The Stationary Distribution of Discretised SPDEs

Jochen Voss

University of Leeds

8th March 2012, CREST workshop, Beppu, Japan

Outline

MCMC Methods using SPDEs

example 1. The stochastic heat equation

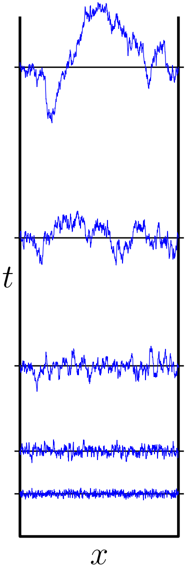
$$\partial_t u(t,x) = \partial_x^2 u(t,x) + \sqrt{2}\,\partial_t w(t,x)$$

with Dirichlet boundary conditions

$$u(t,0) = 0, \quad u(t,1) = 0 \qquad \forall t > 0$$

has the distribution of a Brownian bridge on $[0,1]$ as its stationary distribution.

- $\partial_t w$ is space-time white noise
- $t \in [0,\infty)$ is "time" of the SPDE
- $x \in [0,1]$ ("space" of the SPDE) is "time" of the Brownian bridge.



example 2. Consider the stochastic partial differential equation (SPDE)

$$\partial_t u(t,x) = \partial_x^2 u(t,x) - \left(gg' + \frac{1}{2}g''\right)\left(u(t,x)\right) + \sqrt{2}\partial_t w(t,x)$$

with Dirichlet boundary conditions

$$u(t,0) = 0, \quad u(t,1) = 0 \qquad \forall t > 0.$$

- The stationary distribution of this SPDE on $C\left([0,1],\mathbb{R}\right)$ coincides with the conditional distribution of the process X given by

  $$\begin{aligned} dX_\tau &= g(X_\tau)\,d\tau + dW_\tau \qquad \forall \tau \in [0,1] \\ X_0 &= 0. \end{aligned}$$

  conditioned on $X_1 = 0$.
- We can study X by studying $x \mapsto u(t,x)$ for large times t.

In general, we aim to construct SPDEs such that

- $u(t,\cdot) \in C\big([0,1],\mathbb{R}\big)$ for all $t \geq 0$
- in stationarity, the paths $x \mapsto u(t,x)$ have the distribution of some "interesting" process X, *e.g.* of a conditioned diffusion
- u is ergodic: for $\varphi\colon C\big([0,1],\mathbb{R}\big) \to \mathbb{R}$ we have

$$\mathbb{E}\big(\varphi(X)\big) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi\big(u(t,\cdot)\big)\, dt$$

If we can solve the SPDE on a computer, this leads to Markov Chain Monte Carlo (MCMC) methods: the process u generates samples of X which we can use to study the distribution of X.

How to solve SPDEs on a computer?

We consider SPDEs of the form

$$\partial_t u(t,x) = \partial_x^2 u(t,x) + f\big(u(t,x)\big) + \sqrt{2}\,\partial_t w(t,x)$$

where

- $(t,x) \in [0,\infty) \times [0,1]$,
- $\partial_t w$ is space-time white noise,
- the drift $f : \mathbb{R} \to \mathbb{R}$ is a smooth function,
- the differential operator $\mathcal{L} = \partial_x^2$ is equipped with boundary conditions such that it is a negative operator on the space $L^2\big([0,1],\mathbb{R}\big)$.

Lemma. For $f = 0$, let ν be the stationary distribution of the linear SPDE

$$\partial_t u(t,x) = \partial_x^2 u(t,x) + \sqrt{2}\,\partial_t w(t,x).$$

Then ν coincides with the distribution of the process U given by

$$U(x) = (1-x)L + xR + B(x) \qquad \forall x \in [0,1]$$

where

- B is a Brownian bridge, independent of L and R,
- $L \sim \mathcal{N}(0, \sigma_L^2)$, $R \sim \mathcal{N}(0, \sigma_R^2)$ with $\mathrm{Cov}(L,R) = \sigma_{LR}$,
- σ_L^2, σ_R^2, σ_{LR} are determined by the boundary conditions of $\mathcal{L}$.

Lemma. For $f = F'$ where $F\colon \mathbb{R} \to \mathbb{R}$ is bounded from above, let μ be the stationary distribution of the SPDE

$$\partial_t u(t,x) = \partial_x^2 u(t,x) + f\big(u(t,x)\big) + \sqrt{2}\,\partial_t w(t,x).$$

Then μ satisfies

$$\frac{d\mu}{d\nu}(u) = \frac{1}{Z}\exp\Big(\int_0^1 F\big(u(x)\big)\,dx\Big)$$

where ν is the stationary distribution of the linear SPDE.

On $\mathbb{R}^d$ we know that the SDE

$$dX_t = \nabla \log \varphi(X_t)\,dt + \sqrt{2}\,dW_t$$

has invariant density φ. The lemma is an infinite dimensional analogue of this result.

Finite Element Discretisation

In this talk we only consider space discretisation of our SPDE.

