Analysis of SPDEs Arising in Path Sampling
Part I: The Gaussian Case

October 21, 2005

M. Hairer[1], **A. M. Stuart**[1], **J. Voss**[1,2], **and P. Wiberg**[1,3]

[1] Mathematics Institute, The University of Warwick, Coventry CV4 7AL, UK
[2] Previously at University of Kaiserslautern, 67653 Kaiserslautern, Germany;
supported by Marie Curie Fellowship HPMT-CT-2000-00076
[3] Now at Goldman-Sachs, London

Abstract

In many applications it is important to be able to sample paths of SDEs conditional on observations of various kinds. This paper studies SPDEs which solve such sampling problems. The SPDE may be viewed as an infinite dimensional analogue of the Langevin SDE used in finite dimensional sampling. Here the theory is developed for conditioned Gaussian processes for which the resulting SPDE is linear. Applications include the Kalman-Bucy filter/smoother. A companion paper studies the nonlinear case, building on the linear analysis provided here.

KEYWORDS SPDEs, Conditioned Diffusions, Gaussian Processes, Kalman-Bucy Filter, High Dimensional Sampling.

AMS SUBJECT CLASSIFICATIONS 60G15, 60G35, 60H10, 60H15

1 Introduction

An important basic concept in sampling is Langevin dynamics: suppose a target density p on $\mathbb{R}^d$ has the form $p(x) = c\exp(-V(x))$. Then 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. Thus, assuming that (1.1) is ergodic, $x(t)$ produces samples from the target density p as $t \to \infty$. (For details see, for example, [RC99].)

In [SVW04] we give an heuristic approach to generalising the Langevin method to an infinite dimensional setting. We derive stochastic partial differential equations (SPDEs) which are the infinite dimensional analogue of (1.1). These SPDEs sample from paths of stochastic differential equations, conditional on observations. Observations which can be incorporated into this framework include knowledge of the solution at two points (bridges) and a set-up which includes the Kalman-Bucy filter/smoother. For bridge sampling the SPDEs are also derived in [RVE05], their motivation being to understand the invariant measures of SPDEs through bridge processes. The Girsanov transformation is used to study the connection between SPDEs and bridge processes in [RVE05]; it is also used to study Gibbs measures on $\mathbb{R}$ in [BL03]. However the results concerning bridges in this paper are not a linear subcase of those papers because we consider non-symmetric drifts (which are hence not gradient) and covariance of the noise which is not proportional to the identity. Furthermore the nonlinear results in [HSV] include the results stated in [RVE05] as a subset, both because of the form of the nonlinearity and noise, and because of the wide-ranging forms of conditioning that we consider.

In the current paper we give a rigorous treatment of this SPDE based sampling method when the processes to be sampled are linear and Gaussian. The resulting SPDEs for the sampling are also linear and Gaussian in this case. We find it useful to present the Gaussian theory of SPDE based sampling for conditioned diffusions in a self-contained fashion for the following reasons.

- For nonlinear problems the SPDE based samplers can be quite competitive. A companion article [HSV] will build on the analysis in this paper to analyse SPDEs which sample paths from nonlinear SDEs, conditional on observations. The mathematical techniques are quite different from the Gaussian methods used here and hence we present them in a separate paper. However the desired path-space measures there will be characterised by calculating the density with respect to the Gaussian measures calculated here.

- We derive an explicit description of the Kalman/Bucy smoother via the solution of a linear two-point boundary value problem. This is not something that we have found in the existing literature; it is strongly suggestive that for off-line smoothing of Gaussian processes there is the potential for application of a range of fast techniques available in the computational mathematics literature, and different from the usual forward/backward implementation of the filter/smoother. See section 4.

- For Gaussian processes, the SPDEs studied here will not usually constitute the optimal way to sample, because of the time correlation inherent in the SPDE; better methods can be developed to generate independent samples by factorising the covariance operator. However these better methods can be viewed as a particular discretisation of the SPDEs written down in this paper, and this connection is of both theoretical interest and practical use, including as the basis for algorithms in the nonlinear case. See section 5 and [RSV].

In section 2 of this article we will develop a general MCMC method to sample from a given Gaussian process. It transpires that the distribution of a centred Gaussian process coincides with the invariant distribution of the L^2-valued SDE

$$\frac{dx}{dt} = \mathcal{L}x - \mathcal{L}m + \sqrt{2}\,\frac{dw}{dt} \qquad \forall t \in (0,\infty), \tag{1.2}$$

where $\mathcal{L}$ is the inverse of the covariance operator, m is the mean of the process and w is a cylindrical Wiener process.

The first sampling problems we consider are governed by paths of the $\mathbb{R}^d$-valued linear SDE

$$\frac{dX}{du}(u) = AX(u) + B\,\frac{dW}{du}(u) \qquad \forall u \in [0,1] \tag{1.3}$$

subject to observations of the initial point $X(0)$, as well as possibly the end-point $X(1)$. Here we have $A, B \in \mathbb{R}^{d\times d}$ and W is a standard d-dimensional Brownian motion. Since the SDE is linear, the solution X is a Gaussian process. Section 3 identifies the operator $\mathcal{L}$ in the case where we sample solutions of (1.3), subject to end-point conditions. In fact, $\mathcal{L}$ is a second order differential operator with boundary conditions reflecting the nature of the observations and thus we can write (1.2) as an SPDE.

In section 4 we study the situation where two processes X and Y solve the linear system of SDEs

$$\begin{aligned}\frac{dX}{du}(u) &= A_{11}X(u) + B_{11}\,\frac{dW_x}{du}(u)\\ \frac{dY}{du}(u) &= A_{21}X(u) + B_{22}\,\frac{dW_y}{du}(u)\end{aligned}$$

on $[0, 1]$ and we want to sample paths from the distribution of X (the signal) conditioned on Y (the observation). Again, we identify the operator $\mathcal{L}$ in (1.2) as a second order differential operator and derive an SPDE with this distribution as its invariant distribution. We also give a separate proof that the mean of the invariant measure of the SPDE coincides with the standard algorithmic implementation of the Kalman-Bucy filter/smoother through forward/backward sweeps.

Section 5 contains some brief remarks concerning the process of discretising SPDEs to create samplers, and section 6 contains our conclusions.

To avoid confusion we use the following naming convention. Solutions to SDEs like (1.3) which give our target distributions are denoted by upper case letters. Solutions to infinite dimensional Langevin equations like (1.2) which we use to sample from these target distributions are denoted by lower case letters.

2 Gaussian Processes

In this section we will derive a Hilbert space valued SDE to sample from arbitrary Gaussian processes.

Recall that a random variable X taking values in a separable Hilbert space $\mathcal{H}$ is said to be *Gaussian* if the law of $\langle y, X\rangle$ is Gaussian for every $y \in \mathcal{H}$ (Dirac measures are considered as Gaussian for this purpose). It is called *centred* if $\mathbb{E}\langle y, X\rangle = 0$ for every $y \in \mathcal{H}$. Gaussian random variables are determined by their mean $m = \mathbb{E}X \in \mathcal{H}$ and their covariance operator $\mathcal{C}\colon \mathcal{H} \to \mathcal{H}$ defined by

$$\langle y, \mathcal{C}x\rangle = \mathbb{E}(\langle y, X - m\rangle\langle X - m, x\rangle).$$

For details see *e.g.* [DPZ92, section 2.3.2]. The following lemma (see [DPZ92, proposition 2.15]) characterises the covariance operators of Gaussian measures.

