in the SPDE can be justified rigorously by appealing to the Girsanov Formula, and writing the density with respect to Brownian bridge (see [BL03], [HSVW] and [RVE04]). However we pursue a self-contained numerical justification of the choice of sign.

Consider the SPDE (2.2) with (2.2a) replaced by

$$\frac{\partial x}{\partial t} = \frac{1}{\sigma^2}\left(\frac{\partial^2 x}{\partial u^2} - f(x) f'(x) - \xi \frac{\sigma^2}{2} f''(x) \right) + \sqrt{2}\frac{\partial w}{\partial t} \qquad \forall (u,t) \in [0,1] \times (0,\infty), \tag{4.5}$$

where $\xi \in \mathbb{R}$ is the parameter we are interested in and (2.2b) is replaced by (2.4). With the correct choice for ξ, this should sample from the SDE (2.1a) with initial condition $x(0) = x^-$ and no condition at $u = 1$. Thus for the correct value of ξ, a path of (2.1a) with $X(0) = x^-$ should be stationary data for (4.5).

To make the f'' term visible we have to choose a nonlinear drift f. For our experiment we consider the drift function $f(x) = x(1 - x^2)$, the diffusion constant $\sigma = 1$, the initial point $x^- = -1$, and we implement the SPDE (4.5) with $\xi = -1, 0, +1$ in turn.

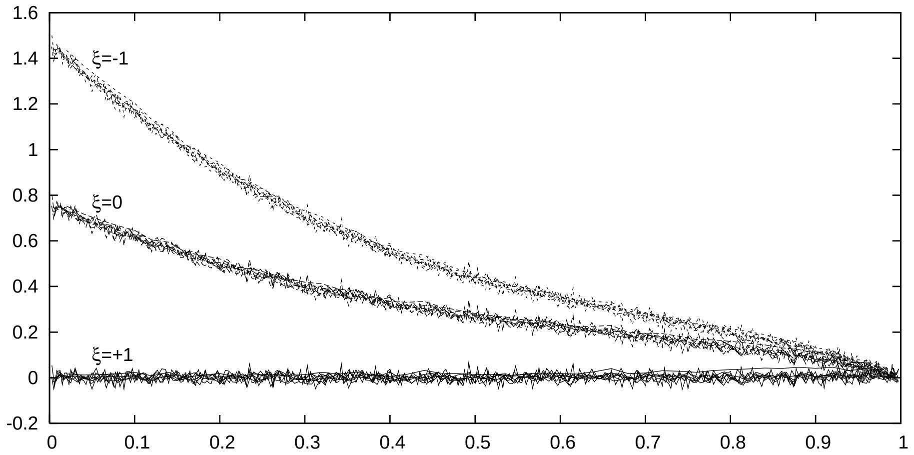

Figure 1: *Demonstration that ξ should take the value $+1$ in (2.3). The graph is of the empirical mean of $Z_n/\Delta u$, from (4.6) for $N = (50\ell)^{-1}$, $\ell = 1, \ldots, 8$. The same noise is used to generate each curve.*

The result is displayed in figure 1. For the picture we simulated (4.5), (2.4) using a Crank-Nicolson discretisation of the heat operator, with explicit Euler treatment of the remaining terms. The discretisation parameters were $\Delta u = 1/N$ and $\Delta t = 10\Delta u^2$ for different values of N. For the initial condition we choose a path generated by the implicit Euler method

$$X_{n+1} = X_n + \frac{f(X_n) + f(X_{n+1})}{2} \Delta u + \Delta W_n, \quad X_0 = -1.$$

This initial distribution should be close to the stationary distribution for the SPDE (4.5), (2.4). Then we simulate the SPDE until time $t = 0.1$ and consider the values

$$Z_n = x(t, u_{n+1}) - x(t, u_n) - \frac{f\big(x(t, u_n)\big) + f\big(x(t, u_{n+1})\big)}{2} \Delta u \tag{4.6}$$

for $n = 0, ..., N-1$ where $u_n = n\Delta u$. If we are still in equilibrium these values should be i.i.d. $\mathcal{N}(0, \Delta u)$ random variables.

Figure 1 shows the empirical expectation of $Z_n/\Delta u$, obtained by taking the mean of 10^6 realisations. One can see that the grid parameter N has almost no influence on the mean and that only the drift from the SPDE (2.2a), i.e. the case $\xi = 1$, leads to the correct mean.