- let $\Delta x = 1/n$, $n \in \mathbb{N}$
- consider x-values on the grid $\{0, \Delta x, \ldots, (n-1)\Delta x, 1\}$
- we use "hat functions" φ_i for $i = 0, 1, \ldots, n$ which have $\varphi_i(i\,\Delta x) = 1$, $\varphi_i(j\,\Delta x) = 0$ for all $j \neq i$, and which are affine between the grid points

Formally, expressing the solution in the basis φ_i as

$$u(t,x) = \sum_j U_j(t)\varphi_j(x)$$

gives

$$\langle \varphi_i, \sum_j \frac{dU_j}{dt}\varphi_j\rangle = \langle \varphi_i, \partial_x^2 \sum_j U_j\varphi_j\rangle + \langle \varphi_i, f\big(\sum_j U_j\varphi_j\big)\rangle + \sqrt{2}\langle \varphi_i, \frac{dw}{dt}\rangle$$

where $\langle \cdot\,, \cdot\,\rangle$ denotes the L^2-inner product. We will see that this is a system of $n+1$ SDEs.

$$\langle \varphi_i, \sum_j \frac{dU_j}{dt}\varphi_j\rangle = \langle \varphi_i, \partial_x^2 \sum_j U_j\varphi_j\rangle + \langle \varphi_i, f\big(\sum_j U_j\varphi_j\big)\rangle + \sqrt{2}\langle \varphi_i, \frac{dw}{dt}\rangle$$

can be written as

$$M\frac{dU}{dt} = L^{\mathrm{FE}}U + f^{\mathrm{FE}}(U) + \sqrt{2}M^{1/2}\frac{dW}{dt}$$

where

- the matrix L^{FE} is defined by $L^{\mathrm{FE}}_{ij} = \langle \varphi_i, \partial_x^2\varphi_j\rangle$,
- the matrix M is defined by $M_{ij} = \langle \varphi_i, \varphi_j\rangle$,
- $f^{\mathrm{FE}}(u)_i = \Big\langle \varphi_i, f\Big(\sum_{j=0}^{n} u_j\varphi_j\Big)\Big\rangle$ for all $u \in \mathbb{R}^{n+1}$, $i = 0, \ldots, n$.
- $\mathrm{Cov}\big(\langle \varphi_i, w_t\rangle, \langle \varphi_j, w_t\rangle\big) = \langle \varphi_i, \varphi_j\rangle t$.

By multiplication with M^{-1} we get the finite element discretisation:

$$\frac{dU}{dt} = M^{-1}L^{\mathrm{FE}}U + M^{-1}f^{\mathrm{FE}}(U) + \sqrt{2}M^{-1/2}\,\frac{dW}{dt}$$

where

- W is an $(n+1)$-dimensional standard Brownian motion
- $L^{\mathrm{FE}} = \dfrac{1}{\Delta x}\begin{pmatrix} -1-\frac{\alpha_1}{\beta_1}\Delta x & 1 & \\ 1 & -2 & 1 \\ & 1 & -1-\frac{\alpha_1}{\beta_1}\Delta x \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$
- $M = \Delta x \begin{pmatrix} 2/6 & 1/6 & \\ 1/6 & 4/6 & 1/6 \\ & 1/6 & 2/6 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$

Lemma. Let $L \in \mathbb{R}^{d\times d}$ be symmetric, negative definite and G be symmetric, positive definite. Then the SDEs

$$\frac{dU}{dt} = L\,U + f(U) + \frac{dW}{dt}$$

and

$$\frac{dU}{dt} = GL\,U + G\,f(U) + G^{1/2}\,\frac{dW}{dt}$$

have the same stationary distribution.

Using the lemma with $G = M^{-1}$ shows that

$$\frac{dU}{dt} = L^{\mathrm{FE}}U + f^{\mathrm{FE}}(U) + \sqrt{2}\,\frac{dW}{dt}$$

has the same stationary distribution as the finite element discretisation.

We first consider the discretised equation for the case $f = 0$:

Lemma. Let $L \in \mathbb{R}^{d \times d}$ be a matrix such that the real part of all eigenvalues is strictly negative. Then the unique stationary distribution of

$$\frac{dU}{dt} = LU + B\,\frac{dW}{dt}$$

is $\mathcal{N}(0, C)$, where the covariance matrix C solves the Lyapunov equation

$$LC + CL^T = -BB^T.$$

Thus, for $f = 0$, the stationary distribution is $\nu_n = \mathcal{N}(0, C^{\mathrm{FE}})$ where C^{FE} is the unique solution of $L^{\mathrm{FE}} C^{\mathrm{FE}} + C^{\mathrm{FE}} L^{\mathrm{FE}} = -2I$, *i.e.* $C^{\mathrm{FE}} = (-L^{\mathrm{FE}})^{-1}$.