Lemma 2.1 *Let X be a Gaussian random variable on a separable Hilbert space. Then the covariance operator $\mathcal{C}$ of X is self-adjoint, positive and trace class.*

A Gaussian random variable is said to be *non-degenerate* if $\langle y, \mathcal{C}y\rangle > 0$ for every $y \in \mathcal{H} \setminus \{0\}$. An equivalent characterisation is that the law of $\langle y, X\rangle$ is a proper Gaussian measure (*i.e.* not a Dirac measure) for every $y \in \mathcal{H} \setminus \{0\}$. Here we will always consider non-degenerate Gaussian measures. Then $\mathcal{C}$ is strictly positive definite and we can define $\mathcal{L}$ to be the inverse of $-\mathcal{C}$. Since $\mathcal{C}$ is trace class, it is also bounded and thus the spectrum of $\mathcal{L}$ is bounded away from 0.

We now construct an infinite dimensional process which, in equilibrium, samples from a prescribed Gaussian measure. Denote by w the cylindrical Wiener process on $\mathcal{H}$. Then one has formally

$$w(t) = \sum_{n=1}^{\infty} \beta_n(t)\phi_n \quad \forall t \in (0,\infty), \tag{2.1}$$

where for $n \in \mathbb{N}$ the β_n are i.i.d. standard Brownian motions and ϕ_n are the (orthonormal) eigenvectors of $\mathcal{C}$. Note that the sum (2.1) does *not* converge in $\mathcal{H}$ but that one can make sense of it by embedding $\mathcal{H}$ into a larger Hilbert space in such a way that the embedding is Hilbert-Schmidt. The choice of this larger space does not affect any of the subsequent expressions (see also [DPZ92] for further details).

Given $\mathcal{C}$ and $\mathcal{L}$ as above, consider the $\mathcal{H}$-valued SDE given by (1.2), interpreted in the following way:

$$x(t) = m + \mathrm{e}^{\mathcal{L}t}(x(0) - m) + \sqrt{2}\int_0^t \mathrm{e}^{\mathcal{L}(t-s)}\, dw(s). \tag{2.2}$$

If $x \in C([0,T],\mathcal{H})$ satisfies (2.2) it is called a *mild* solution of the SDE (1.2). We have the following result.

Lemma 2.2 *Let $\mathcal{C}$ be the covariance operator and m the mean of a non-degenerate Gaussian random variable X on a separable Hilbert space $\mathcal{H}$. Then the corresponding evolution equation (1.2) with $\mathcal{L} = -\mathcal{C}^{-1}$ has continuous $\mathcal{H}$-valued mild solutions. Furthermore, it has a unique invariant measure μ on $\mathcal{H}$ which is Gaussian with mean m and covariance $\mathcal{C}$ and there exists a constant K such that for every initial condition $x_0 \in \mathcal{H}$ one has*

$$\|\mathrm{law}(x(t)) - \mu\|_{\mathrm{TV}} \le K\,(1 + \|x_0 - m\|_{\mathcal{H}})\exp(-\|\mathcal{C}\|^{-1}_{\mathcal{H}\to\mathcal{H}}t),$$

where $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation distance between measures.

Proof. The existence of a continuous $\mathcal{H}$-valued solution of the SDE (1.2) is established in [IMM$^+$90]. The uniqueness of the invariant measure and the convergence rate in the total variation distance follow by combining Theorems 6.3.3 and 7.1.1 from [DPZ96]. The characterisation of the invariant measure is established in [DPZ96, Thm 6.2.1]. □

We can both characterise the invariant measure, and explain the exponential rate of convergence to it, by using the Karhunen-Loève expansion. In particular we give an heuristic argument which illustrates why Lemma 2.2 holds in the case $m = 0$: denote by $(\phi_n)_{n\in\mathbb{N}}$ an orthonormal basis of eigenvectors of $\mathcal{C}$ and by $(\lambda_n)_{n\in\mathbb{N}}$ the corresponding eigenvalues. If X is centred it is possible to expand X as

$$X = \sum_{n=1}^{\infty} \alpha_n \sqrt{\lambda_n}\phi_n, \tag{2.3}$$

for some real-valued random variables α_n. (In contrast to the situation in (2.1) the convergence in (2.3) actually holds in $L^2(\Omega,\mathbb{P},\mathcal{H})$, where $(\Omega,\mathbb{P})$ is the underlying probability space.) A simple calculation shows that the coefficients α_n are i.i.d. $\mathcal{N}(0,1)$ distributed random variables. The expansion (2.3) is called the Karhunen-Loève expansion. Details about this construction can be found in [Adl81].

Now express the solution x of (1.2) in the basis (ϕ_n) as

$$x(t) = \sum_{n=1}^{\infty} \gamma_n(t)\phi_n.$$

Then a formal calculation using (2.1) and (1.2) leads to the SDE

$$\frac{d\gamma_n}{dt} = -\frac{\gamma_n}{\lambda_n} + \sqrt{2}\frac{d\beta_n}{dt}$$

for the time evolution of the coefficients γ_n and hence γ_n is ergodic with stationary distribution $\mathcal{N}(0, \lambda_n)$ for every $n \in \mathbb{N}$. Thus the stationary distribution of (1.2) has the same Karhunen-Loève expansion as the distribution of X and the two distributions are the same. Furthermore, the fact that the rate of convergence to stationarity is bounded independently of n is a manifestation of the exponential rate of convergence to stationarity stated in Lemma 2.2.

In this article, the Hilbert space $\mathcal{H}$ will always be the space $L^2([0,1], \mathbb{R}^d)$ of square integrable $\mathbb{R}^d$-valued functions and the Gaussian measures we consider will be distributions of Gaussian processes. In this case the operator $\mathcal{C}$ has a kernel $C\colon [0,1]^2 \to \mathbb{R}^{d\times d}$ such that

$$(\mathcal{C}x)(u) = \int_0^1 C(u,v)\, x(v)\, dv. \tag{2.4}$$

If the covariance function C is Hölder continuous, then the Kolmogorov continuity criterion (see *e.g.* [DPZ92, Thm 3.3]) ensures that X is almost surely a continuous function from $[0,1]$ to $\mathbb{R}^d$. In this case C is given by the formula

$$C(u,v) = \mathbb{E}\Big((X(u) - m(u))(X(v) - m(v))^*\Big)$$

and the convergence of the expansion (2.3) is uniform with probability one.

Remark 2.3 The solution of (1.2) may be viewed as the basis for an MCMC method for sampling from a given Gaussian process. The key to exploiting this fact is the identification of the operator $\mathcal{L}$ for a given Gaussian process. In the next section we show that, for a variety of linear SDEs, $\mathcal{L}$ is a second order differential operator and hence (1.2) is a stochastic partial differential equation. If $\mathcal{C}$ has a Hölder continuous kernel C, it follows from (2.4) and the relation $\mathcal{C} = (-\mathcal{L})^{-1}$ that it suffices to find a differential operator $\mathcal{L}$ such that $C(u,v)$ is the Green's function of $-\mathcal{L}$.

3 Conditioned Linear SDEs

In this section we apply our sampling technique from section 2 to Gaussian measures which are given as the distributions of a number of conditioned linear SDEs. We condition on, in turn, a single known point (subsection 3.1), a single point with Gaussian distribution (subsection 3.2) and finally a bridge between two points (subsection 3.3).

Throughout we consider the $\mathbb{R}^d$-valued SDE

$$\frac{dX}{du}(u) = AX(u) + B\,\frac{dW}{du}(u), \qquad \forall u \in [0,1], \tag{3.1}$$

where $A, B \in \mathbb{R}^{d\times d}$ and W is the standard d-dimensional Brownian motion. We assume that the matrix BB^* is invertible. We associate to (3.1) the second order differential operator L formally given by

$$L = (\partial_u + A^*)(BB^*)^{-1}(\partial_u - A). \tag{3.2}$$

When equipped with homogeneous boundary conditions through its domain of definition, we will denote the operator (3.2) by $\mathcal{L}$. We will always consider boundary conditions of the general form $D_0 x(0) = 0$ and $D_1 x(1) = 0$, where $D_i = A_i\partial_u + b_i$ are first-order differential operators.

