Continuous-discrete smoothing of diffusions

Frank van der Meulen   Moritz Schauer

Imperial College London 2017-15-12

Delft University of Technology   Mathematical Institute Leiden

Overview

Problem description

Simulating diffusion bridges

The general case: blockwise updating

The general case: single block updating

A novel smoothing algorithm

Numerical illustrations

- Lorenz attractor
- Simple pendulum

Problem description

Problem description

- Suppose X is a multivariate diffusion process in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t,$$

where $b \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^{d \times d'}$ and W a $\mathbb{R}^{d'}$-valued Wiener process.

Problem description

- Suppose X is a multivariate diffusion process in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t,$$

where $b \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^{d \times d'}$ and W a $\mathbb{R}^{d'}$-valued Wiener process.

- Informally, for h small and $Z \sim N_{d'}(0, I)$

$$X_{t+h} = X_t + hb(t, X_t) + \sigma(t, X_t)\sqrt{h}Z,$$

Problem description

- Suppose X is a multivariate diffusion process in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t,$$

  where $b \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^{d \times d'}$ and W a $\mathbb{R}^{d'}$-valued Wiener process.
- Informally, for h small and $Z \sim N_{d'}(0, I)$

$$X_{t+h} = X_t + hb(t, X_t) + \sigma(t, X_t)\sqrt{h}Z,$$

- Let $t_0 < t_1 < \ldots < t_n$ and assume observations

$$V_i = L_i X_{t_i} + \eta_i \qquad i = 0, \ldots, n,$$

  with each L_i an $m_i \times d$ matrix with $m_i \leq d$ and $\{\eta_i\}$ a sequence of IID random variables (independent of X).

Examples for L_i

1. Suppose X is two-dimensional.
    - $L_i = I_2$: observe all components.
    - $L_i = \begin{bmatrix} 1 & 0 \end{bmatrix}$: observe only first component.
    - $L_i = \begin{bmatrix} 1 & 1 \end{bmatrix}$: observe the sum of the two components.

Examples for L_i

1. Suppose X is two-dimensional.
   - $L_i = I_2$: observe all components.
   - $L_i = \begin{bmatrix} 1 & 0 \end{bmatrix}$: observe only first component.
   - $L_i = \begin{bmatrix} 1 & 1 \end{bmatrix}$: observe the sum of the two components.
2. Suppose X is three-dimensional.

$$L_i = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

   Observe the difference between
   - components one and two;
   - components two and three.

Problem description

Let

$$\mathcal{D} = \{V_i,\, i = 0, \ldots, n\}.$$

Continuous-discrete smoothing: reconstruct the path $X := (X_t,\, t \in [0, t_n])$, conditional on $\mathcal{D}$.

Problem description

Let

$$\mathcal{D} = \{V_i,\, i = 0, \ldots, n\}.$$

Continuous-discrete smoothing: reconstruct the path $X := (X_t,\, t \in [0, t_n])$, conditional on $\mathcal{D}$.

Motivation: suppose b and σ depend on parameter θ.

Problem description

Let

$$\mathcal{D} = \{V_i,\, i = 0, \dots, n\}.$$

Continuous-discrete smoothing: reconstruct the path $X := (X_t,\, t \in [0, t_n])$, conditional on $\mathcal{D}$.

Motivation: suppose b and σ depend on parameter θ.

- Equip θ with a prior distribution and compute $\theta \mid \mathcal{D}$.

Problem description

Let

$$\mathcal{D} = \{V_i,\, i = 0, \ldots, n\}.$$

Continuous-discrete smoothing: reconstruct the path $X := (X_t,\, t \in [0, t_n])$, conditional on $\mathcal{D}$.

Motivation: suppose b and σ depend on parameter θ.

- Equip θ with a prior distribution and compute $\theta \mid \mathcal{D}$.
- Main difficulty: transition densities of X are intractable.

Problem description

Let

$$\mathcal{D} = \{V_i,\, i = 0, \ldots, n\}.$$

Continuous-discrete smoothing: reconstruct the path $X := (X_t,\, t \in [0, t_n])$, conditional on $\mathcal{D}$.

Motivation: suppose b and σ depend on parameter θ.

- Equip θ with a prior distribution and compute $\theta \mid \mathcal{D}$.
- Main difficulty: transition densities of X are intractable.

Data-augmentation: Sample from $(\theta, X) \mid \mathcal{D}$ by alternating the steps

1. sample $\theta \mid X$;
2. sample $X \mid (\theta, \mathcal{D})$.

A simplified problem

Problem is easier in case of

- full observations and
- no noise on the observations.

Simulation of independent diffusion bridges.

Simulating diffusion bridges

Simplest example: Brownian bridge

Notation: $X_0 = x_0$, we condition on $X_T = x_T$.

- Consider $dX_t = \sigma dW_t$, with $\sigma \in (0, \infty)$.

Simplest example: Brownian bridge

Notation: $X_0 = x_0$, we condition on $X_T = x_T$.

- Consider $dX_t = \sigma dW_t$, with $\sigma \in (0, \infty)$.

  The conditioned process $X^\star$ satisfies the SDE

$$dX_t^\star = \frac{x_T - X_t^\star}{T - t} dt + \sigma dW_t, \qquad X_0^\star = x_0.$$

Simplest example: Brownian bridge

Notation: $X_0 = x_0$, we condition on $X_T = x_T$.

- Consider $dX_t = \sigma dW_t$, with $\sigma \in (0, \infty)$.

  The conditioned process $X^\star$ satisfies the SDE

$$dX_t^\star = \frac{x_T - X_t^\star}{T - t} dt + \sigma dW_t, \qquad X_0^\star = x_0.$$

- Assume transition densities p of X so that

$$P^{(t,x)}(X_T \in A) = \int_A p(t, x; T, v) dv.$$

Simplest example: Brownian bridge

Notation: $X_0 = x_0$, we condition on $X_T = x_T$.

- Consider $dX_t = \sigma dW_t$, with $\sigma \in (0, \infty)$.

  The conditioned process $X^\star$ satisfies the SDE

$$dX_t^\star = \frac{x_T - X_t^\star}{T - t} dt + \sigma dW_t, \qquad X_0^\star = x_0.$$

- Assume transition densities p of X so that

$$P^{(t,x)}(X_T \in A) = \int_A p(t, x; T, v) dv.$$

- Define $r(t, x) \equiv r(t, x; T, x_T) = \nabla_x \log p(t, x; T, x_T)$. Then

$$dX_t^\star = ar(t, X_t^\star) dt + \sigma dW_t, \qquad X_0^\star = u,$$

  where $a = \sigma\sigma'$.

Simplest example: Brownian bridge

Notation: $X_0 = x_0$, we condition on $X_T = x_T$.

- Consider $dX_t = \sigma dW_t$, with $\sigma \in (0, \infty)$.

  The conditioned process $X^\star$ satisfies the SDE

$$dX_t^\star = \frac{x_T - X_t^\star}{T - t} dt + \sigma dW_t, \qquad X_0^\star = x_0.$$

- Assume transition densities p of X so that

$$P^{(t,x)}(X_T \in A) = \int_A p(t, x; T, v) dv.$$

- Define $r(t, x) \equiv r(t, x; T, x_T) = \nabla_x \log p(t, x; T, x_T)$. Then

$$dX_t^\star = a r(t, X_t^\star) dt + \sigma dW_t, \qquad X_0^\star = u,$$

  where $a = \sigma \sigma'$.

- This holds in much greater generality.