4.5 Extension to the Smoothing Problem

Now consider the system (3.1). The discretisation with step size $\Delta u = 1/N$ is

$$\begin{aligned} X_n &= X_{n-1} + f(X_{n-1})\Delta u + \sigma \Delta W_n & \qquad X_0 &\sim \mathcal{N}(a, \delta^2) \\ Y_n &= Y_{n-1} + g(X_{n-1})\Delta u + \gamma \Delta B_n & Y_0 &= 0 \end{aligned}$$

for $n = 1, \ldots, N$. The potential $\mathcal{V}$ from equation (4.2) should be replaced by $\mathcal{V}_0(x) + \mathcal{V}(x) + \mathcal{U}(x, y)$, where $\mathcal{V}$ is defined by (4.3),

$$\mathcal{V}_0(x) = \frac{1}{\Delta u} \frac{(x_0 - a)^2}{2\delta^2}$$

represents the initial condition for X, and

$$\mathcal{U}(x, y) = \sum_{n=1}^{N} \frac{1}{2\gamma^2} \Big(\frac{y_n - y_{n-1}}{\Delta u} - g(x_{n-1}) \Big)^2.$$

The Langevin equation (4.4) is now replaced by

$$\frac{dX}{dt} = -(\nabla_x \mathcal{V}_0 + \nabla_x \mathcal{V} + \nabla_x \mathcal{U})(X) + \sqrt{\frac{2}{\Delta u}}\frac{dW}{dt}, \tag{4.7}$$

where $\partial \mathcal{V}_0 / \partial x_n = 0$ and

$$-\frac{\partial \mathcal{U}}{\partial x_n} = \frac{1}{\gamma^2}\left(\frac{y_{n+1}-y_n}{\Delta u} - g(x_n)\right) g'(x_n)$$

for $n = 1, \ldots, N-1$. As $N \to \infty$ this gives the extra term appearing in equation (3.2a).

The derivation for the boundary condition at $u = 1$ is the same as in section 4.2, because $\partial \mathcal{V}_0/\partial x_N = \partial \mathcal{U}/\partial x_N = 0$. For $u = 0$ we find

$$\begin{aligned} &-\frac{\partial(\mathcal{V}_0 + \mathcal{V} + \mathcal{U})}{\partial x_0} \\ &\quad = -\frac{1}{\Delta u}\frac{x_0 - a}{\delta^2} + \frac{1}{\sigma^2}\Big(\frac{x_1 - x_0}{\Delta u} - f(x_0)\Big)\Big(\frac{1}{\Delta u} + f'(x_0)\Big) + \frac{1}{\gamma^2}\Big(\frac{y_1 - y_0}{\Delta u} - g(x_0)\Big) g'(x_0) \end{aligned}$$

and multiplying this by Δu and letting $\Delta u \to 0$ in (4.7) formally leads to the boundary condition from (3.2b).

5 Numerical Examples

In this section we present three numerical examples, illustrating the methods from sections 2 and 3. All the figures in this section are produced after 10^3 time units; some cursory diagnostics indicate that we have then reached statistical stationarity in all three examples. The numerical method employed is a semi-implicit discretisation based on a Crank-Nicolson approximation of the heat operator, together with explicit treatment of the remaining terms. For the filtering problem the time-derivative of the observation y is approximated in a one-sided fashion. The discretisation parameters used in all the experiments reported here are $\Delta u = 10^{-2}$ and $\Delta t = \Delta u^2$.

5.1 Bridge Paths

Consider the SDE

$$\frac{dX}{du} = f(X) + \frac{dW}{du} \tag{5.1}$$

where the drift f is the (negative) gradient of a double well potential with two stable equilibrium points at -1 and $+1$:

$$f(x) = -\Big(\frac{(x-1)^2(x+1)^2}{1+x^2}\Big)' = x\Big(\frac{8}{(1+x^2)^2} - 2\Big). \tag{5.2}$$

Because the phase transition behaviour of the system (5.1) is not visible on $O(1)$ time scales, we consider the time interval $[0, 100]$. The technique from section 2 is readily extended to arbitrary bounded time intervals and so we can sample from solutions of (5.1) on the interval $[0, 100]$ and subject to the boundary conditions

$$X(0) = -1, \quad X(100) = +1.$$

Figure 2 shows the result of a simulation of (2.2).