Remark 3.1 We will repeatedly write $\mathbb{R}^d$-valued SPDEs with inhomogeneous boundary conditions of the type

$$\begin{aligned} \partial_t x(t,u) = Lx(t,u) + g(u) + \sqrt{2}\,\partial_t w(t,u) \quad &\forall (t,u) \in (0,\infty)\times[0,1],\\ D_0 x(t,0) = a, \quad D_1 x(t,1) = b \qquad &\forall t \in (0,\infty),\\ x(0,u) = x_0(u) \qquad &\forall u \in [0,1] \end{aligned} \tag{3.3}$$

where $g\colon [0,1] \to \mathbb{R}^d$ is a function, $\partial_t w$ is space-time white noise, and $a, b \in \mathbb{R}^d$. We call a process x a solution of this SPDE if it solves (2.2) with $x(0) = x_0$ where $\mathcal{L}$ is L equipped with the boundary conditions $D_0 f(0) = 0$ and $D_1 f(1) = 0$, and $m\colon [0,1] \to \mathbb{R}^d$ is the solution of the boundary value problem $-Lm = g$ with boundary conditions $D_0 m(0) = a$ and $D_1 m(1) = b$.

To understand the connection between (3.3) and (2.2) note that, if w is a smooth function, then the solutions of both equations coincide.

3.1 Fixed Left End-Point

Consider the problem of sampling paths of (3.1) subject only to the initial condition

$$X(0) = x^- \in \mathbb{R}^d. \tag{3.4}$$

The solution of this SDE is a Gaussian process with mean

$$m(u) = E(X(u)) = \mathrm{e}^{uA}x^- \tag{3.5}$$

and covariance function

$$C_0(u,v) = \mathrm{e}^{uA}\Big(\int_0^{u\wedge v} \mathrm{e}^{-rA}BB^*\mathrm{e}^{-rA^*}\,dr\Big)\mathrm{e}^{vA^*} \tag{3.6}$$

(see *e.g.* [KS91, section 5.6] for reference). Let $\mathcal{L}$ denote the differential operator L from (3.2) with the domain of definition

$$\mathcal{D}(\mathcal{L}) = \{f \in H^2([0,1],\mathbb{R}^d) \mid f(0) = 0, \frac{d}{du}f(1) = Af(1)\}. \tag{3.7}$$

Lemma 3.2 *With $\mathcal{L}$ given by (3.2) and (3.7) the function C_0 is the Green's function for $-\mathcal{L}$. That is*

$$LC_0(u,v) = -\delta(u-v)I$$

and

$$C_0(0,v) = 0, \quad \partial_u C_0(1,v) = A\,C_0(1,v) \qquad \forall v \in (0,1).$$

Proof. From (3.6) it is clear that the left-hand boundary condition $C_0(0,v) = 0$ is satisfied for all $v \in [0,1]$. It also follows that, for $u \neq v$, the kernel is differentiable with derivative

$$\partial_u C_0(u,v) = \begin{cases} AC_0(u,v) + BB^* \mathrm{e}^{-uA^*} \mathrm{e}^{vA^*}, & \text{for } u < v, \text{ and} \\ AC_0(u,v) & \text{for } u > v. \end{cases} \tag{3.8}$$

Thus the kernel C_0 satisfies the boundary condition $\partial_u C_0(1,v) = AC_0(1,v)$ for all $v \in [0,1)$.

Equation (3.8) shows

$$(BB^*)^{-1}(\partial_u - A)C_0(u,v) = \begin{cases} \mathrm{e}^{-uA^*} \mathrm{e}^{vA^*}, & \text{for } u < v, \text{ and} \\ 0 & \text{for } u > v \end{cases} \tag{3.9}$$

and thus we get

$$LC_0(u,v) = (\partial_u + A^*)(BB^*)^{-1}(\partial_u - A)C_0(u,v) = 0 \qquad \forall u \neq v.$$

Now let $v \in (0,1)$. Then we get

$$\lim_{u \uparrow v} (BB^*)^{-1}(\partial_u - A)C_0(u,v) = I$$

and

$$\lim_{u \downarrow v} (BB^*)^{-1}(\partial_u - A)C_0(u,v) = 0$$

This shows $LC_0(u,v) = -\delta(u-v)I$ for all $v \in (0,1)$. □

Now that we have identified the operator $\mathcal{L} = (-\mathcal{C})^{-1}$ we are in the situation of Lemma 2.2 and can derive an SPDE to sample paths of (3.1), subject to the initial condition (3.4). We formulate this result precisely in the following theorem.

Theorem 3.3 *For every $x_0 \in \mathcal{H}$ the $\mathbb{R}^d$-valued SPDE*

$$\partial_t x(t,u) = Lx(t,u) + \sqrt{2}\,\partial_t w(t,u) \qquad \forall (t,u) \in (0,\infty) \times (0,1) \tag{3.10a}$$

$$x(t,0) = x^-, \quad \partial_u x(t,1) = Ax(t,1) \qquad \forall t \in (0,\infty) \tag{3.10b}$$

$$x(0,u) = x_0(u) \qquad \forall u \in [0,1] \tag{3.10c}$$

where $\partial_t w$ is space-time white noise has a unique mild solution. The SPDE is ergodic and in equilibrium samples paths of the SDE (3.1) with initial condition $X(0) = x^-$.

Proof. The solution of SDE (3.1) with initial condition (3.4) is a Gaussian process where the mean m is given by (3.5). The mean m solves the boundary value problem $Lm(u) = 0$ for all $u \in (0,1)$, $m(0) = x^-$ and $m'(1) = Am(1)$. From Remark 3.1 we find that x is a solution of the Hilbert space valued SDE (1.2) for this function m.

Lemma 3.2 shows that $\mathcal{L}$, given by (3.2) with the boundary conditions from (3.10b), is the inverse of $-\mathcal{C}$ where $\mathcal{C}$ is the covariance operator of the distribution we want to sample from (and with covariance function given by (3.6)). Lemma 2.2 then shows that the SPDE (3.10) is ergodic and that its stationary distribution coincides with the distribution of solutions of the SDE (3.1) with initial condition $X(0) = x^-$. □

3.2 Gaussian Left End-Point

An argument similar to the one in section 3.1 deals with sampling paths of (3.1) where $X(0)$ is a Gaussian random variable distributed as

$$X(0) \sim \mathcal{N}(x^-, \Sigma) \tag{3.11}$$

with an invertible covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$ and independent of the Brownian motion W.

Theorem 3.4 *For every $x_0 \in \mathcal{H}$ the $\mathbb{R}^d$-valued SPDE*

$$\partial_t x(t,u) = Lx(t,u) + \sqrt{2}\,\partial_t w(t,u) \qquad \forall (t,u) \in (0,\infty)\times(0,1) \tag{3.12a}$$

$$\partial_u x(t,0) = Ax(t,0) + BB^*\Sigma^{-1}(x - x^-), \quad \partial_u x(t,1) = Ax(t,1) \qquad \forall t \in (0,\infty) \tag{3.12b}$$

$$x(0,u) = x_0(u) \qquad \forall u \in [0,1] \tag{3.12c}$$

where $\partial_t w$ is space-time white noise has a unique mild solution. The SPDE is ergodic and in equilibrium samples paths of the SDE (3.1) with Gaussian initial condition (3.11).

Proof. The solution X of SDE (3.1) with initial condition (3.11) is a Gaussian process with mean (3.5) and covariance function

$$C(u,v) = \mathrm{e}^{uA}\Sigma\mathrm{e}^{vA^*} + C_0(u,v), \tag{3.13}$$

where C_0 is the covariance function from (3.6) for the case $X(0) = 0$ (see Problem 6.1 in Section 5.6 of [KS91] for a reference). The mean m from (3.5) solves the boundary value problem $Lm(u) = 0$ for all $u \in (0,1)$ with boundary conditions $m'(0) = Am(0) + BB^*\Sigma^{-1}(m(0) - x^-)$ and $m'(1) = Am(1)$.