Diffusion bridges

Bridge from $(0, x_0)$ to (T, x_T)

$$\mathrm{d}X_t^\star = b^\star(t, X_t^\star)\,\mathrm{d}t + \sigma(t, X_t^\star)\,\mathrm{d}W_t,$$

with drift

$$b^\star(t,x) = b(t,x) + a(t,x)\underbrace{\nabla_x \log p(t,x;T,x_T)}_{r(t,x)}.$$

Diffusion bridges

Bridge from $(0, x_0)$ to (T, x_T)

$$\mathrm{d}X_t^\star = b^\star(t, X_t^\star)\,\mathrm{d}t + \sigma(t, X_t^\star)\,\mathrm{d}W_t,$$

with drift

$$b^\star(t,x) = b(t,x) + a(t,x)\,\underbrace{\nabla_x \log p(t,x;T,x_T)}_{r(t,x)}.$$

Idea: $b^\star$ is intractable, so approximate it by something tractable.

Proposal (proxy for bridge) from $(0, x_0)$ to (T, x_T)

$$dX_t^\circ = b^\circ(t, X_t^\circ)dt + \sigma(t, X_t^\circ)dW_t, \quad X_0^\circ = x_0$$

Diffusion bridges: related work

- Delyon & Hu (2006): Take

$$b^\circ(t, X_t^\circ) = \lambda b(t, X_t^\circ) + \frac{x_T - X_t^\circ}{T - t}.$$

with $\lambda \in \{0, 1\}$.

Diffusion bridges: related work

- Delyon & Hu (2006): Take

$$b^\circ(t, X_t^\circ) = \lambda b(t, X_t^\circ) + \frac{x_T - X_t^\circ}{T - t}.$$

  with $\lambda \in \{0, 1\}$.

- Durham & Gallant (2002): case $\lambda = 0$ with correction on Euler discretisation.

Diffusion bridges: related work

- Delyon & Hu (2006): Take

$$b^\circ(t, X_t^\circ) = \lambda b(t, X_t^\circ) + \frac{x_T - X_t^\circ}{T - t}.$$

  with $\lambda \in \{0, 1\}$.

- Durham & Gallant (2002): case $\lambda = 0$ with correction on Euler discretisation.
- Whitaker et al.(2017): suppose $x(t)$ solves

$$\mathrm{d}x(t) = b(t, x(t))\,\mathrm{d}t.$$

  Take

$$b^\circ(t, X_t^\circ) = b(t, x(t)) + \frac{x_T - X_t^\circ - \int_t^T b(s, x(s))\,\mathrm{d}s}{T - t}.$$

Diffusion bridges: related work

Other approaches

- Bladt, Finch, Sörensen (2016) coupling methods, requires ergodicity.
- Beskos and Roberts (2006) exact algorithm, requires diffusion to be reducible
- Lindström (2012) refinement of work by Durham & Gallant
- $\cdots$ many others

Guided proposals for diffusion bridges

Bridge from $(0, x_0)$ to (T, x_T)

$$dX_t^\star = b^\star(t, X_t^\star)dt + \sigma(t, X_t^\star)dW_t, \quad X_0^\star = x_0$$

with drift

$$b^\star(t,x) = b(t,x) + a(t,x)\underbrace{\nabla_x \log p(t,x;T,x_T)}_{r(t,x)}.$$

Guided proposals for diffusion bridges

Bridge from $(0, x_0)$ to (T, x_T)

$$dX_t^\star = b^\star(t, X_t^\star)dt + \sigma(t, X_t^\star)dW_t, \quad X_0^\star = x_0$$

with drift

$$b^\star(t,x) = b(t,x) + a(t,x)\underbrace{\nabla_x \log p(t,x;T,x_T)}_{r(t,x)}.$$

Proposal (proxy for bridge) from $(0, x_0)$ to (T, x_T)

$$dX_t^\circ = b^\circ(t, X_t^\circ)dt + \sigma(t, X_t^\circ)dW_t, \quad X_0^\circ = x_0$$

with drift

$$b^\circ(t,x) = b(t,x) + a(t,x)\underbrace{\nabla_x \log \widetilde{p}(t,x;T,x_T)}_{\widetilde{r}(t,x)}.$$

Guided proposals for diffusion bridges

Bridge from $(0, x_0)$ to (T, x_T)

$$dX_t^\star = b^\star(t, X_t^\star)dt + \sigma(t, X_t^\star)dW_t, \quad X_0^\star = x_0$$

with drift

$$b^\star(t,x) = b(t,x) + a(t,x)\underbrace{\nabla_x \log p(t,x;T,x_T)}_{r(t,x)}.$$

Proposal (proxy for bridge) from $(0, x_0)$ to (T, x_T)

$$dX_t^\circ = b^\circ(t, X_t^\circ)dt + \sigma(t, X_t^\circ)dW_t, \quad X_0^\circ = x_0$$

with drift

$$b^\circ(t,x) = b(t,x) + a(t,x)\underbrace{\nabla_x \log \widetilde{p}(t,x;T,x_T)}_{\widetilde{r}(t,x)}.$$

Take $\widetilde{p}$ the transition density of auxiliary process $\widetilde{X}$ with

$$d\widetilde{X}_t = \left(\widetilde{\beta}(t) + \widetilde{B}(t)\widetilde{X}_t\right)dt + \widetilde{\sigma}(t)dW_t.$$

Absolute continuity result for guided proposals

Theorem

Assume the diffusion is uniformly elliptic:

eigenvals of $a(t,x) = \sigma(t,x)\sigma(t,x)'$ *are bounded away from zero.*

Absolute continuity result for guided proposals

Theorem

Assume the diffusion is uniformly elliptic:

eigenvals of $a(t,x) = \sigma(t,x)\sigma(t,x)'$ are bounded away from zero.

If $\widetilde{a}(T) = a(T, x_T)$, then

$$\frac{d\mathbb{P}^\star}{d\mathbb{P}^\circ}(X^\circ) = \frac{\widetilde{p}(0,u;T,x_T)}{p(0,u;T,x_T)}\Psi(X^\circ)$$

where $\Psi(X^\circ) = \exp\left(\int_0^T G(s, X_s^\circ)\,\mathrm{d}s\right)$ is tractable.

Metropolis Hastings step

Independent proposals sampler:

1. Initialise $X = (X_t,\, t \in [0,T])$.
2. Propose $X^\circ = (X_t^\circ,\, t \in [0,T])$.
3. Update X to X° with probability $1 \wedge \Psi(X^\circ)\Psi(X)^{-1}$.
4. Return to step (2).

Autoregressive proposals sampler:

There exists a mapping g such that $X^\circ = g(W)$, with W the driving Wiener process.

Choose persistence parameter $\lambda \in [0,1)$.

1. Initialise a Wiener process $Z = (Z_t,\, t \in [0,T])$. Set $X^\circ = g(Z)$.
2. Let W be Wiener process, independent of Z. Propose
$$Z^\circ = \sqrt{\lambda} Z + \sqrt{1-\lambda} W.$$
3. Update Z to Z° with probability $1 \wedge \Psi(g(Z^\circ))\Psi(g(Z))^{-1}$.
4. Return to step (2).

Metropolis Hastings step

Independent proposals sampler:

1. Initialise $X = (X_t,\, t \in [0, T])$.
2. Propose $X^\circ = (X_t^\circ,\, t \in [0, T])$.
3. Update X to X° with probability $1 \wedge \Psi(X^\circ)\Psi(X)^{-1}$.
4. Return to step (2).

Autoregressive proposals sampler:

There exists a mapping g such that $X^\circ = g(W)$, with W the driving Wiener process.

Choose persistence parameter $\lambda \in [0, 1)$.

1. Initialise a Wiener process $Z = (Z_t,\, t \in [0, T])$. Set $X^\circ = g(Z)$.
2. Let W be Wiener process, independent of Z. Propose
$$Z^\circ = \sqrt{\lambda} Z + \sqrt{1 - \lambda} W.$$
3. Update Z to Z° with probability $1 \wedge \Psi(g(Z^\circ))\Psi(g(Z))^{-1}$.
4. Return to step (2).