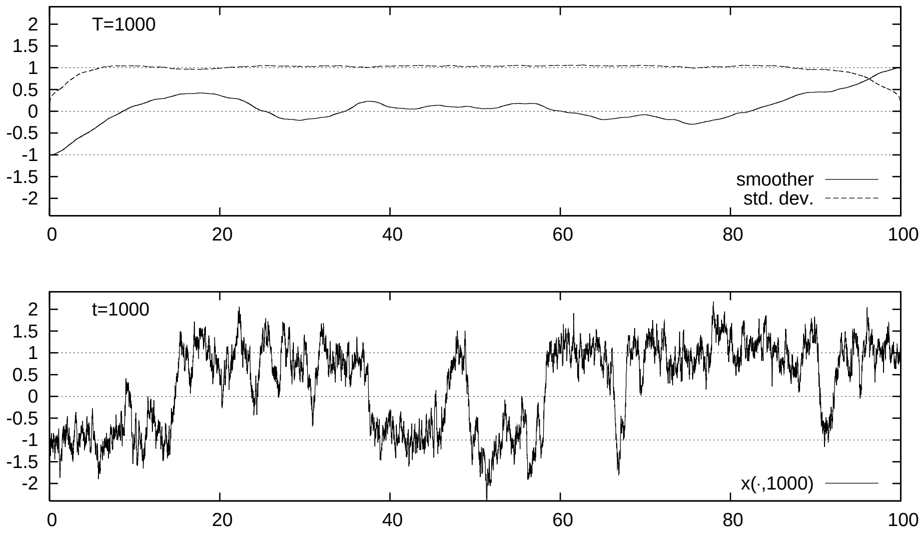

Figure 2: *The bridge double-well process from section 5.1. The first diagram shows the empirical expectation and standard deviation for solutions of* (5.1)*, obtained by evaluating the ergodic average from* (1.3) *at* $T = 1000$*. The second diagram shows a sample from the SDE, namely* $x(\cdot, 1000)$ *where* x *is a solution of the corresponding SPDE* (2.2)*.*

5.2 Gaussian Smoothing

To illustrate the smoothing method from section 3 we consider the following linear smoothing problem.

$$\begin{aligned} \text{signal:} \quad & \frac{dX}{du} = -X + \frac{dW}{du}, \qquad X(0) = 0 \\ \text{observation:} \quad & \frac{dY}{du} = \quad X + \frac{1}{10}\frac{dV}{du}, \qquad Y(0) = 0, \end{aligned} \tag{5.3}$$

where W and V are one-dimensional Brownian motions. This corresponds to the case $\delta^2 = 0$ in section 3. Because these SDEs are linear, SPDE (3.2) simplifies to

$$\frac{\partial x}{\partial t} = \Big(\frac{\partial^2 x}{\partial u^2} - x\Big) + 100\Big(\frac{dY}{du} - x\Big) + \sqrt{2}\frac{\partial w}{\partial t} \qquad \forall (u,t) \in [0,1] \times (0,\infty), \tag{5.4a}$$

$$x(0,t) = 0, \quad \frac{\partial x}{\partial u}(1,t) = -x(1,t) \qquad \forall t \in (0,\infty), \tag{5.4b}$$

$$x(u,0) = x_0(u) \qquad \forall u \in [0,1]. \tag{5.4c}$$

The stationary distribution of this SPDE is the distribution of X given Y. By taking ergodic limits as in (1.3) we can, for a given observation Y, numerically estimate the expectation and variance of X given Y. The result of a simulation is shown in figure 3.

5.3 Nonlinear Smoothing

While the linear case from the preceding section could easily be treated by standard techniques as in [Øks98], the following nonlinear example requires more sophisticated methods, such as the particle filtering techniques described in [DdFG01]. Here we illustrate our method.