In order to identify the inverse of the covariance operator $\mathcal{C}$ we can use (3.8) to find

$$\partial_u C(u,v) = \begin{cases} AC(u,v) + BB^*\mathrm{e}^{-uA^*}\mathrm{e}^{vA^*}, & \text{for } u < v, \text{ and} \\ AC(u,v) & \text{for } u > v \end{cases}$$

and, since $C(0,v) = \Sigma\,\mathrm{e}^{vA^*}$, we get the boundary conditions

$$\partial_u C(0,v) = AC(0,v) + BB^*\Sigma^{-1}C(0,v)$$

and

$$\partial_u C(1,v) = AC(1,v).$$

From $(\partial_u - A)\mathrm{e}^{uA}\Sigma\mathrm{e}^{vA^*} = 0$ we also get

$$LC(u,v) = L\mathrm{e}^{uA}\Sigma\mathrm{e}^{vA^*} + LC_0(u,v) = 0$$

for all $u \neq v$ and $LC(u,v) = LC_0(u,v) = -\delta(u,v)I$ for all $u,v \in (0,1)$. Thus C is again the Green's function for $-\mathcal{L}$ and the claim follows from Remark 2.3 and Lemma 2.2. □

Remark 3.5 If A is negative-definite symmetric, then the solution X of SDE (3.1) has a stationary distribution which is a centred Gaussian measure with covariance $\Sigma = -\frac{1}{2}A^{-1}BB^*$. Choosing this distribution in (3.11), the boundary condition (3.12b) becomes

$$\partial_u x(t,0) = -Ax(t,0), \quad \partial_u x(t,1) = Ax(t,1) \qquad \forall t \in (0,\infty).$$

3.3 Bridge Sampling

In this section we apply our sampling method to sample from solutions of the linear SDE (3.1) with fixed end-points, *i.e.* we sample from the distribution of X conditioned on

$$X(0) = x^-, \quad X(1) = x^+. \tag{3.14}$$

The conditional distribution transpires to be absolutely continuous with respect to the Brownian bridge measure satisfying (3.14).

Let m and C_0 be the mean and covariance of the unconditioned solution X of the SDE (3.1) with initial condition $X(0) = x^-$. As we will show in Lemma 4.4 below, the solution conditioned on $X(1) = x^+$ is again a Gaussian process. The mean and covariance of the conditioned process can be found by conditioning the random variable $(X(u), X(v), X(1))$ for $u \leq v \leq 1$ on the value of $X(1)$. Since this is a finite dimensional Gaussian random variable, mean and covariance of the conditional distribution can be explicitly calculated. The result for the mean is

$$\tilde{m}(u) = m(u) + C_0(u,1)C_0(1,1)^{-1}(x^+ - m(1)) \tag{3.15}$$

and for the covariance function we get

$$\tilde{C}(u,v) = C_0(u,v) - C_0(u,1)C_0(1,1)^{-1}C_0(1,v). \tag{3.16}$$

Theorem 3.6 *For every $x_0 \in \mathcal{H}$ the $\mathbb{R}^d$-valued SPDE*

$$\partial_t x = Lx + \sqrt{2}\,\partial_t w \qquad \forall (t,u) \in (0,\infty) \times (0,1) \tag{3.17a}$$

$$x(t,0) = x^-, \quad x(t,1) = x^+ \qquad \forall t \in (0,\infty) \tag{3.17b}$$

$$x(0,u) = x_0(u) \qquad \forall u \in [0,1] \tag{3.17c}$$

where $\partial_t w$ is white noise has a unique mild solution. The SPDE is ergodic and in equilibrium samples paths of the SDE (3.1) subject to the bridge conditions (3.14).

Proof. The solution of the SDE (3.1) with boundary conditions (3.14) is a Gaussian process where the mean $\tilde{m}$ is given by (3.15) and the covariance function $\tilde{C}$ is given by (3.16). From formula (3.9) we know $LC_0(u,1) = 0$ and thus $\tilde{m}$ satisfies $L\tilde{m} = Lm = 0$. Since $\tilde{m}(0) = x^-$ and $\tilde{m}(t) = m(1) + C_0(1,1)C_0(1,1)^{-1}(x^+ - m(1)) = x^+$, the mean $\tilde{m}$ solves the boundary value problem $L\tilde{m}(u) = 0$ for all $u \in (0,1)$ with boundary conditions $\tilde{m}(0) = x^-$ and $\tilde{m}(1) = x^+$.

It remains to show that $\tilde{C}$ is the Green's function for the operator L with homogeneous Dirichlet boundary conditions: we have $\tilde{C}(0,v) = 0$,

$$\tilde{C}(1,v) = C_0(1,v) - C_0(1,1)C_0(1,1)^{-1}C_0(1,v) = 0$$

and using $LC_0(u,1) = 0$ we find

$$L\tilde{C}(u,v) = LC_0(u,v) = -\delta(u-v)I.$$

This completes the proof. □

4 The Kalman-Bucy Filter/Smoother

Consider (3.1) with X replaced by the $\mathbb{R}^m \times \mathbb{R}^n$-valued process (X, Y) and $A, B \in \mathbb{R}^{(m+n)\times(m+n)}$ chosen so as to obtain the linear SDE

$$\frac{d}{du}\begin{pmatrix} X(u) \\ Y(u) \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ A_{21} & 0 \end{pmatrix}\begin{pmatrix} X(u) \\ Y(u) \end{pmatrix} + \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix}\frac{d}{du}\begin{pmatrix} W_x(u) \\ W_y(u) \end{pmatrix}. \tag{4.1a}$$

We impose the conditions

$$X_0 \sim \mathcal{N}(x^-, \Lambda), \quad Y_0 = 0 \tag{4.1b}$$

and try to sample from paths of X given paths of Y. We derive an SPDE whose invariant measure is the conditional distribution of X given Y. Formally this SPDE is found by writing the SPDE for sampling from the solution (X, Y) of (4.1) and considering the equation for the evolution of x, viewing $y \equiv Y$ as known. This leads to the following result.

Theorem 4.1 *Given a path Y sampled from (4.1) consider the SPDE*

$$\begin{aligned}\partial_t x = \Big((\partial_u + A_{11}^*)(B_{11}B_{11}^*)^{-1}(\partial_u - A_{11})\Big)x \\ + A_{21}^*(B_{22}B_{22}^*)^{-1}\Big(\frac{dY}{du} - A_{21}x\Big) + \sqrt{2}\,\partial_t w,\end{aligned} \tag{4.2a}$$

equipped with the inhomogeneous boundary conditions

$$\begin{aligned}\partial_u x(t,0) &= A_{11}x(t,0) + B_{11}B_{11}^*\Lambda^{-1}(x(t,0) - x^-), \\ \partial_u x(t,1) &= A_{11}x(t,1)\end{aligned} \tag{4.2b}$$

and initial condition

$$x(0,u) = x_0(u) \qquad \forall u \in [0,1]. \tag{4.2c}$$

Then for every $x_0 \in \mathcal{H}$ the SPDE has a unique mild solution and is ergodic. Its stationary distribution coincides with the conditional distribution of X given Y for X, Y solving (4.1).

The proof of this theorem is based on the following three lemmas concerning conditioned Gaussian processes. After deriving these three lemmas we give the proof of Theorem 4.1. The section finishes with a direct proof that the mean of the invariant measure coincides with the standard algorithmic implementation of the Kalman-Bucy filter/smoother through forward/backward sweeps (this fact is implicit in Theorem 4.1).