Metropolis Hastings step

Independent proposals sampler:

1. Initialise $X = (X_t,\, t \in [0, T])$.
2. Propose $X^\circ = (X_t^\circ,\, t \in [0, T])$.
3. Update X to X° with probability $1 \wedge \Psi(X^\circ)\Psi(X)^{-1}$.
4. Return to step (2).

Autoregressive proposals sampler:

There exists a mapping g such that $X^\circ = g(W)$, with W the driving Wiener process.

Choose persistence parameter $\lambda \in [0, 1)$.

1. Initialise a Wiener process $Z = (Z_t,\, t \in [0, T])$. Set $X^\circ = g(Z)$.
2. Let W be Wiener process, independent of Z. Propose
$$Z^\circ = \sqrt{\lambda} Z + \sqrt{1 - \lambda} W.$$
3. Update Z to Z° with probability $1 \wedge \Psi(g(Z^\circ))\Psi(g(Z))^{-1}$.
4. Return to step (2).

Metropolis Hastings step

Independent proposals sampler:

1. Initialise $X = (X_t,\, t \in [0,T])$.
2. Propose $X^\circ = (X^\circ_t,\, t \in [0,T])$.
3. Update X to X° with probability $1 \wedge \Psi(X^\circ)\Psi(X)^{-1}$.
4. Return to step (2).

Autoregressive proposals sampler:

There exists a mapping g such that $X^\circ = g(W)$, with W the driving Wiener process.

Choose persistence parameter $\lambda \in [0,1)$.

1. Initialise a Wiener process $Z = (Z_t,\, t \in [0,T])$. Set $X^\circ = g(Z)$.
2. Let W be Wiener process, independent of Z. Propose
$$Z^\circ = \sqrt{\lambda} Z + \sqrt{1-\lambda} W.$$
3. Update Z to Z° with probability $1 \wedge \Psi(g(Z^\circ))\Psi(g(Z))^{-1}$.
4. Return to step (2).

Metropolis Hastings step

Independent proposals sampler:

1. Initialise $X = (X_t,\, t \in [0,T])$.
2. Propose $X^\circ = (X^\circ_t,\, t \in [0,T])$.
3. Update X to X° with probability $1 \wedge \Psi(X^\circ)\Psi(X)^{-1}$.
4. Return to step (2).

Autoregressive proposals sampler:

There exists a mapping g such that $X^\circ = g(W)$, with W the driving Wiener process.

Choose persistence parameter $\lambda \in [0,1)$.

1. Initialise a Wiener process $Z = (Z_t,\, t \in [0,T])$. Set $X^\circ = g(Z)$.
2. Let W be Wiener process, independent of Z. Propose
$$Z^\circ = \sqrt{\lambda}Z + \sqrt{1-\lambda}W.$$
3. Update Z to Z° with probability $1 \wedge \Psi(g(Z^\circ))\Psi(g(Z))^{-1}$.
4. Return to step (2).

Metropolis Hastings step

Independent proposals sampler:

1. Initialise $X = (X_t,\, t \in [0,T])$.
2. Propose $X^\circ = (X_t^\circ,\, t \in [0,T])$.
3. Update X to X° with probability $1 \wedge \Psi(X^\circ)\Psi(X)^{-1}$.
4. Return to step (2).

Autoregressive proposals sampler:

There exists a mapping g such that $X^\circ = g(W)$, with W the driving Wiener process.

Choose persistence parameter $\lambda \in [0,1)$.

1. Initialise a Wiener process $Z = (Z_t,\, t \in [0,T])$. Set $X^\circ = g(Z)$.
2. Let W be Wiener process, independent of Z. Propose
$$Z^\circ = \sqrt{\lambda} Z + \sqrt{1-\lambda} W.$$
3. Update Z to Z° with probability $1 \wedge \Psi(g(Z^\circ))\Psi(g(Z))^{-1}$.
4. Return to step (2).

On the choice of $\widetilde{X}$

Auxiliary process used for defining X° is given by

$$\mathrm{d}\widetilde{X}_t = \left(\widetilde{\beta}(t) + \widetilde{B}(t)\widetilde{X}_t\right)\,\mathrm{d}t + \widetilde{\sigma}(t)\,\mathrm{d}W_t.$$

On the choice of $\widetilde{X}$

Auxiliary process used for defining X° is given by

$$\mathrm{d}\widetilde{X}_t = \left(\widetilde{\beta}(t) + \widetilde{B}(t)\widetilde{X}_t\right)\,\mathrm{d}t + \widetilde{\sigma}(t)\,\mathrm{d}W_t.$$

1. Simplest choice is

   $$\widetilde{\beta} \equiv 0 \qquad \widetilde{B} \equiv 0 \qquad \widetilde{\sigma} = \sigma(T, x_T).$$

   In case of strong nonlinearities in the drift b this choice will not do.

On the choice of $\widetilde{X}$

Auxiliary process used for defining X° is given by

$$\mathrm{d}\widetilde{X}_t = \left(\widetilde{\beta}(t) + \widetilde{B}(t)\widetilde{X}_t\right)\,\mathrm{d}t + \widetilde{\sigma}(t)\,\mathrm{d}W_t.$$

1. Simplest choice is
$$\widetilde{\beta} \equiv 0 \qquad \widetilde{B} \equiv 0 \qquad \widetilde{\sigma} = \sigma(T, x_T).$$
In case of strong nonlinearities in the drift b this choice will not do.
2. Suppose $x(t)$ solves
$$\mathrm{d}x(t) = b(t, x(t))\,\mathrm{d}t, \qquad x(0) = x_0.$$
Take $\widetilde{B} \equiv 0$ and $\widetilde{\beta}(t) = b(t, x(t))$.

Proposals for the hypo-elliptic case

Apply Girsanov's theorem to X and X° on $[0, T-\epsilon]$.

Find solution $\gamma(s,x)$ to

$$\sigma\gamma = b - b^\circ.$$

Proposals for the hypo-elliptic case

Apply Girsanov's theorem to X and X° on $[0, T-\epsilon]$.

Find solution $\gamma(s,x)$ to

$$\sigma\gamma = b - b^\circ.$$

- Delyon-Hu

$$\sigma(s,x)\gamma(s,x) = \frac{x_T - x}{T-s}$$

No solution unless σ is invertible. No extension to HE case.

Proposals for the hypo-elliptic case

Apply Girsanov's theorem to X and X° on $[0, T-\epsilon]$.

Find solution $\gamma(s,x)$ to

$$\sigma\gamma = b - b^\circ.$$

- Delyon-Hu

$$\sigma(s,x)\gamma(s,x) = \frac{x_T - x}{T-s}$$

  No solution unless σ is invertible. No extension to HE case.

- Guided proposals

$$\sigma\gamma =$$

Proposals for the hypo-elliptic case

Apply Girsanov's theorem to X and X° on $[0, T-\epsilon]$.

Find solution $\gamma(s,x)$ to

$$\sigma\gamma = b - b^\circ.$$

- Delyon-Hu

$$\sigma(s,x)\gamma(s,x) = \frac{x_T - x}{T-s}$$

  No solution unless σ is invertible. No extension to HE case.

- Guided proposals

$$\sigma\gamma = a\widetilde{r} =$$

Proposals for the hypo-elliptic case

Apply Girsanov's theorem to X and X° on $[0, T-\epsilon]$.

Find solution $\gamma(s,x)$ to

$$\sigma\gamma = b - b^\circ.$$

- Delyon-Hu

$$\sigma(s,x)\gamma(s,x) = \frac{x_T - x}{T - s}$$

  No solution unless σ is invertible. No extension to HE case.