We consider the nonlinear system

$$\begin{aligned} \text{signal:} \quad & \frac{dX}{du} = f(X) + \frac{dW}{du}, \qquad X(0) = -1 \\ \text{observation:} \quad & \frac{dY}{du} = \quad X \ + \frac{dV}{du}, \qquad Y(0) = 0, \end{aligned} \tag{5.5}$$

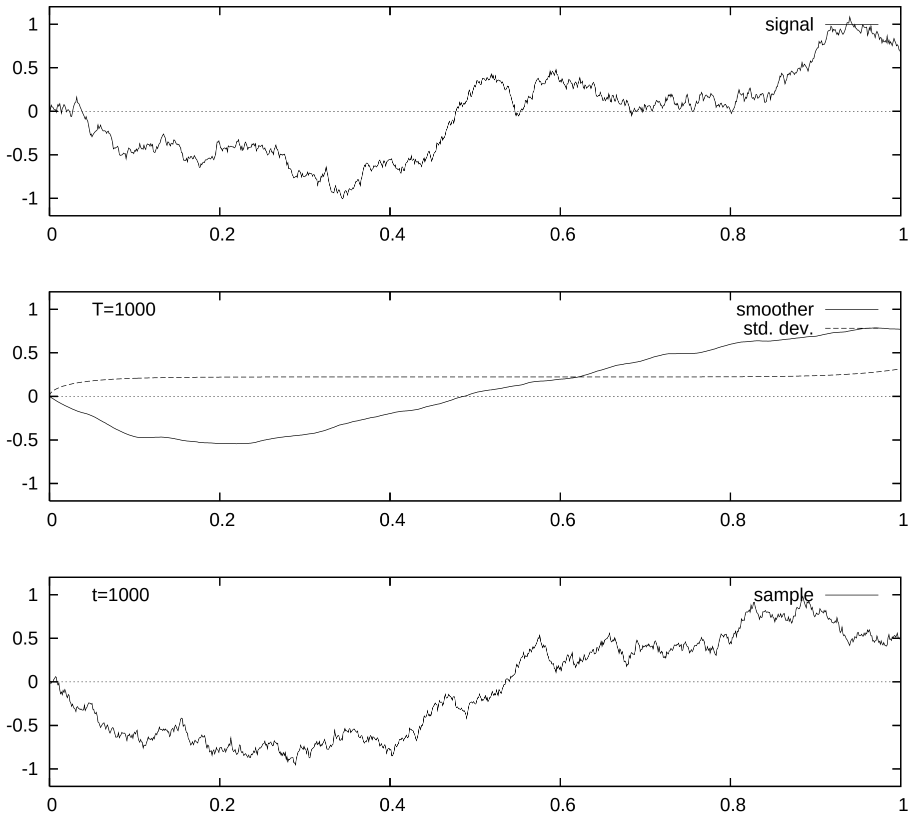

Figure 3: *The linear filtering problem from section 5.2. The first diagram shows the signal X. The corresponding observation Y is not displayed. The middle diagram shows the conditional expectation and standard deviation of X given Y, obtained by evaluating the ergodic average from* (1.3) *at* $T = 1000$. *The third diagram gives a sample from the conditional distribution of X given Y, namely $x(\,\cdot\,, 1000)$ where x solves the SPDE* (3.2).

where f is the double-well drift from equation (5.2). Figure 4 shows a simulation of the nonlinear Kalman sampler for this case, using the SPDE (3.1).

Acknowledgements. We are grateful to Eric Vanden Eijnden for helpful suggestions. This project was partially funded by the EPSRC. Jochen Voss was additionally funded by the European Union with a Marie Curie scholarship. The computing facilities were provided by the Centre for Scientific Computing of the University of Warwick.

References

[BCDG02] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Ann. Rev. Phys. Chems.*, 53:291–318, 2002.

[BL03] V. Betz and J. Lőrinczi. Uniqueness of Gibbs measures relative to Brownian motion. *Ann. I. H. Poincaré*, 39(4):877–889, 2003.

[BR] A. Beskos and G.O. Roberts. Exact simulations of diffusions. Submitted to Annals of Applied Probability.

