Sampling Conditioned Hypoelliptic Diffusions

Jochen Voss

University of Leeds

16th August 2010

Joint work with Martin Hairer and Andrew Stuart

Outline

Sampling on Path Space

The solution of an SDE, *e.g.* of the form

$$dX_t = b(X_t)\,dx + a(X_t)\,dW_t \qquad \forall t \in [0, T],$$

defines a probability distribution μ on the space $C\big([0, T], \mathbb{R}^d\big)$.

Idea. Use a MCMC method, *i.e.* find a stochastic process x with values in $C\big([0, T], \mathbb{R}^d\big)$ whose stationary distribution coincides with the target distribution μ. Assuming ergodicity, we can probe all statistical properties of μ using ergodic averages:

$$\int_{C\left([0,T],\mathbb{R}^d\right)} f(x)\,d\mu(x) = \lim_{S\to\infty} \frac{1}{S}\int_0^S f(x_\tau)\,d\tau.$$

This point of view is particularly useful, if there are additional contraints on the solution X which destroy the basic Markovian structure of the process. Example: sampling bridges with $X(0) = a$ and $X(T) = b$.

basic example: sampling Brownian bridges

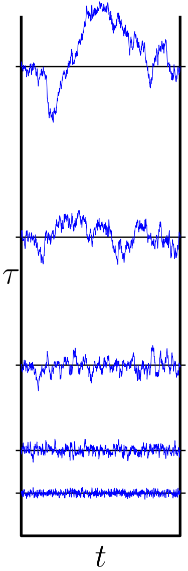
The stochastic heat equation

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

with Dirichlet boundary conditions

$$x(\tau, 0) = 0, \qquad x(\tau, T) = 0$$

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



- $\partial_\tau w$ is space-time white noise
- $t \in [0, T]$ is *physical time* ("space" of the SPDE, time of the Brownian bridge)
- $\tau \in [0, \infty)$ is *algorithmic time* (time of the SPDE)

One can obtain results like the following:

theorem 1. Let X be the solution of

$$dX_t = f(X_t)\,dt + dW_t, \qquad X(0) = 0,\ X(T) = 0.$$

Then the stationary distribution of

$$\partial_\tau x(\tau, t) = \partial_t^2 x(\tau, t) - \big(ff' + \frac{1}{2} f''\big)(x) + \sqrt{2}\,\partial_\tau w(\tau, t)$$

with Dirichlet boundary conditions

$$x(\tau, 0) = 0, \qquad x(\tau, T) = 0$$

coincides with the distribution of X on $C\big([0,1], \mathbb{R}\big)$.

The result needs (among other assumptions) that f is a gradient.

Main Result

We consider hypoelliptic diffusions of the form

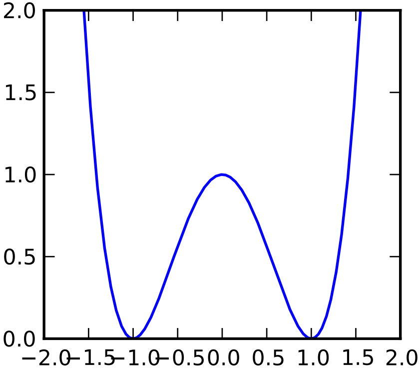
$$m\ddot{X}_t = F(X_t) - \dot{X}_t + \sqrt{2/\beta}\,\dot{W}_t$$

where $X_t \in \mathbb{R}^d$ for $t \in [0, T]$, $F\colon \mathbb{R}^d \to \mathbb{R}^d$, $\beta > 0$ and $\dot{W}$ is white noise. This could, for example, describe a physical system with friction and noise.

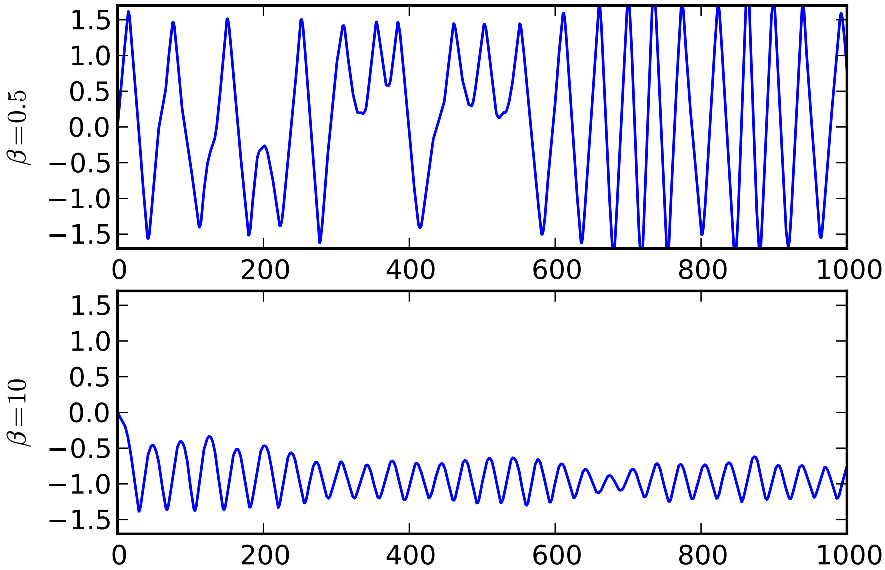
Example. We can consider the case $F = -V'$ where V is a double-well potential:

$$V(x) = (x-1)^2(x+1)^2 \quad \forall x \in \mathbb{R}.$$

Depending on the amount of noise, the system exhibits metastable behaviour.



$$m\ddot{X}_t = F(X_t) - \dot{X}_t + \sqrt{2/\beta}\,\dot{W}_t \qquad X_0 = 0$$



Sometimes we want to simulate the dynamics of the system conditioned on certain events.

Examples.

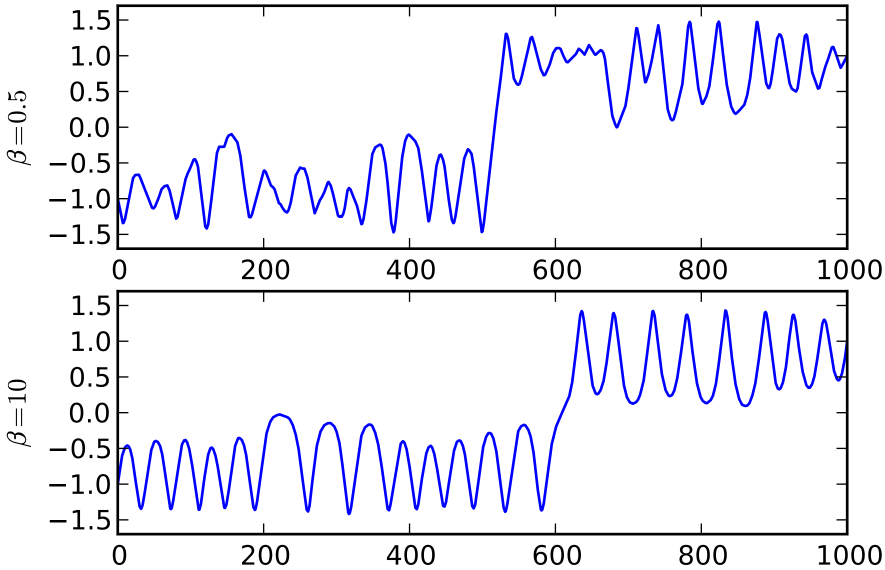
- We can study the transitions between meta-stable states by simulating paths conditioned on a transition happening.
- In signal processing we want to find the conditional distribution of the system given (noisy) observations.

Problem. How can we sample from the distribution μ of

$$m\,\ddot{X}_t = F(X_t) - \dot{X}_t + \sqrt{2/\beta}\,\dot{W}_t,$$

conditioned on $X_0 = x_-$ and $X_T = x_+$?

$$m\ddot{X}_t = F(X_t) - \dot{X}_t + \sqrt{2/\beta}\,\dot{W}_t \qquad X_0 = -1, \quad X_{1000} = +1$$



Main Result

theorem 2. Let $x \colon \Omega \times \mathbb{R}_+ \to C\big([0,T],\mathbb{R}^d\big)$ be the solution of

$$\partial_\tau x(\tau,t) = \mathcal{L}\big(x(\tau,t) - \bar{x}(t)\big) + \mathcal{N}(x) + \sqrt{2}\,\partial_\tau w(\tau,t)$$

where $\mathcal{L} = -\frac{\beta}{2}\big(m^2\partial_t^4 - \partial_t^2\big)$ with certain boundary conditions,

$$\begin{aligned}\mathcal{N}_k(x) = &-\frac{\beta}{2}F_i(x)\partial_k F_i(x) + \frac{m\beta}{2}\partial_t x_i \partial_t x_j \partial_{ij}^2 F_k(x)\\ &-\frac{\beta}{2}\partial_t x_j\big(\partial_j F_k(x) - \partial_k F_j(x)\big)\\ &+\frac{m\beta}{2}\partial_t^2 x_j\big(\partial_j F_k(x) + \partial_k F_j(x)\big)\\ &+\frac{m\beta}{2}\big(F_k(x_-)\partial_t\delta_0 - F_k(x_+)\partial_t\delta_T\big)\end{aligned}$$

and w is a cylindrical Wiener process. Then, in stationarity, the distribution of $t \mapsto x(\tau,t)$ coincides with the target distribution μ.