- Guided proposals

$$\sigma\gamma = a\widetilde{r} = \sigma\sigma' \,\widetilde{r}.$$

Proposals for the hypo-elliptic case

Apply Girsanov's theorem to X and X° on $[0, T-\epsilon]$.

Find solution $\gamma(s,x)$ to

$$\sigma\gamma = b - b^\circ.$$

- DELYON-HU

  $$\sigma(s,x)\gamma(s,x) = \frac{x_T - x}{T-s}$$

  No solution unless σ is invertible. No extension to HE case.
- GUIDED PROPOSALS

  $$\sigma\gamma = a\widetilde{r} = \sigma\sigma'\,\widetilde{r}.$$

  So take

  $$\gamma = \sigma'\,\widetilde{r}.$$

  Straightforward extension to HE case.

Work in progress... really need to incorporate $\widetilde{B}$ for this to work.

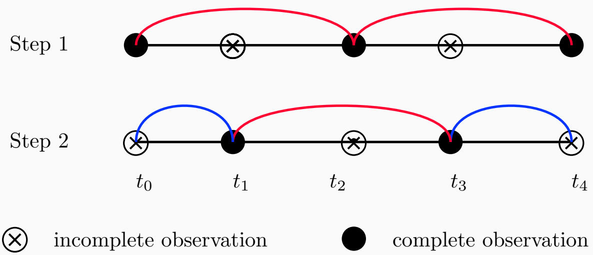
The general case: blockwise updating

Sampling in blocks

- In case of incomplete observations bridges can no longer be sampled independently.
- Dynamics of the bridge depends on *all future conditionings*.

Sampling in blocks

- In case of incomplete observations bridges can no longer be sampled independently.
- Dynamics of the bridge depends on *all future conditionings*.
- Main idea: sample bridges in overlapping blocks.



Related work by GOLIGHTLY & WILKINSON (2008), FUCHS (2013) and DITLEVSEN, JENSEN, PAPASPILIOPOULOS.

Sampling filtered diffusion bridges

Let $0 < S < T$. It suffices to consider one bridge, denoted $X^\star$, connecting x_0 at time 0 to x_T at time T while satisfying $V_S = v_S$.

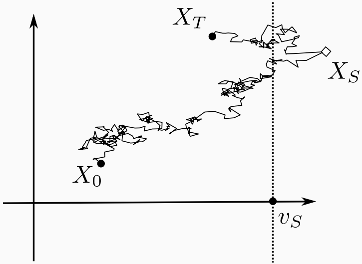


Illustration of filtered diffusion bridge with $L = [1 \;\; 0]$.

SDE for filtered bridge

Theorem

For $t \in [0, T)$, the diffusion conditioned on $V_S = v_S$ and $X_T = x_T$ satisfies the SDE

$$\mathrm{d}X_t^\star = [b(t, X_t^\star) + a(t, X_t^\star) r(t, X_t^\star)]\,\mathrm{d}t + \sigma(t, X_t^\star)\,\mathrm{d}W_t,$$

where

SDE for filtered bridge

Theorem

For $t \in [0,T)$, the diffusion conditioned on $V_S = v_S$ and $X_T = x_T$ satisfies the SDE

$$\mathrm{d}X_t^\star = [b(t, X_t^\star) + a(t, X_t^\star) r(t, X_t^\star)]\,\mathrm{d}t + \sigma(t, X_t^\star)\,\mathrm{d}W_t,$$

where $a = \sigma\sigma'$ and $r(t,x) = \nabla_x \log \rho(t,x)$ with $\rho(t,x)$ equal to

$$\begin{cases} \int p(t,x;S,\xi) p(S,\xi;T,x_T) q(v_S - L\xi)\,\mathrm{d}\xi & t \in [0,S) \\ p(t,x;T,x_T) & t \in [S,T) \end{cases}.$$

SDE for filtered bridge

Theorem

For $t \in [0, T)$, the diffusion conditioned on $V_S = v_S$ and $X_T = x_T$ satisfies the SDE

$$\mathrm{d}X_t^\star = \left[b(t, X_t^\star) + a(t, X_t^\star) r(t, X_t^\star)\right] \mathrm{d}t + \sigma(t, X_t^\star)\, \mathrm{d}W_t,$$

where $a = \sigma\sigma'$ and $r(t, x) = \nabla_x \log \rho(t, x)$ with $\rho(t, x)$ equal to

$$\begin{cases} \int p(t, x; S, \xi) p(S, \xi; T, x_T) q(v_S - L\xi)\, \mathrm{d}\xi & t \in [0, S) \\ p(t, x; T, x_T) & t \in [S, T) \end{cases}.$$

The proof is based on the theory of initial enlargement of filtrations.

Guided proposals for filtered bridges

As p is intractable, we define Metropolis-Hastings proposals.

Define process X° by

$$\mathrm{d}X^\circ_t = [b(t, X^\circ_t) + a(t, X^\circ_t)\widetilde{r}(t, X^\circ_t)]\,\mathrm{d}t + \sigma(t, X^\circ_t)\,\mathrm{d}W_t,$$

with $\widetilde{r}(t,x)$ derived from $\widetilde{p}(\cdot,\cdot;\cdot,\cdot)$ which is the transition density of

$$\mathrm{d}\widetilde{X}_t = \left(\widetilde{\beta}(t) + \widetilde{B}(t)\widetilde{X}_t\right)\,\mathrm{d}t + \widetilde{\sigma}(t)\,\mathrm{d}W_t.$$

Absolute continuity result for guided proposals

Theorem

Assume uniform ellipticity. Let $\mathbb{P}^\star$ and $\mathbb{P}^\circ$ denote the laws of $X^\star$ and X° on $C([0,T])$ respectively. If $\widetilde{a}(T) = a(T, x_T)$, then

$$\frac{\mathrm{d}\mathbb{P}^\star}{\mathrm{d}\mathbb{P}^\circ}(X^\circ) = \frac{\widetilde{\rho}(0, x_0)}{\rho(0, x_0)} \Psi(X^\circ)$$

where

Absolute continuity result for guided proposals

Theorem

Assume uniform ellipticity. Let $\mathbb{P}^\star$ and $\mathbb{P}^\circ$ denote the laws of $X^\star$ and X° on $C([0,T])$ respectively. If $\widetilde{a}(T) = a(T, x_T)$, then

$$\frac{\mathrm{d}\mathbb{P}^\star}{\mathrm{d}\mathbb{P}^\circ}(X^\circ) = \frac{\widetilde{\rho}(0,x_0)}{\rho(0,x_0)}\Psi(X^\circ)$$

where

$$\Psi(X^\circ) = \exp\left(\int_0^T G(s, X_s^\circ)\,\mathrm{d}s\right),$$

with G tractable and depending on b, a, $\widetilde{\beta}$, $\widetilde{B}$ and $\widetilde{a}$.

Related work

Conditioning on one partial noiseless observation ahead, say $v_T = L x_T$.

- J.L. MARCHAND (2011): guiding term

$$a(t, X_t^\circ) L' \left(L a(t, X_t^\circ) L'\right)^{-1} \frac{v_T - L X_t^\circ}{T - t}.$$

Related work

Conditioning on one partial noiseless observation ahead, say $v_T = Lx_T$.

- J.L. Marchand (2011): guiding term

$$a(t, X_t^\circ)L' \left(La(t, X_t^\circ)L'\right)^{-1} \frac{v_T - LX_t^\circ}{T - t}.$$

- Arnaudon, Holm, Sommer (2017), hypo-elliptic case where

$$\sigma = \begin{bmatrix} \Sigma_1 \\ \Sigma_2 \end{bmatrix} \qquad L = \begin{bmatrix} I & 0 \end{bmatrix}$$

Related work

Conditioning on one partial noiseless observation ahead, say $v_T = Lx_T$.