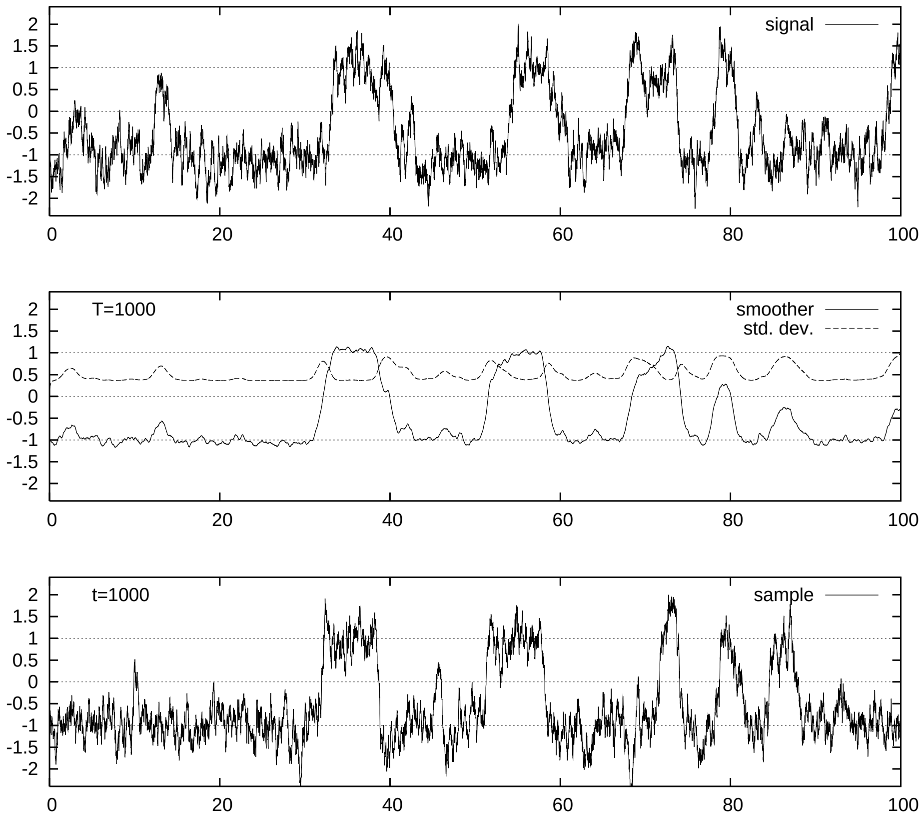

Figure 4: *Nonlinear Kalman filtering for the double-well process from section 5.3. The first diagram shows the signal X. The corresponding observation Y is not displayed. The middle diagram shows the conditional expectation and standard deviation of X given Y, obtained by evaluating the ergodic average from* (1.3) *at $T = 1000$. The third diagram gives a sample from the conditional distribution of X given Y, namely $x(\,\cdot\,, 1000)$ where x solves the SPDE* (5.4).

[BS94] J.-M. Bernardo and A.F.M. Smith. *Bayesian theory.* John Wiley & Sons Ltd., Chichester, 1994.

[CPS] S. Chib, M. Pitt, and N. Shephard. Likelihood based inference for diffusion driven models. In preparation.

[DBG02] C. Dellago, P. G. Bolhuis, and P. L. Geissler. Transition path sampling. *Adv. Chem. Phys.*, 123:1–78, 2002.

[DdFG01] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice.* Springer, 2001.

[DPZ92] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions.* Cambridge University Press, 1992.

[ERVE] W. E, W. Ren, and E. Vanden-Eijnden. Finite temperature string method for the study of rare events. In preparation.

[ERVE02] W. E, W. Ren, and E. Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B.*, 66, 2002.

[HSVW] M. Hairer, A. M. Stuart, J. Voss, and P. Wiberg. Analysis of some SPDEs arising in path sampling. In preparation.

[KP99] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations.* Number 23 in Applications of Mathematics. Springer, 1999. Corrected Third Printing.

[Liu01] J. S. Liu. *Monte Carlo strategies in scientific computing.* Springer Series in Statistics. Springer, New York, 2001.

[Øks98] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer, Berlin, fifth edition, 1998.

[RC99] C. P. Robert and G. Casella. *Monte Carlo statistical methods.* Springer, New York, 1999.

[RGG97] G.O. Roberts, A. Gelman, and W.R. Gilks. Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Prob.*, 7:110–120, 1997.

[RS01] G.O. Roberts and O. Stramer. Bayesian inference for incomplete observations of diffusion processes. *Biometrika*, 88:603–221, 2001.

[RSV] G.O. Roberts, A.M. Stuart, and J. Voss. Langevin sampling in path space: Implicit methods and pre-conditioning. In preparation.

[RVE04] M. Reznikoff and E. Vanden Eijnden. Invariant measures of SPDEs and conditioned diffusions. Preprint, 2004.