Lemma 4.2 *Let $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ be a separable Hilbert space with projectors $\Pi_i \colon \mathcal{H} \to \mathcal{H}_i$. Let $\mathcal{C} \colon \mathcal{H} \to \mathcal{H}$ be a positive definite, bounded, linear, self-adjoint operator and denote $\mathcal{C}_{ij} = \Pi_i \mathcal{C} \Pi_j^*$. Then $\mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21}$ is positive definite and if $\mathcal{C}_{11}$ is trace class then the operator $\mathcal{C}_{12}\mathcal{C}_{22}^{-\frac{1}{2}}$ is Hilbert-Schmidt.*

Proof. Since $\mathcal{C}$ is positive definite, one has

$$2|\langle \mathcal{C}_{21}x, y\rangle| \le \langle x, \mathcal{C}_{11}x\rangle + \langle y, \mathcal{C}_{22}y\rangle,$$

for every $(x, y) \in \mathcal{H}$. It follows that

$$|\langle \mathcal{C}_{21}x, y\rangle|^2 \leq \langle x, \mathcal{C}_{11}x\rangle\langle y, \mathcal{C}_{22}y\rangle, \tag{4.3}$$

and so

$$|\langle \mathcal{C}_{21}x, \mathcal{C}_{22}^{-1/2}y\rangle|^2 \leq \langle x, \mathcal{C}_{11}x\rangle\|y\|^2 \tag{4.4}$$

for every $y \neq 0$ in the range of $\mathcal{C}_{22}^{1/2}$. Equation (4.3) implies that $\mathcal{C}_{21}x$ is orthogonal to $\ker \mathcal{C}_{22}$ for every $x \in \mathcal{H}_1$. Therefore the operator $\mathcal{C}_{22}^{-1/2}\mathcal{C}_{21}$ can be defined on all of $\mathcal{H}_1$ and thus is bounded. Taking $y = \mathcal{C}_{22}^{-1/2}\mathcal{C}_{21}x$ in (4.4) gives $\|\mathcal{C}_{22}^{-1/2}\mathcal{C}_{21}x\|^2 \leq \langle x, \mathcal{C}_{11}x\rangle$ and thus $\langle x, (\mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21})x\rangle \geq 0$ for every $x \in \mathcal{H}_1$. This implies that $\mathcal{C}_{22}^{-\frac{1}{2}}\mathcal{C}_{21}$ and $\mathcal{C}_{12}\mathcal{C}_{22}^{-\frac{1}{2}}$ are both Hilbert-Schmidt, and completes the proof. □

Remark 4.3 Note that $\mathcal{C}$ being strictly positive definite is not sufficient to imply that $\mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21}$ is also strictly positive definite. A counter-example can be constructed by considering the Wiener measure on $\mathcal{H} = L^2([0,1])$ with $\mathcal{H}_1$ being the linear space spanned by the constant function 1.

Lemma 4.4 *Let $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ be a separable Hilbert space with projectors $\Pi_i : \mathcal{H} \to \mathcal{H}_i$. Let (X_1, X_2) be an $\mathcal{H}$-valued Gaussian random variable with mean $m = (m_1, m_2)$ and positive definite covariance operator $\mathcal{C}$ and define $\mathcal{C}_{ij} = \Pi_i\mathcal{C}\Pi_j^*$. Then the conditional distribution of X_1 given X_2 is Gaussian with mean*

$$m_{1|2} = m_1 + \mathcal{C}_{12}\mathcal{C}_{22}^{-1}(X_2 - m_2) \tag{4.5}$$

and covariance operator

$$\mathcal{C}_{1|2} = \mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21}. \tag{4.6}$$

Proof. Note that by Lemma 2.1 the operator $\mathcal{C}$ is trace class. Thus $\mathcal{C}_{11}$ and $\mathcal{C}_{22}$ are also trace class. Let μ be the law of X_2 and let $\mathcal{H}_0$ be the range of $\mathcal{C}_{22}^{1/2}$ equipped with the inner product

$$\langle x, y\rangle_0 = \langle C_{22}^{-1/2}x, C_{22}^{-1/2}y\rangle.$$

If we embed $\mathcal{H}_0 \hookrightarrow \mathcal{H}_2$ via the trivial injection $i(f) = f$, then we find $i^*(f) = C_{22}f$. Since $i \circ i^* = C_{22}$ is the covariance operator of μ, the space $\mathcal{H}_0$ is its reproducing kernel Hilbert space. From Lemma 4.2 we know that $\mathcal{C}_{12}\mathcal{C}_{22}^{-1/2}$ is Hilbert-Schmidt from $\mathcal{H}_2$ to $\mathcal{H}_1$ and hence bounded. Thus we can define

$$A = \mathcal{C}_{12}\mathcal{C}_{22}^{-1/2}\,\mathcal{C}_{22}^{-1/2} = \mathcal{C}_{12}\mathcal{C}_{22}^{-1}$$

as a bounded operator from $\mathcal{H}_0$ to $\mathcal{H}_1$.

Let $(\phi_n)_n$ be an orthonormal basis of $\mathcal{H}_2$. Then $\psi_n = C_{22}^{1/2}\phi_n$ defines an orthonormal basis on $\mathcal{H}_0$ and we get

$$\sum_{n\in\mathbb{N}} \|A\psi_n\|_{\mathcal{H}_1}^2 = \sum_{n\in\mathbb{N}} \|\mathcal{C}_{12}\mathcal{C}_{22}^{-1}\,\mathcal{C}_{22}^{1/2}\phi_n\|_{\mathcal{H}_1}^2 = \sum_{n\in\mathbb{N}} \|\mathcal{C}_{12}\mathcal{C}_{22}^{-1/2}\phi_n\|_{\mathcal{H}_1}^2 < \infty,$$

where the last inequality comes from Lemma 4.2. This shows that the operator A is Hilbert-Schmidt on the reproducing kernel Hilbert space $\mathcal{H}_0$. Theorem II.3.3 of [DF91] shows that A can be extended in a measurable way to a subset of $\mathcal{H}_2$ which has full measure, so that (4.5) is well-defined.

Now consider the process Y defined by

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} I_{\mathcal{H}_1} & -A \\ 0_{\mathcal{H}_2} & I_{\mathcal{H}_2} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}.$$

This process is also Gaussian, but with mean

$$m^Y = \begin{pmatrix} I_{\mathcal{H}_1} & -A \\ 0_{\mathcal{H}_2} & I_{\mathcal{H}_2} \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} m_1 - Am_2 \\ m_2 \end{pmatrix}$$

and covariance operator

$$\mathcal{C}^Y = \begin{pmatrix} I_{\mathcal{H}_1} & -A \\ 0_{\mathcal{H}_2} & I_{\mathcal{H}_2} \end{pmatrix} \begin{pmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} \\ \mathcal{C}_{21} & \mathcal{C}_{22} \end{pmatrix} \begin{pmatrix} I_{\mathcal{H}_1} & 0_{\mathcal{H}_2} \\ -A^* & I_{\mathcal{H}_2} \end{pmatrix} = \begin{pmatrix} \mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21} & 0 \\ 0 & \mathcal{C}_{22} \end{pmatrix}.$$

This shows that $Y_1 = X_1 - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}X_2$ and $Y_2 = X_2$ are uncorrelated and thus independent. So we get

$$\begin{aligned} \mathbb{E}(X_1 \mid X_2) &= \mathbb{E}(X_1 - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}X_2 \mid X_2) + \mathbb{E}(\mathcal{C}_{12}\mathcal{C}_{22}^{-1}X_2 \mid X_2) \\ &= \mathbb{E}(X_1 - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}X_2) + \mathcal{C}_{12}\mathcal{C}_{22}^{-1}X_2 \\ &= m_1 - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}m_2 + \mathcal{C}_{12}\mathcal{C}_{22}^{-1}X_2. \end{aligned}$$

This proves (4.5) and a similar calculation gives equality (4.6). □

Remark 4.5 If we define as above $\mathcal{L} = (-\mathcal{C})^{-1}$ and formally define $\mathcal{L}_{ij} = \Pi_i\mathcal{L}\Pi_j^*$ (note that without additional information on the domain of $\mathcal{L}$ these operators may not be densely defined), then a simple formal calculation shows that $m_{1|2}$ and $\mathcal{C}_{1|2}$ are expected to be given by

$$m_{1|2} = m_1 - \mathcal{L}_{11}^{-1}\mathcal{L}_{12}(X_2 - m_2), \qquad \mathcal{C}_{1|2} = -\mathcal{L}_{11}^{-1}. \tag{4.7}$$

We now justify these relations in a particular situation which is adapted to the case that will be considered in the remaining part of this section.