- J.L. Marchand (2011): guiding term

$$a(t, X_t^\circ)L' \left(La(t, X_t^\circ)L'\right)^{-1} \frac{v_T - LX_t^\circ}{T - t}.$$

- Arnaudon, Holm, Sommer (2017), hypo-elliptic case where

$$\sigma = \begin{bmatrix} \Sigma_1 \\ \Sigma_2 \end{bmatrix} \qquad L = \begin{bmatrix} I & 0 \end{bmatrix}$$

Guiding term

$$\sigma(t, X_t^\circ)\Sigma_1(t, X_t^\circ)^+ \frac{v_T - \phi(X_t^\circ)}{T - t},$$

where ideally

$$\phi(X_t^\circ) = E[V_T \mid X_t^\circ].$$

A recursive scheme

It turns out that

$$\widetilde{r}(t,x) = \widetilde{H}(t)(\nu(t) - x).$$

A recursive scheme

It turns out that

$$\widetilde{r}(t,x) = \widetilde{H}(t)(\nu(t) - x).$$

Define $H^\dagger(t) = \widetilde{H}^{-1}(t)$.

We actually only need to keep track of $\nu(t)$ and $H^\dagger(t)$.

A recursive scheme

It turns out that

$$\widetilde{r}(t,x) = \widetilde{H}(t)(\nu(t) - x).$$

Define $H^\dagger(t) = \widetilde{H}^{-1}(t)$.

We actually only need to keep track of $\nu(t)$ and $H^\dagger(t)$.

(I) For $t \in (S,T]$,

$$\begin{aligned} \frac{\mathrm{d}H^\dagger(t)}{\mathrm{d}t} &= \widetilde{B}(t)H^\dagger(t) + H^\dagger(t)\widetilde{B}(t)' - \widetilde{a}(t), \qquad H^\dagger(T) = 0 \\ \frac{\mathrm{d}\nu(t)}{\mathrm{d}t} &= \widetilde{B}(t)\nu(t) + \widetilde{\beta}(t), \qquad \nu(T) = x_T. \end{aligned}$$

A recursive scheme

(II) Compute

$$H^{\dagger}(S) = H^{\dagger}(S+) - H^{\dagger}(S+)L' \left(\Sigma + LH^{\dagger}(S+)L'\right)^{-1} LH^{\dagger}(S+)$$

and

$$\nu(S) = H^{\dagger}(S) \left(L'\Sigma^{-1}v_S + \widetilde{H}(S+)\nu(S+)\right)$$

A recursive scheme

(II) Compute

$$H^\dagger(S) = H^\dagger(S+) - H^\dagger(S+)L'\left(\Sigma + LH^\dagger(S+)L'\right)^{-1} LH^\dagger(S+)$$

and

$$\nu(S) = H^\dagger(S)\left(L'\Sigma^{-1}v_S + \widetilde{H}(S+)\nu(S+)\right)$$

(III) For $t \in [0, S]$,

$$\frac{\mathrm{d}H^\dagger(t)}{\mathrm{d}t} = \widetilde{B}(t)H^\dagger(t) + H^\dagger(t)\widetilde{B}(t)' - \widetilde{a}(t), \qquad H^\dagger(S)$$

$$\frac{\mathrm{d}\nu(t)}{\mathrm{d}t} = \widetilde{B}(t)\nu(t) + \widetilde{\beta}(t), \qquad \nu(S).$$

A recursive scheme

1. Solve **backward** ODEs for $H^\dagger(t)$ and $\nu(t)$.
    - Solve for $t \in (S, T]$.
    - Extract $H^\dagger(S)$ and $\nu(S)$.
    - Solve for $t \in [0, S]$.
2. Simulate **forward** SDE for X°.

A recursive scheme

1. Solve **backward** ODEs for $H^{\dagger}(t)$ and $\nu(t)$.
    - Solve for $t \in (S, T]$.
    - Extract $H^{\dagger}(S)$ and $\nu(S)$.
    - Solve for $t \in [0, S]$.
2. Simulate **forward** SDE for X°.

This algorithm easily extends to conditioning on multiple future observations.

The general case: single block updating

Single block updating

Let $\mathcal{D} = (V_0, \ldots, V_n)$ be the data.

Main idea:
Update $X_0 \mid \mathcal{D}$ and next $X_{(0:n]} := (X_t,\, t \in (0, t_n])$ conditional on $(X_0, \mathcal{D})$.

Single block updating

Let $\mathcal{D} = (V_0, \ldots, V_n)$ be the data.

Main idea:
Update $X_0 \mid \mathcal{D}$ and next $X_{(0:n]} := (X_t,\, t \in (0, t_n])$ conditional on $(X_0, \mathcal{D})$.

Questions:

- What is the guiding term?

Single block updating

Let $\mathcal{D} = (V_0, \ldots, V_n)$ be the data.

Main idea:
Update $X_0 \mid \mathcal{D}$ and next $X_{(0:n]} := (X_t,\, t \in (0, t_n])$ conditional on $(X_0, \mathcal{D})$.

Questions:

- What is the guiding term?
- If we know its form, is it tractable from a computational point of view?

Single block updating

Let $\mathcal{D} = (V_0, \ldots, V_n)$ be the data.

Main idea:
Update $X_0 \mid \mathcal{D}$ and next $X_{(0:n]} := (X_t,\, t \in (0, t_n])$ conditional on $(X_0, \mathcal{D})$.

Questions:

- What is the guiding term?
- If we know its form, is it tractable from a computational point of view?
- Won't this result in extremely small acceptance probabilities?

SDE for the conditioned process

For $t \in [t_{i-1}, t_i)$, the conditioned process satisfies the SDE

$$\mathrm{d}X_t^\star = b(t, X_t^\star)\,\mathrm{d}t + a(t, X_t^\star) r(t, X_t^\star)\,\mathrm{d}t + \sigma(t, X_t^\star)\,\mathrm{d}W_t, \qquad X_{t_{i-1}}^\star = x_{t_{i-1}}.$$

SDE for the conditioned process

For $t \in [t_{i-1}, t_i)$, the conditioned process satisfies the SDE

$$\mathrm{d}X^\star_t = b(t, X^\star_t)\,\mathrm{d}t + a(t, X^\star_t) r(t, X^\star_t)\,\mathrm{d}t + \sigma(t, X^\star_t)\,\mathrm{d}W_t, \qquad X^\star_{t_{i-1}} = x_{t_{i-1}}.$$

Here

- $a(t,x) = \sigma(t,x)\sigma(t,x)'$

SDE for the conditioned process

For $t \in [t_{i-1}, t_i)$, the conditioned process satisfies the SDE

$$\mathrm{d}X_t^\star = b(t, X_t^\star)\,\mathrm{d}t + a(t, X_t^\star) r(t, X_t^\star)\,\mathrm{d}t + \sigma(t, X_t^\star)\,\mathrm{d}W_t, \qquad X_{t_{i-1}}^\star = x_{t_{i-1}}.$$

Here

- $a(t,x) = \sigma(t,x)\sigma(t,x)'$
- $r(t,x) = \nabla_x \log \rho(t,x)$, where ρ is defined by

$$\rho(t,x) = \int p(t,x;t_i,\xi_i) \prod_{j=i}^{n} p(t_j,\xi_j;t_{j+1},\xi_{j+1}) q_j(v_j - L_j\xi_j)\,\mathrm{d}\xi_i \cdots \mathrm{d}\xi_n,$$

- q_j the density of the $N(0, \Sigma_j)$ distribution.

Guided proposals

Replace p with $\widetilde{p}$.

