

Skew-symmetric numerical schemes for SDEs and where to find them

G. Vasdekis

Joint work with S. Livingstone, N. Nusken, R. Zhang

UCL

16 February 2024, Warwick

Metropolis-Hastings

Suppose that we want to target a measure with density π .

Metropolis-Hastings

Suppose that we want to target a measure with density π .

First define a transition Kernel q .

Suppose that we want to target a measure with density π .

First define a transition Kernel q .

Assuming the Markov Chain is on state $X_n = x$, choose a proposed next state $y \sim q(x, \cdot)$

Set $X_{n+1} = y$ with probability

$$a(x, y) = \min \left\{ 1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)} \right\}.$$

Metropolis-Hastings

Suppose that we want to target a measure with density π .

First define a transition Kernel q .

Assuming the Markov Chain is on state $X_n = x$, choose a proposed next state $y \sim q(x, \cdot)$

Set $X_{n+1} = y$ with probability

$$a(x, y) = \min \left\{ 1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)} \right\}.$$

Else set $X_{n+1} = x$.

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$.

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$. "Too slow".

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$. "Too slow".
- **MALA:** $q(x, y) = N(x + \frac{1}{2}h\nabla \log \pi(x), h)$.

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$. "Too slow".
- **MALA:** $q(x, y) = N(x + \frac{1}{2}h\nabla \log \pi(x), h)$. "Too fast".

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$. "Too slow".
- **MALA:** $q(x, y) = N(x + \frac{1}{2}h\nabla \log \pi(x), h)$. "Too fast".
- **Barker:** Try $q(x, y) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|)$,

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$. "Too slow".
- **MALA:** $q(x, y) = N(x + \frac{1}{2}h\nabla \log \pi(x), h)$. "Too fast".
- **Barker:** Try $q(x, y) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|)$, $g(t) = \frac{t}{1+t}$.

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$. "Too slow".
- **MALA:** $q(x, y) = N(x + \frac{1}{2}h\nabla \log \pi(x), h)$. "Too fast".
- **Barker:** Try $q(x, y) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|)$, $g(t) = \frac{t}{1+t}$.

Barker is exact $\pi(x)q(x, y) = \pi(y)q(y, x)$.

Motivation

- **Random Walk:** $q(x, y) = \gamma(|x - y|)$. "Too slow".
- **MALA:** $q(x, y) = N(x + \frac{1}{2}h\nabla \log \pi(x), h)$. "Too fast".
- **Barker:** Try $q(x, y) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|)$, $g(t) = \frac{t}{1+t}$.

Barker is exact $\pi(x)q(x, y) = \pi(y)q(y, x)$.

Barker is not probability transition Kernel $\int q(x, y)dy \neq 1$. (Herd et.al., 2022).

Correction

Want to "simulate" from

$$q(x, dy) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|) dy,$$

with $g(t) = \frac{t}{1+t}$.

Correction

Want to "simulate" from

$$q(x, dy) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|) dy,$$

with $g(t) = \frac{t}{1+t}$.

Instead write

$$\frac{\pi(y)}{\pi(x)} = \exp \{ \log(\pi(y)) - \log \pi(x) \} \approx \exp \left\{ (\log \pi(x))' (y - x) \right\}.$$

Correction

Want to "simulate" from

$$q(x, dy) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|) dy,$$

with $g(t) = \frac{t}{1+t}$.

Instead write

$$\frac{\pi(y)}{\pi(x)} = \exp \{ \log(\pi(y)) - \log \pi(x) \} \approx \exp \left\{ (\log \pi(x))' (y - x) \right\}.$$

So instead use the proposal

$$\tilde{q}(x, dy) = 2 \frac{1}{1 + \exp \left\{ - (\log \pi(x))' (y - x) \right\}} \gamma(|x - y|) dy.$$

Correction

Want to "simulate" from

$$q(x, dy) = g\left(\frac{\pi(y)}{\pi(x)}\right) \gamma(|x - y|) dy,$$

with $g(t) = \frac{t}{1+t}$.

Instead write

$$\frac{\pi(y)}{\pi(x)} = \exp \{ \log(\pi(y)) - \log \pi(x) \} \approx \exp \left\{ (\log \pi(x))' (y - x) \right\}.$$

So instead use the proposal

$$\tilde{q}(x, dy) = 2 \frac{1}{1 + \exp \left\{ - (\log \pi(x))' (y - x) \right\}} \gamma(|x - y|) dy.$$

Logistic CDF.