Lemma 4.6 *Consider the setup of Lemma 4.4 and Remark 4.5 and assume furthermore that the following properties are satisfied:*

a. *The operator $\mathcal{L}$ can be extended to a closed operator $\tilde{\mathcal{L}}$ on $\Pi_1\mathcal{D}(\mathcal{L}) \oplus \Pi_2\mathcal{D}(\mathcal{L})$.*
b. *Define the operators $\mathcal{L}_{ij} = \Pi_i\tilde{\mathcal{L}}\Pi_j^*$. Then, the operator $\mathcal{L}_{11}$ is self-adjoint and one has* $\ker \mathcal{L}_{11} = \{0\}$.
c. *The operator $-\mathcal{L}_{11}^{-1}\mathcal{L}_{12}$ can be extended to a bounded operator from $\mathcal{H}_2$ into $\mathcal{H}_1$.*

Then $\mathcal{C}_{12}\mathcal{C}_{22}^{-1}$ can be extended to a bounded operator from $\mathcal{H}_2$ into $\mathcal{H}_1$ and one has $\mathcal{C}_{12}\mathcal{C}_{22}^{-1} = -\mathcal{L}_{11}^{-1}\mathcal{L}_{12}$. Furthermore, $\mathcal{C}_{21}$ maps $\mathcal{H}_1$ into the range of $\mathcal{C}_{22}$ and one has

$$\mathcal{L}_{11}^{-1}x = (\mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21})x,$$

for every $x \in \mathcal{H}_1$.

Proof. We first show that $\mathcal{C}_{12}\mathcal{C}_{22}^{-1} = -\mathcal{L}_{11}^{-1}\mathcal{L}_{12}$. By property *a.* and the definition of $\mathcal{L}$, we have the equality

$$\tilde{\mathcal{L}}\Pi_1^*\Pi_1\mathcal{C}x + \tilde{\mathcal{L}}\Pi_2^*\Pi_2\mathcal{C}x = -x \tag{4.8}$$

for every $x \in \mathcal{H}$, and thus $\mathcal{L}_{11}\mathcal{C}_{12}x = -\mathcal{L}_{12}\mathcal{C}_{22}x$ for every $x \in \mathcal{H}_2$. It follows immediately that $\mathcal{L}_{11}\mathcal{C}_{12}\mathcal{C}_{22}^{-1}x = -\mathcal{L}_{12}x$ for every $x \in \mathcal{R}(\mathcal{C}_{22})$. Since $\mathcal{R}(\mathcal{C}_{22})$ is dense in $\mathcal{H}_2$, the statement follows from assumptions *b.* and *c.*

Let us now turn to the second equality. By property *a.* the operator $\mathcal{C}_{21}$ maps $\mathcal{H}_1$ into the domain of $\mathcal{L}_{12}$ so that

$$x = x - \mathcal{L}_{12}\mathcal{C}_{21}x + \mathcal{L}_{12}\mathcal{C}_{21}x = \mathcal{L}_{11}\mathcal{C}_{11}x + \mathcal{L}_{12}\mathcal{C}_{21}x, \tag{4.9}$$

for every $x \in \mathcal{H}_1$ (the second equality follows from an argument similar to the one that yields (4.8)). Since the operator $\mathcal{C}_{22}^{-1}$ is self-adjoint, we know from [Yos95, p. 195] that $(\mathcal{C}_{12}\mathcal{C}_{22}^{-1})^* = \mathcal{C}_{22}^{-1}\mathcal{C}_{21}$. Since the left hand side operator is densely defined and bounded, its adjoint is defined on all of $\mathcal{H}_1$, so that $\mathcal{C}_{21}$ maps $\mathcal{H}_1$ into the range of $\mathcal{C}_{22}$. It follows from (4.9) that

$$x = \mathcal{L}_{11}\mathcal{C}_{11}x + \mathcal{L}_{12}\mathcal{C}_{22}\mathcal{C}_{22}^{-1}\mathcal{C}_{21}x,$$

for every $x \in \mathcal{H}_1$. Using (4.8), this yields $x = \mathcal{L}_{11}\mathcal{C}_{11}x - \mathcal{L}_{11}\mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21}x$, so that $\mathcal{L}_{11}^{-1}$ is an extension of $\mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21}$. Since both of these operators are self-adjoint, they must agree. □

Corollary 4.7 *Let (X, Y) be Gaussian with covariance $\mathcal{C}$ and mean m on a separable Hilbert space $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. Assume furthermore that $\mathcal{C}$ satisfies the assumptions of Lemmas 4.4 and 4.6. Then, the conditional law of X given Y is given by the invariant measure of the ergodic SPDE*

$$\frac{dx}{dt} = \mathcal{L}_{11}x - \mathcal{L}_{11}\Pi_1 m + \mathcal{L}_{12}(Y - \Pi_2 m) + \sqrt{2}\,\frac{dw}{dt}, \tag{4.10}$$

where w is a cylindrical Wiener process on $\mathcal{H}_1$ and the operators $\mathcal{L}_{ij}$ are defined as in Lemma 4.6. SPDE (4.10) is again interpreted in the mild sense (2.2).

Proof. Note that $\mathcal{L}_{11}^{-1}\mathcal{L}_{12}$ can be extended to a bounded operator by assumption and the mild interpretation of (4.10) is

$$x_t = M + e^{\mathcal{L}_{11}t}(x_0 - M) + \sqrt{2}\int_0^t e^{\mathcal{L}_{11}(t-s)}\,dw(s), \tag{4.11}$$

with $M = \Pi_1 m - \mathcal{L}_{11}^{-1}\mathcal{L}_{12}(Y - \Pi_2 m)$. The result follows by combining Lemma 4.4 and Lemma 4.6 with Lemma 2.2. □

These abstract results enable us to prove the main result of this section.

Proof of Theorem 4.1. Consider a solution (X, Y) to the SDE (4.1). Introducing the shorthand notations

$$\Sigma_1 = (B_{11}B_{11}^*)^{-1}, \qquad \Sigma_2 = (B_{22}B_{22}^*)^{-1},$$

it follows by the techniques used in the proof of Theorem 3.4 that the operator $\mathcal{L}$ corresponding to its covariance is formally given by

$$\begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} := \begin{pmatrix} \partial_u + A_{11}^* & A_{21}^* \\ 0 & \partial_u \end{pmatrix} \begin{pmatrix} B_{11}B_{11}^* & 0 \\ 0 & B_{22}B_{22}^* \end{pmatrix}^{-1} \begin{pmatrix} \partial_u - A_{11} & 0 \\ -A_{21} & \partial_u \end{pmatrix}$$

$$= \begin{pmatrix} (\partial_u + A_{11}^*)\Sigma_1(\partial_u - A_{11}) - A_{21}^*\Sigma_2 A_{21} & A_{21}^*\Sigma_2\partial_u \\ -\partial_u\Sigma_2 A_{21} & \partial_u\Sigma_2\partial_u \end{pmatrix}.$$

In order to identify its domain, we consider (3.12b) with

$$\Sigma = \begin{pmatrix} \Lambda & 0 \\ 0 & \Gamma \end{pmatrix}$$

and we take the limit $\Gamma \to 0$. This leads to the boundary conditions

$$\begin{aligned} \partial_u x(0) &= A_{11}x(0) + (\Lambda\Sigma_1)^{-1}(x(0) - x^-), & \partial_u x(1) &= A_{11}x(1), \\ y(0) &= 0, & \partial_u y(1) &= A_{21}x(1). \end{aligned} \tag{4.12a}$$

The domain of $\mathcal{L}$ is thus $H^2([0,1], \mathbb{R}^m \times \mathbb{R}^n)$, equipped with the homogeneous version of these boundary conditions.