Guided proposals

Replace p with $\widetilde{p}$.

- Formula for Radon-Nikodym derivative remains the same.

Guided proposals

Replace p with $\widetilde{p}$.

- Formula for Radon-Nikodym derivative remains the same.
- Need to evaluate

$$\widetilde{\rho}(t,x) = \int \widetilde{p}(t,x;t_i,\xi_i) \prod_{j=i}^{n} \widetilde{p}(t_j,\xi_j;t_{j+1},\xi_{j+1}) q_j(v_j - L_j\xi_j)\,\mathrm{d}\xi_i \cdots \mathrm{d}\xi_n,$$

but also

$$\widetilde{r}(t,x) = \nabla_x \log \widetilde{\rho}(t,x)$$

and

$$\widetilde{H}(t) = -\nabla_x^2 \log \widetilde{\rho}(t,x)$$

A novel smoothing algorithm

The basic idea

- Compute $\nu(t)$ and $H^\dagger(t)$ recursively on $[0, t_n]$.
- Propose a new starting point, and conditional on that a new driving Wiener process.
- Use Metropolis-Hastings to decide acceptance of the proposal.

The basic idea

- Compute $\nu(t)$ and $H^\dagger(t)$ recursively on $[0, t_n]$.
- Propose a new starting point, and conditional on that a new driving Wiener process.
- Use Metropolis-Hastings to decide acceptance of the proposal.

Choice of auxiliary process $\widetilde{X}$:

- Simply take $\widetilde{B} = 0$ and $\widetilde{\beta} = 0$ (works for uniformly elliptic diffusions).

The basic idea

- Compute $\nu(t)$ and $H^{\dagger}(t)$ recursively on $[0, t_n]$.
- Propose a new starting point, and conditional on that a new driving Wiener process.
- Use Metropolis-Hastings to decide acceptance of the proposal.

Choice of auxiliary process $\widetilde{X}$:

- Simply take $\widetilde{B} = 0$ and $\widetilde{\beta} = 0$ (works for uniformly elliptic diffusions).
- Determine $\widetilde{B}$ and $\widetilde{\beta}$ adaptively.

The basic idea

- Compute $\nu(t)$ and $H^{\dagger}(t)$ recursively on $[0, t_n]$.
- Propose a new starting point, and conditional on that a new driving Wiener process.
- Use Metropolis-Hastings to decide acceptance of the proposal.

Choice of auxiliary process $\widetilde{X}$:

- Simply take $\widetilde{B} = 0$ and $\widetilde{\beta} = 0$ (works for uniformly elliptic diffusions).
- Determine $\widetilde{B}$ and $\widetilde{\beta}$ adaptively.
    1. Initialise auxiliary process.
    2. Run MCMC for k iterations. Compute the average of all simulated paths to obtain $\{\bar{x}(t),\, t \in [0, t_n]\}$.

The basic idea

- Compute $\nu(t)$ and $H^\dagger(t)$ recursively on $[0, t_n]$.
- Propose a new starting point, and conditional on that a new driving Wiener process.
- Use Metropolis-Hastings to decide acceptance of the proposal.

Choice of auxiliary process $\widetilde{X}$:

- Simply take $\widetilde{B} = 0$ and $\widetilde{\beta} = 0$ (works for uniformly elliptic diffusions).
- Determine $\widetilde{B}$ and $\widetilde{\beta}$ adaptively.
  1. Initialise auxiliary process.
  2. Run MCMC for k iterations. Compute the average of all simulated paths to obtain $\{\bar{x}(t),\, t \in [0, t_n]\}$.
  3. Compute $\widetilde{b}(t, x) = b(t, \bar{x}(t)) + J_b(t, \bar{x}(t))(x - \bar{x}(t))$ where J_b denotes the Jacobian matrix of b.

The basic idea

- Compute $\nu(t)$ and $H^\dagger(t)$ recursively on $[0, t_n]$.
- Propose a new starting point, and conditional on that a new driving Wiener process.
- Use Metropolis-Hastings to decide acceptance of the proposal.

Choice of auxiliary process $\widetilde{X}$:

- Simply take $\widetilde{B} = 0$ and $\widetilde{\beta} = 0$ (works for uniformly elliptic diffusions).
- Determine $\widetilde{B}$ and $\widetilde{\beta}$ adaptively.
  1. Initialise auxiliary process.
  2. Run MCMC for k iterations. Compute the average of all simulated paths to obtain $\{\bar{x}(t),\, t \in [0, t_n]\}$.
  3. Compute $\widetilde{b}(t,x) = b(t, \bar{x}(t)) + J_b(t, \bar{x}(t))(x - \bar{x}(t))$ where J_b denotes the Jacobian matrix of b.
  4. Recompute $H^\dagger(t)$ and $\nu(t)$ on $[0, t_n]$ based on $\widetilde{b}^{(i)}(t,x)$.
  5. Return to step 2.

Choose a regularisation parameter $\epsilon \geq 0$ and a persistence parameter $\lambda \in [0, 1)$. Denote the number of MCMC iterations by N.

Choose a regularisation parameter $\epsilon \geq 0$ and a persistence parameter $\lambda \in [0, 1)$. Denote the number of MCMC iterations by N.

1. Initialise

$$H^{\dagger}(t_n) = \left(L_n' \Sigma_n^{-1} L_n + \epsilon I\right)^{-1}$$
$$\nu(t_n) = H^{\dagger}(t_n) L_n' \Sigma_n^{-1} v_n.$$

Choose a regularisation parameter $\epsilon \geq 0$ and a persistence parameter $\lambda \in [0, 1)$. Denote the number of MCMC iterations by N.

1. Initialise

$$H^\dagger(t_n) = \left(L_n' \Sigma_n^{-1} L_n + \epsilon I\right)^{-1}$$
$$\nu(t_n) = H^\dagger(t_n) L_n' \Sigma_n^{-1} v_n.$$

2. For $i = n - 1$ to 0
   2.1 For $t \in (t_i, t_{i+1}]$, backwards solve the ordinary differential equations

$$\frac{\mathrm{d}H^\dagger(t)}{\mathrm{d}t} = \widetilde{B}(t) H^\dagger(t) + H^\dagger(t) \widetilde{B}(t)' - \widetilde{a}(t),$$
$$\frac{\mathrm{d}\nu(t)}{\mathrm{d}t} = \widetilde{B}(t)\nu(t) + \widetilde{\beta}(t).$$

Choose a regularisation parameter $\epsilon \geq 0$ and a persistence parameter $\lambda \in [0,1)$. Denote the number of MCMC iterations by N.

1. Initialise

$$H^\dagger(t_n) = \left(L_n' \Sigma_n^{-1} L_n + \epsilon I\right)^{-1}$$
$$\nu(t_n) = H^\dagger(t_n) L_n' \Sigma_n^{-1} v_n.$$

2. For $i = n-1$ to 0
   2.1 For $t \in (t_i, t_{i+1}]$, backwards solve the ordinary differential equations

$$\frac{\mathrm{d}H^\dagger(t)}{\mathrm{d}t} = \widetilde{B}(t) H^\dagger(t) + H^\dagger(t) \widetilde{B}(t)' - \widetilde{a}(t),$$
$$\frac{\mathrm{d}\nu(t)}{\mathrm{d}t} = \widetilde{B}(t)\nu(t) + \widetilde{\beta}(t).$$

   2.2 Compute

$$H^\dagger(t_i) = H^\dagger(t_i+) - H^\dagger(t_i+) L_i' \left(\Sigma_i + L_i H^\dagger(t_i+) L_i'\right)^{-1} L_i H^\dagger(t_i+),$$
$$\nu(t_i) = H^\dagger(t_i) \left(L_i' \Sigma_i^{-1} v_i + \widetilde{H}(t_i+)\nu(t_i+)\right).$$

3. Sample $X_0 \sim N\left(\nu(0), H^{\dagger}(0)\right)$ and a Wiener process Z on $[0, t_n]$.

3. Sample $X_0 \sim N\left(\nu(0), H^\dagger(0)\right)$ and a Wiener process Z on $[0, t_n]$.

   Simulate the guided proposal $X^\circ = g(X_0, Z)$, i.e.

$$\mathrm{d}X_t^\circ = \left(b(t, X_t^\circ) + a(t, X_t^\circ)\widetilde{H}(t)(\nu(t) - X_t^\circ)\right)\mathrm{d}t + \sigma(t, X_t^\circ)\,\mathrm{d}Z_t.$$

   Initialise X by defining $X = (X_t^\circ,\, t \in [0, t_n])$.