The stationary distribution μ_n for the discretised SPDE with $f \neq 0$ can be found using the following lemma:

Lemma. Let $f\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ be a vector field with $f = \nabla F$ for some $F\colon \mathbb{R}^{n+1} \to \mathbb{R}$. Then the SDE

$$dU = LU\,dt + f(U)\,dt + \sqrt{2}\,dW$$

has stationary distribution μ_n with

$$\frac{d\mu_n}{d\nu_n}(u) = \frac{1}{Z_n}\exp\big(F(u)\big)$$

where ν_n is the stationary distribution of the linear equation and Z_n is a normalisation constant.

Once we show that f^{FE} can be written as a gradient, the lemma allows to find μ_n.

Discretisation Error

We have seen how to find

- the stationary distribution μ of the SPDE on $C\big([0,1],\mathbb{R}\big)$
- the stationary distribution μ_n of the discretised SPDE on $\mathbb{R}^{n+1}$

We want to show $\mu_n \to \mu$ as $n \to \infty$.

questions. What metric to use? On which space?

Here we project everything to $\mathbb{R}^{n+1}$: We define

$$\Pi_n \colon C\big([0,1],\mathbb{R}\big) \to \mathbb{R}^{n+1}$$

by

$$\Pi_n u = \big(u(0\Delta x), u(1\Delta x), \ldots, u(n\Delta x)\big).$$

Again, we start with the linear equation.

Lemma. For $f = 0$, let ν be the stationary distribution of the linear SPDE

$$\partial_t u(t,x) = \partial_x^2 u(t,x) + \sqrt{2}\,\partial_t w(t,x).$$

and let ν_n be the stationary distribution of the (linear) finite element discretisation with $f \equiv 0$ on $\mathbb{R}^{n+1}$. Then we have

$$\nu_n = \nu \circ \Pi_n^{-1}$$

for every $n \in \mathbb{N}$.

This shows that for the linear equation there is no discretisation error at all!

Theorem. For $f \neq 0$, let μ be the stationary distribution of the SPDE and let μ_n be the stationary distribution of the finite element discretisation. Assume $f = F'$ where $F \in C^2(\mathbb{R})$ is bounded from above with bounded second derivatives. Then we have

$$\big\|\mu_n - \mu \circ \Pi_n^{-1}\big\|_{\mathrm{TV}} = O\big(\frac{1}{n}\big) \quad \text{as } n \to \infty$$

where $\|\cdot\|_{\mathrm{TV}}$ denotes total-variation distance.

If μ and ν both have densities w.r.t. a common reference measure λ, then the total variation distance can be computed as follows:

$$\|\mu - \nu\|_{\mathrm{TV}} = \int \Big|\frac{d\mu}{d\lambda} - \frac{d\nu}{d\lambda}\Big|\, d\lambda.$$

Main Idea of the Proof

We want to compare

- the stationary distribution μ of the SPDE on $C\big([0,1],\mathbb{R}\big)$
- the stationary distribution μ_n of the discretised SPDE on $\mathbb{R}^{n+1}$

Steps of the proof:

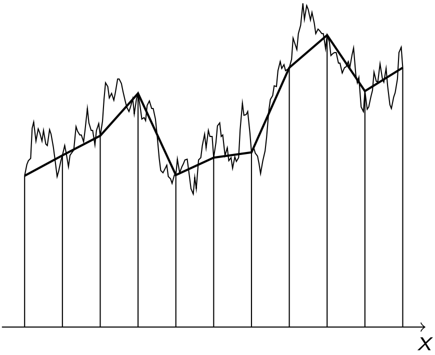
1. find a common space for both measures
2. rewrite the total variation distance using the densities

$$\frac{d\mu}{d\nu} = \frac{1}{Z}\exp\Big(\int_0^1 F\big(U(x)\big)\,dx\Big) \qquad \frac{d\mu_n}{d\nu_n} = \frac{1}{Z_n}\exp\Big(\int_0^1 F\big(U_n(x)\big)\,dx\Big)$$

   where U is distributed according to the stationary distribution ν and $U_n = \sum_{j=0}^{n} U(j\Delta x)\varphi_j(t)$.

3. deal with the normalisation constants
4. compare the two exponentials

Using the above steps, the theorem can be reduced to the question how fast $\|U - U_n\|_\infty$ converges to 0.



The difference $U - U_n$ is a chain of independent Brownian bridges, the resulting questions are easy to answer.

Conclusion

- We have seen that

$$\left\|\mu \circ \Pi_n^{-1} - \mu_n\right\|_{\mathrm{TV}} = O\Big(\frac{1}{n}\Big) \quad \text{as } n \to \infty.$$

  One can show that this bound is sharp.
- Instead of projecting μ onto $\mathbb{R}^{n+1}$ one can embed $\mathbb{R}^{n+1}$ in $C\big([0,1],\mathbb{R}\big)$ by interpolating the discretisation with Brownian bridges. Nearly no changes are required in the proof and the result is the same.
- One would expect for a similar result to hold for SPDEs with values in $\mathbb{R}^d$ instead of in $\mathbb{R}$ (but notation will be more challenging).