We now check that the conditions of Lemma 4.6 hold. Condition *a.* is readily verified, the operator $\tilde{\mathcal{L}}$ being equipped with the boundary conditions

$$\begin{aligned} \partial_u x(0) &= A_{11}x(0) + (\Lambda\Sigma_1)^{-1}x(0), & \partial_u x(1) &= A_{11}x(1), \\ y(0) &= 0, & \Pi\partial_u y(1) &= 0, \end{aligned} \tag{4.12b}$$

where Π is the projection on the orthogonal complement of the range of A_{21}. Note that the operator $\tilde{\mathcal{L}}$ is closed, but no longer self-adjoint (unless $A_{21} = 0$). The operator $\mathcal{L}_{11}$ is therefore given by

$$\mathcal{L}_{11} = (\partial_u + A_{11}^*)\Sigma_1(\partial_u - A_{11}) - A_{21}^*\Sigma_2 A_{21},$$

equipped with the boundary condition

$$\partial_u x(0) = A_{11}x(0) + (\Lambda\Sigma_1)^{-1}x(0), \quad \partial_u x(1) = A_{11}x(1).$$

It is clear that this operator is self-adjoint. The fact that its spectrum is bounded away from 0 follows from the fact that the form domain of $\mathcal{L}$ contains $\Pi_1^*\Pi_1\mathcal{D}(\mathcal{L})$ and that there is a $c > 0$ with $\langle a, \mathcal{L}a\rangle \le -c\|a\|^2$ for all $a \in \mathcal{D}(\mathcal{L})$. Thus condition *b.* holds.

The operator $\mathcal{L}_{12}$ is given by the first-order differential operator $A_{21}^*\Sigma_2\partial_u$ whose domain is given by functions with square-integrable second derivative that vanish at 0. Since the kernel of $\mathcal{L}_{11}^{-1}$ has a square-integrable derivative, it is easy to check that $\mathcal{L}_{11}^{-1}\mathcal{L}_{12}$ extends to a bounded operator on $\mathcal{H}$, so that condition *c.* is also verified.

We can therefore apply Lemma 4.6 and Lemma 2.2. The formulation of the equation with inhomogeneous boundary conditions is an immediate consequence of Remark 3.1: a short calculation to remove the inhomogeneity in the boundary conditions (4.2b) and change the inhomogeneity in the PDE (4.2a) shows that (4.2) can be written in the form (4.10) or (4.11) with the desired value for M, the conditional mean. Since $\mathcal{L}_{11}$ is indeed the conditional covariance operator, the proof is complete. □

Remark 4.8 For Y solving (4.1) the derivative $\frac{dY}{du}$ only exists in a distributional sense (it is in the Sobolev space $H^{-1/2-\epsilon}$ for every $\epsilon > 0$). But the definition (2.2) of a mild solution which we use here applies the inverse of the second order differential operator $\mathcal{L}_{11}$ to $\frac{dY}{du}$, resulting in an element of $H^{3/2-\epsilon}$ in the solution.

Remark 4.9 Denote by $x(t,u)$ a solution of the SPDE (4.2) and write the mean as $\bar{x}(t,u) = \mathbb{E}x(t,u)$. Then, as $t \to \infty$, $\bar{x}(t,u)$ converges to its limit $\tilde{x}(u)$ strongly in

$L^2([0,1],\mathbb{R}^m)$ and $\tilde{x}(u)$ must coincide with the Kalman-Bucy filter/smoother. This follows from the fact that $\tilde{x}$ equals $\mathbb{E}(X \mid Y)$. It is instructive to demonstrate this result directly and so we do so.

The mean $\tilde{x}(u)$ of the invariant measure of (4.2) satisfies the linear two point boundary value problem

$$\begin{aligned}(\frac{d}{du}+A_{11}^*)(B_{11}B_{11}^*)^{-1}(\frac{d}{du}-A_{11})\tilde{x}(u)\\ +A_{21}^*(B_{22}B_{22}^*)^{-1}\Big(\frac{dY}{du}-A_{21}\tilde{x}(u)\Big)=0 \qquad \forall u\in(0,1),\end{aligned} \tag{4.13a}$$

$$\frac{d}{du}\tilde{x}(0)=A_{11}\tilde{x}(0)+B_{11}B_{11}^*\Lambda^{-1}(\tilde{x}(0)-x^-), \tag{4.13b}$$

$$\frac{d}{du}\tilde{x}(1)=A_{11}\tilde{x}(1). \tag{4.13c}$$

The standard implementation of the Kalman filter is to calculate the conditional expectation $\hat{X}(u)=\mathbb{E}(X(u)\mid Y(v),0\le v\le u)$ by solving the initial value problem

$$\begin{aligned}\frac{d}{du}S(u)&=A_{11}S(u)+S(u)A_{11}^*-S(u)A_{21}^*(B_{22}B_{22}^*)^{-1}A_{21}S(u)+B_{11}B_{11}^*\\ S(0)&=\Lambda\end{aligned} \tag{4.14}$$

and

$$\begin{aligned}\frac{d}{du}\hat{X}(u)&=(A_{11}-S(u)A_{21}^*(B_{22}B_{22}^*)^{-1}A_{21})\hat{X}+S(u)A_{21}^*(B_{22}B_{22}^*)^{-1}\frac{dY}{du}\\ \hat{X}(0)&=x^-.\end{aligned} \tag{4.15}$$

The Kalman smoother $\tilde{X}$, designed to find $\tilde{X}(u)=\mathbb{E}(X(u)\mid Y(v),0\le v\le 1)$, is then given by the backward sweep

$$\begin{aligned}\frac{d}{du}\tilde{X}(u)&=A_{11}\tilde{X}(u)+B_{11}B_{11}^*S(u)^{-1}(\tilde{X}(u)-\hat{X}(u)) \qquad \forall u\in(0,1)\\ \tilde{X}(1)&=\hat{X}(1).\end{aligned} \tag{4.16}$$

See [Øks98, section 6.3 and exercise 6.6] for a reference. We wish to demonstrate that $\tilde{x}(u)=\tilde{X}(u)$.

Equation (4.16) evaluated for $u=1$ gives equation (4.13c). When evaluating (4.16) at $u=0$ we can use the boundary conditions from (4.14) and (4.15) to get equation (4.13b). Thus it remains to show that $\tilde{X}(u)$ satisfies equation (4.13a). We proceed as follows: equation (4.16) gives

$$\begin{aligned}(\frac{d}{du}&+A_{11}^*)(B_{11}B_{11}^*)^{-1}(\frac{d}{du}-A_{11})\tilde{X}\\ &=(\frac{d}{du}+A_{11}^*)(B_{11}B_{11}^*)^{-1}B_{11}B_{11}^*S^{-1}(\tilde{X}-\hat{X})\\ &=(\frac{d}{du}+A_{11}^*)S^{-1}(\tilde{X}-\hat{X})\end{aligned}$$

and so

$$\begin{aligned}(\frac{d}{du}&+A_{11}^*)(B_{11}B_{11}^*)^{-1}(\frac{d}{du}-A_{11})\tilde{X}\\ &=\Big(A_{11}^*S^{-1}+\frac{d}{du}S^{-1}\Big)(\tilde{X}-\hat{X})+S^{-1}\frac{d}{du}(\tilde{X}-\hat{X}).\end{aligned} \tag{4.17}$$