4. Repeat N times
    4.1 Propose a new value for X_0° as follows

$$X_0^\circ = \nu(0) + \sqrt{\lambda}(X_0 - \nu(0)) + \sqrt{1-\lambda}Z,$$

with $Z \sim \mathrm{N}(0, H^\dagger(0))$.

4. Repeat N times

    4.1 Propose a new value for X_0° as follows

$$X_0^\circ = \nu(0) + \sqrt{\lambda}(X_0 - \nu(0)) + \sqrt{1-\lambda} Z,$$

with $Z \sim \mathrm{N}(0, H^\dagger(0))$. Sample independently a Wiener process W and set

$$Z^\circ = \sqrt{\lambda} Z + \sqrt{1-\lambda} W.$$

Compute

$$X^\circ = g(X_0^\circ, Z^\circ).$$

4. Repeat N times

   4.1 Propose a new value for X_0° as follows

   $$X_0^\circ = \nu(0) + \sqrt{\lambda}(X_0 - \nu(0)) + \sqrt{1-\lambda}Z,$$

   with $Z \sim \mathrm{N}(0, H^\dagger(0))$. Sample independently a Wiener process W and set

   $$Z^\circ = \sqrt{\lambda}Z + \sqrt{1-\lambda}W.$$

   Compute

   $$X^\circ = g(X_0^\circ, Z^\circ).$$

   4.2 Compute

   $$A = \Psi(X^\circ)\Psi(X)^{-1}.$$

   Draw $U \sim \mathcal{U}(0,1)$. If $U < A$ then set $X = X^\circ$ and $Z = Z^\circ$.

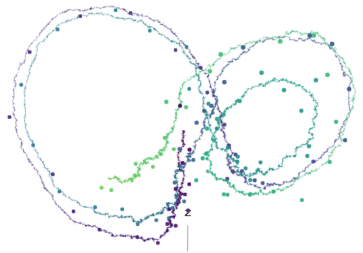
Numerical illustrations

Lorenz attractor

SDE with highly nonlinear dynamics

$$b(x) = \begin{bmatrix} \theta_1(x_2 - x_1) \\ \theta_2 x_1 - x_2 - x_1 x_3 \\ x_1 x_2 - \theta_3 x_3 \end{bmatrix} \qquad \text{and} \qquad \sigma = \sigma_0 I_{3\times 3}.$$

Take $\theta = \begin{bmatrix} 10 & 28 & 8/3 \end{bmatrix}'$, $\sigma_0 = 3$, $\Sigma_i = I$.



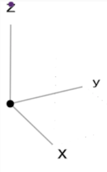


Figure 1: A sample path for $t = [0, 4]$, together with 101 equidistant complete noisy observations. Colours indicate progress of time, with $t = 0$ being violet/dark.

Lorenz attractor

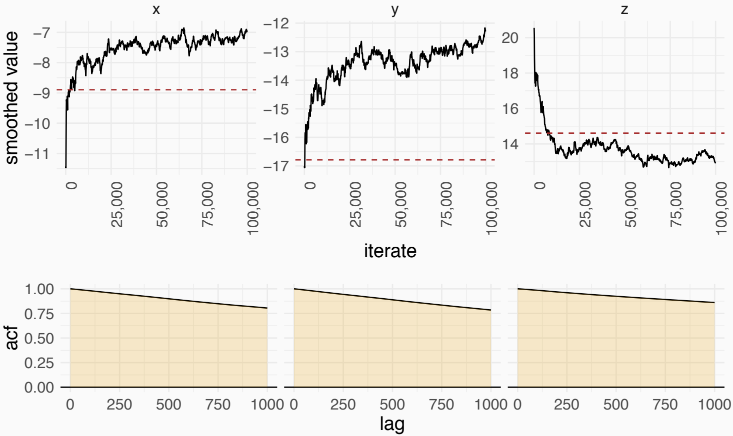
SDE with highly nonlinear dynamics

$$b(x) = \begin{bmatrix} \theta_1(x_2 - x_1) \\ \theta_2 x_1 - x_2 - x_1 x_3 \\ x_1 x_2 - \theta_3 x_3 \end{bmatrix} \qquad \text{and} \qquad \sigma = \sigma_0 I_{3\times 3}.$$

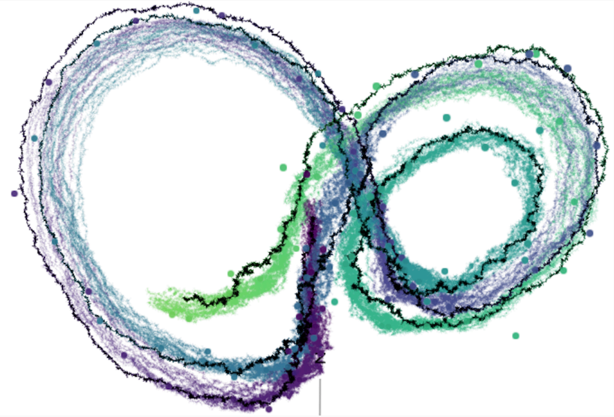
Take $\theta = \begin{bmatrix} 10 & 28 & 8/3 \end{bmatrix}'$, $\sigma_0 = 3$, $\Sigma_i = I$.

Figure 1: A sample path for $t = [0, 4]$, together with 101 equidistant complete noisy observations. Colours indicate progress of time, with $t = 0$ being violet/dark.

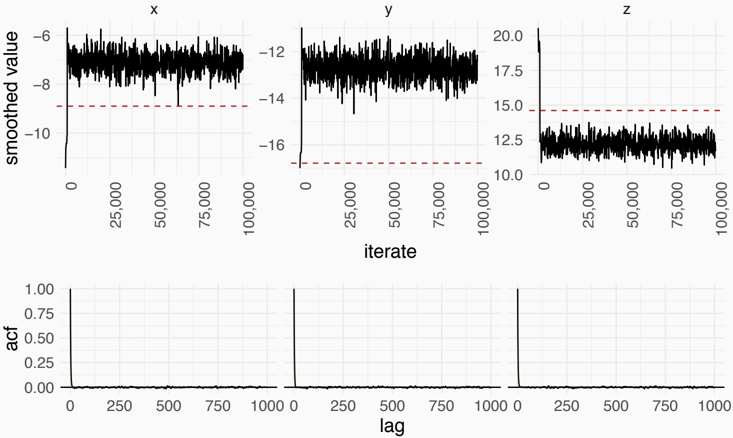
Trace and ACF-plots for X_2: case $\widetilde{B} = 0$, $\widetilde{\beta} = 0$, $\widetilde{\sigma} = \sigma$