Remarks about the Proof

As usual, we can rewrite the second order SDE as a system of first order SDEs. Let $q_t = X_t$ and $p_t = m\dot{X}_t$, then

$$\begin{aligned} dq_t &= \frac{1}{m} p_t \, dt, && q_0 = x_- \\ dp_t &= -\frac{1}{m} p_t \, dt + F(q) \, dt + \sqrt{2/\beta} \, dW_t, && p_0 \sim \mathcal{N}(0, \frac{m}{\beta}). \end{aligned}$$

remark. q is a deterministic function of p. Using this function we can solve the second equation to get p. Finally we can compute q from p.

The linear case ($F = 0$)

For $F = 0$, the hypoelliptic SDE simplifies to

$$m\ddot{X}_t = -\dot{X}_t + \sqrt{2/\beta}\,\dot{W}_t.$$

Since this equation is linear, X is a Gaussian process and its distribution is completely characterised by the mean $\bar{x}$ and the covariance operator $\mathcal{C}$.

lemma. Let $\mathcal{L}$ be a linear, negative, self-adjoint operator on $L^2\big([0,T],\mathbb{R}^d\big)$ such that $\mathcal{C} = -\mathcal{L}^{-1}$ is trace class and let $\bar{x} \in L^2\big([0,T],\mathbb{R}^d\big)$. Then

$$\partial_\tau x(\tau,t) = \mathcal{L}(x - \bar{x})\,d\tau + \sqrt{2}\partial_\tau w(\tau,t)$$

has stationary distribution $\mathcal{N}(\bar{x},\mathcal{C})$.

In our situation we get $\mathcal{L} = -\frac{\beta}{2}\big(m^2\partial_t^4 - \partial_t^2\big)$ (with certain boundary conditions).

The non-linear case ($F \neq 0$)

lemma (on $\mathbb{R}^n$). Let μ, ν be probability distributions. Assume that ν is the stationary distribution of

$$dz(\tau) = Lz(\tau)\,d\tau \qquad\qquad\qquad + \sqrt{2}\,dw(\tau).$$

and that $\frac{d\mu}{d\nu} = \varphi$. Then

$$dx(\tau) = Lx(\tau)\,d\tau + \nabla \log \varphi(x(\tau)) + \sqrt{2}\,dw(\tau)$$

has stationary distribution μ.

The result can be carried over to infinite dimensional situations by finite dimensional approximation.

note. Since the equation for z is linear, we know $\nu = \mathcal{N}(0, -L^{-1})$.

In our case:

- ν is the target distribution with $F = 0$,
- μ is the target distribution with $F \neq 0$.

Girsanov's formula gives

$$\varphi(q) = \exp\Big(\sqrt{\frac{\beta}{2}} \int_0^T \langle F(q(t)), dW(t)\rangle - \frac{\beta}{4}\int_0^T |F(q(t))|^2\, dt\Big).$$

The (variational) derivative of φ is given by

$$\begin{aligned} D \log \varphi(q) h = {} & \frac{m\beta}{2}\big(F_k(q_+)h_k'(T) - F_k(q_-)h_k'(0)\big) \\ & - \frac{\beta}{2}\int_0^T \Big(F_i \partial_k F_i - m\dot{q}_i\dot{q}_j\partial_{ij}^2 F_k \\ & \qquad + \dot{q}_j(\partial_j F_k - \partial_k F_j) - m\ddot{q}_j(\partial_j F_k + \partial_k F_j)\Big) h_k(t)\, dt \\ = {} & \langle \mathcal{N}(q), h\rangle. \end{aligned}$$

Remarks.

- Existence of local solution follows from the fact that the non-linearity $\mathcal{N}$ is a Lipschitz function from $H^{3/2+\epsilon}$ to $H^{-3/2-\epsilon}$ (for good enough F). One can get the required a-priori bounds to prove the existence of global solutions. The most "dangerous" term in the non-linearity is
$$\partial_t^2 x_j \big(\partial_j F_k(x) + \partial_k F_j(x)\big).$$
- Differently from the earlier result (for first order SDEs), we do not require the drift F to be a gradient.

Conclusion

- The method provides a generic framework to derive sampling equations, many applications are possible (*e.g.* nonlinear filtering).
- Different from the first-order SDE case, we do not require a gradient structure.
- Interesting problems in the theory of the method, implementation, and applications.

References

- M. Hairer, A.M. Stuart and J. Voss, *Sampling Conditioned Diffusions*. Pages 159–186 in Trends in Stochastic Analysis, Cambridge University Press, vol. 353 of London Mathematical Society Lecture Note Series, 2009.
- M. Hairer, A.M. Stuart and J. Voss, *Sampling Conditioned Hypoelliptic Diffusions*. To appear in the Annals of Applied Probability, 2010.