We have

$$\frac{d}{du}S^{-1} = -S^{-1}\frac{dS}{du}S^{-1}$$

and hence, using equation (4.14), we get

$$\frac{d}{du}S^{-1} = -S^{-1}A_{11} - A_{11}^*S^{-1} + A_{21}^*(B_{22}B_{22}^*)^{-1}A_{21} - S^{-1}B_{11}B_{11}^*S^{-1}. \quad (4.18)$$

Subtracting (4.15) from (4.16) leads to

$$S^{-1}\frac{d}{du}(\tilde{X} - \hat{X}) = S^{-1}A_{11}(\tilde{X} - \hat{X}) + S^{-1}B_{11}B_{11}^*S^{-1}(\tilde{X} - \hat{X}) - A_{21}^*(B_{22}B_{22}^*)^{-1}\Big(\frac{dY}{du} - A_{21}\hat{X}\Big). \quad (4.19)$$

By substituting (4.18), (4.19) into (4.17) and collecting all the terms we find

$$(\frac{d}{du} + A_{11}^*)(B_{11}B_{11}^*)^{-1}(\frac{d}{du} - A_{11})\tilde{X} = -A_{21}^*(B_{22}B_{22}^*)^{-1}\Big(\frac{dY}{du} - A_{21}\tilde{X}\Big)$$

which is equation (4.13a).

We note in passing that equations (4.14) to (4.16) constitute a factorisation of the two-point boundary value problem (4.13) reminiscent of a continuous LU-factorisation of $\mathcal{L}_{11}$.

5 Numerical Approximation of the SPDEs and Sampling

A primary objective when introducing SPDEs in this paper, and in the nonlinear companion [HSV], is to construct MCMC methods to sample conditioned diffusions. In this section we illustrate briefly how this can be implemented.

If we discretise the SDE (1.2) in time by the θ-method, we obtain the following implicitly defined mapping from (x^k, ξ^k) to $x^\star$:

$$\frac{x^\star - x^k}{\Delta t} = \Big(\theta\mathcal{L}x^* + (1-\theta)\mathcal{L}x^k\Big) - \mathcal{L}m + \sqrt{\frac{2}{\Delta t}}\xi^k,$$

where ξ^k is a sequence of i.i.d Gaussian random variables in $\mathcal{H}$ with covariance operator I (i.e. white noise in in $\mathcal{H}$). The Markov chain implied by the map is well-defined on $\mathcal{H}$ for every $\theta \in [\frac{1}{2}, 1]$. This Markov chain can be used as a proposal distribution for an MCMC method, using the Metropolis-Hastings criterion to accept or reject steps. To make a practical algorithm it is necessary to discretise in the Hilbert space $\mathcal{H}$, as well as in time t. This idea extends to nonlinear problems.

Straightforward calculation using the Karhunen-Loève expansion, similar to the calculations following Lemma 2.2, shows that the invariant measure of the SPDE (1.2) is preserved if the SPDE is replaced by

$$\frac{dx}{dt} = -x + m + \sqrt{2\mathcal{C}}\,\frac{dw}{dt}. \quad (5.1)$$

Such pre-conditioning of Langevin equations can be beneficial algorithmically because it equalises convergence rates in different modes. This in turn allows for optimisation of the time-step choice for a Metropolis-Hastings algorithm across all modes simultaneously. We illustrate this issue for the linear Gaussian processes of interest here.

Equation (5.1) can be discretised in time by the θ-method to obtain the following implicitly defined mapping from (x^k, ξ^k) into $x^\star$:

$$\frac{x^\star - x^k}{\Delta t} = -\Big(\theta x^\star + (1-\theta)x^k\Big) + m + \sqrt{\frac{2}{\Delta t}}\xi^k.$$

Now ξ^k is a sequence of i.i.d. Gaussian random variables in $\mathcal{H}$ with covariance operator $\mathcal{C}$. Again, this leads to a well-defined Markov chain on $\mathcal{H}$ for every $\theta \in [\frac{1}{2}, 1]$. Furthermore the invariant measure is $\mathcal{C}/(1 + (\theta - \frac{1}{2})\Delta t)$. Thus the choice $\theta = \frac{1}{2}$ has a particular advantage: it preserves the exact invariant measure, for all $\Delta t > 0$. (These observations can be justified by using the Karhunen-Loève expansion). Note that

$$\Big(1 + \theta\Delta t\Big)x^\star = \Big(1 - (1-\theta)\Delta t\Big)x^k + \sqrt{2\Delta t}\xi_k.$$

When $\theta = \frac{1}{2}$, choosing $\Delta t = 2$ generates independent random variables which therefore sample the invariant measure independently. This illustrates in a simple Gaussian setting the fact that it is possible to choose a globally optimal time-step for the MCMC method. To make a practical algorithm it is necessary to discretise in the Hilbert space $\mathcal{H}$, as well as in time t. The ideas provide useful insight into nonlinear problems.

6 Conclusions

In this text we derived and exploited a method to construct linear SPDEs which have a prescribed Gaussian measure as their stationary distribution. The fundamental relation between the diffusion operator $\mathcal{L}$ in the SPDE and the covariance operator $\mathcal{C}$ of the Gaussian measure is $\mathcal{L} = (-\mathcal{C})^{-1}$ and, using this, we showed that the kernel of the covariance operator (the covariance function) is the Green's functions for $\mathcal{L}$. We illustrated this technique by constructing SPDEs which sample from the distributions of linear SDEs conditioned on several different types of observations.

These abstract Gaussian results were used to produce some interesting results about the structure of the Kalman-Bucy filter/smoother. Connections were also made between discretisations of the resulting SPDEs and MCMC methods for the Gaussian processes of interest.

In the companion article [HSV] we build on the present analysis to extend this technique beyond the linear case. There we consider conditioned SDEs where the drift is a gradient (or more generally a linear function plus a gradient). The resulting SPDEs can be derived from the SPDEs in the present text by the addition of an extra drift term to account for the additional gradient. The stationary distributions of the new nonlinear SPDEs are identified by calculating their Radon-Nikodym derivative with respect to the corresponding stationary distributions of the linear equations as identified in the present article; this is achieved via the Girsanov transformation.

References

[Adl81] R. J. Adler. *The Geometry of Random Fields*. Wiley, 1981.

[BL03] V. Betz and J. Lörinczi. Uniqueness of gibbs measures relative to brownian motion. *Ann. I. H. Poincaré*, 39:877–889, 2003.

[DF91] Yu. L. Dalecky and S. V. Fomin. *Measures and Differential Equations in Infinite-Dimensional Space*. Mathematics and Its Applications. Kluwer Academic Publishers, 1991.

[DPZ92] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1992.

[DPZ96] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite-Dimensional Systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1996.

[HSV] M. Hairer, A. M. Stuart, and J. Voss. Analysis of SPDEs arising in path sampling, part II: The nonlinear case. In preparation.

[IMM+90] I. Iscoe, M. B. Marcus, D. MacDonald, M. Talagrand, and J. Zinn. Continuity of L^2-valued Ornstein-Uhlenbeck processes. *Ann. Probab.*, 18:68–84, 1990.

[KS91] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, second edition, 1991.

[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.

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

[RVE05] M. G. Reznikoff and E. Vanden-Eijnden. Invariant measures of SPDEs and conditioned diffusions. *C. R. Acad. Sci. Paris, Ser. I*, 340:305–308, 2005.

[SVW04] A. M. Stuart, J. Voss, and P. Wiberg. Conditional path sampling of SDEs and the Langevin MCMC method. *Communications in Mathematical Sciences*, 2(4):685–697, 2004.

[Yos95] K. Yosida. *Functional analysis*. Classics in Mathematics. Springer, Berlin, 1995. Reprint of the sixth (1980) edition.