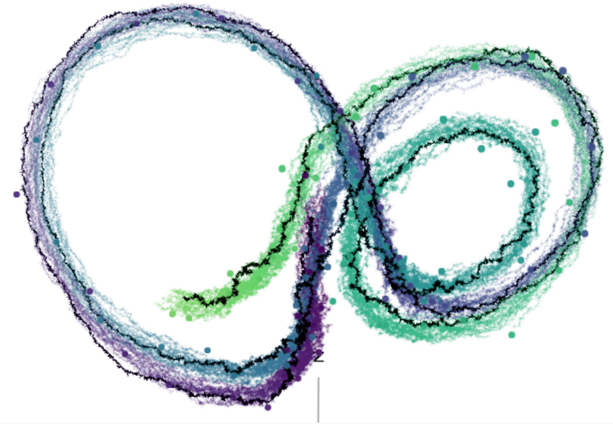
Twenty samples of the posterior: case $\widetilde{B} = 0$, $\widetilde{\beta} = 0$, $\widetilde{\sigma} = \sigma$



Trace and ACF-plots for X_2: adaptive choice auxiliary process



Twenty samples of the posterior: case $\widetilde{B} = 0$, $\widetilde{\beta} = 0$, $\widetilde{\sigma} = \sigma$: adaptive choice auxiliary process



Simple pendulum

Assume $2D$-hypo-elliptic diffusion with SDE

$$\mathrm{d}X_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} X_t \,\mathrm{d}t + \begin{bmatrix} 0 \\ -\theta^2 \sin(X_{1t}) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \gamma \end{bmatrix} \mathrm{d}W_t,$$

where $X_t = \begin{bmatrix} X_{t1} & X_{t2} \end{bmatrix}'$.

Simple pendulum

Assume $2D$-hypo-elliptic diffusion with SDE

$$\mathrm{d}X_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} X_t \,\mathrm{d}t + \begin{bmatrix} 0 \\ -\theta^2 \sin(X_{1t}) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \gamma \end{bmatrix} \mathrm{d}W_t,$$

where $X_t = \begin{bmatrix} X_{t1} & X_{t2} \end{bmatrix}'$.

Assume only the position is observed, i.e. $L_i = \begin{bmatrix} 1 & 0 \end{bmatrix}'$.

Simple pendulum

Assume $2D$-hypo-elliptic diffusion with SDE

$$\mathrm{d}X_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} X_t \,\mathrm{d}t + \begin{bmatrix} 0 \\ -\theta^2 \sin(X_{1t}) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \gamma \end{bmatrix} \mathrm{d}W_t,$$

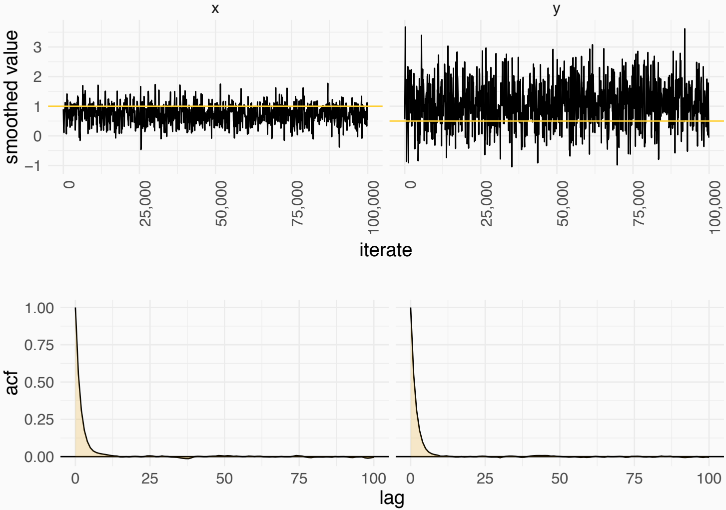
where $X_t = \begin{bmatrix} X_{t1} & X_{t2} \end{bmatrix}'$.

Assume only the position is observed, i.e. $L_i = \begin{bmatrix} 1 & 0 \end{bmatrix}'$.

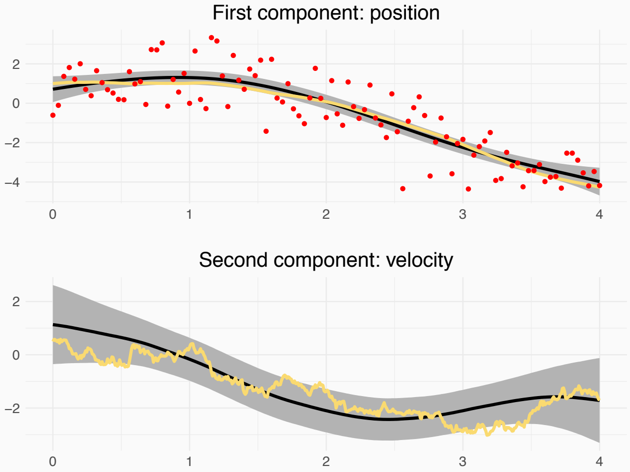
Adaptive MCMC with auxiliary process initialised with

$$\widetilde{B}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \widetilde{\beta}(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \widetilde{\sigma} = \begin{bmatrix} 0 \\ \gamma \end{bmatrix}.$$

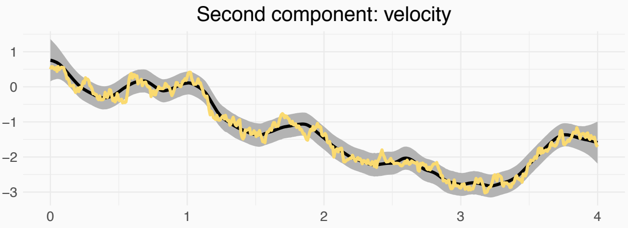
Trace- and ACF plots for X_0: $\sigma^2 = 1$



Posterior for pendulum example: $\sigma^2 = 1$



Posterior for pendulum example: $\sigma^2 = 0.001$



Concluding remarks

Summary:

1. Diffusion bridge simulation: dealing with nonlinearity in the drift.
2. Extension to simulating bridges that take multiple incomplete noisy observations into account.

Concluding remarks

Summary:

1. Diffusion bridge simulation: dealing with nonlinearity in the drift.
2. Extension to simulating bridges that take multiple incomplete noisy observations into account.

Challenges:

1. Improving scalability of higher dimensions.
2. Dealing with nonlinearity in the observation equation.
3. Many more...

Main references

Derivation of the proposal process:

SCHAUER, M. AND VAN DER MEULEN, F. H. AND VAN ZANTEN, J. H. (2017), *Guided proposals for simulating multi-dimensional diffusion bridges*, Bernoulli **23**(4A) (2017), 2917–2950.

Bayesian estimation (full observations without noise):

VAN DER MEULEN, F. H. AND SCHAUER, M. (2017), *Bayesian estimation of discretely observed multi-dimensional diffusion processes using guided proposals*, Electronic Journal of Statistics **11**(1), 2358–2396.

Incomplete observations:

VAN DER MEULEN, F. H. AND SCHAUER, M. (2017) *Bayesian estimation of incompletely observed diffusions*, Stochastics, 1–22.

Single block updating:

VAN DER MEULEN, F. H. AND SCHAUER, M. (2017) *Continuous-discrete smoothing of diffusions*, arXiv:1712.03807.

Advertisement: INFERENCE FOR DIFFUSION MODELS

Parallel Session D CFE-CMStatistics
Saturday 16.12.2017 at 11:25 - 13:05
EO581 *Room CLO 102*

- *Nonparametric learning of stochastic differential equations*
  Andreas Ruttor, TU Berlin, Germany
- *Correlated pseudo marginal schemes for partially observed diffusion processes*
  Andrew Golightly, Newcastle University, United Kingdom
- **Inference for diffusion processes from observations of passage times**
  Moritz Schauer, Leiden University, Netherlands
- *MCMC inference for discretely-observed diffusions: Improving efficiency*
  Christiane Fuchs, Helmholtz Center Munich, Germany