Algorithm (Sampling from \tilde{q})

- ① *Simulate $z \sim \gamma(\cdot)$*

Algorithm (Sampling from \tilde{q})

- ① Simulate $z \sim \gamma(\cdot)$
- ② Sample b such that

$$b = \begin{cases} 1, & \text{with probability } \frac{1}{1 + \exp\{-(\log \pi(x))' z\}} \\ -1 & \text{with probability } 1 - \frac{1}{1 + \exp\{-(\log \pi(x))' z\}}. \end{cases}$$

Algorithm (Sampling from \tilde{q})

- ① Simulate $z \sim \gamma(\cdot)$
- ② Sample b such that

$$b = \begin{cases} 1, & \text{with probability } \frac{1}{1 + \exp\{-(\log \pi(x))' z\}} \\ -1 & \text{with probability } 1 - \frac{1}{1 + \exp\{-(\log \pi(x))' z\}}. \end{cases}$$

- ③ Set $y = x + bz$.

Sampling

Algorithm (Sampling from \tilde{q})

- ① Simulate $z \sim \gamma(\cdot)$
- ② Sample b such that

$$b = \begin{cases} 1, & \text{with probability } \frac{1}{1 + \exp\{-(\log \pi(x))' z\}} \\ -1 & \text{with probability } 1 - \frac{1}{1 + \exp\{-(\log \pi(x))' z\}}. \end{cases}$$

- ③ Set $y = x + bz$.

Then

$$y \sim \tilde{q}(x, \cdot).$$

Numerical Schemes for diffusions ($d = 1$)

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t, Y_0 = x \in \mathbb{R}.$$

Numerical Schemes for diffusions ($d = 1$)

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t, Y_0 = x \in \mathbb{R}.$$

Euler: $Y_{\Delta t} \approx x + \mu(x)\Delta t + \sigma(x)\Delta t^{1/2}\nu, \nu \sim N(0, 1)$

If μ changes rapidly this could be a bad approximation.

Numerical Schemes for diffusions ($d = 1$)

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t, Y_0 = x \in \mathbb{R}.$$

Euler: $Y_{\Delta t} \approx x + \mu(x)\Delta t + \sigma(x)\Delta t^{1/2}\nu, \nu \sim N(0, 1)$

If μ changes rapidly this could be a bad approximation.

Skew-symmetric: $Y_{\Delta t} \approx \begin{cases} x + \sigma(x)\Delta t^{1/2}\nu, & \text{with probability } p \\ x - \sigma(x)\Delta t^{1/2}\nu, & \text{with probability } 1 - p. \end{cases}$

p will take μ into account.

Choice for p

The probability p depends on the **current position** x and the **proposed jump** ξ .

Choice for p

The probability p depends on the **current position** x and the **proposed jump** ξ . For the function $p(x, \xi)$ we must ask that

Assumption

$$p(x, 0) = \frac{1}{2},$$

Choice for p

The probability p depends on the **current position** x and the **proposed jump** ξ . For the function $p(x, \xi)$ we must ask that

Assumption

$$p(x, 0) = \frac{1}{2},$$

and

$$\partial_\xi p(x, 0) = \frac{\mu(x)}{2\sigma(x)^2}.$$

Choice for p

The probability p depends on the **current position** x and the **proposed jump** ξ . For the function $p(x, \xi)$ we must ask that

Assumption

$$p(x, 0) = \frac{1}{2},$$

and

$$\partial_\xi p(x, 0) = \frac{\mu(x)}{2\sigma(x)^2}.$$

For example: Let F be the CDF of a R.V. with PDF f which is symmetric with respect to zero and $f(0) \neq 0$.

Choice for p

The probability p depends on the **current position** x and the **proposed jump** ξ . For the function $p(x, \xi)$ we must ask that

Assumption

$$p(x, 0) = \frac{1}{2},$$

and

$$\partial_\xi p(x, 0) = \frac{\mu(x)}{2\sigma(x)^2}.$$

For example: Let F be the CDF of a R.V. with PDF f which is symmetric with respect to zero and $f(0) \neq 0$. Then

$$p(x, \xi) = F\left(\xi \frac{\mu(x)}{2f(0) \sigma^2(x)}\right)$$

is a valid choice.

Skew-symmetric scheme for diffusions ($d = 1$)

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t.$$

Fix a step size $\Delta t \in (0, 1)$ and a function $p : \mathbb{R} \times \mathbb{R} \rightarrow (0, 1)$.

Skew-symmetric scheme for diffusions ($d = 1$)

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t.$$

Fix a step size $\Delta t \in (0, 1)$ and a function $p : \mathbb{R} \times \mathbb{R} \rightarrow (0, 1)$.

Algorithm (Skew-symmetric)

- ① Suppose $X_t = x \in \mathbb{R}$.
- ② Sample $\nu \sim N(0, 1)$, let $\xi = \Delta t^{1/2} \sigma(x)\nu$.

Skew-symmetric scheme for diffusions ($d = 1$)

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t.$$

Fix a step size $\Delta t \in (0, 1)$ and a function $p : \mathbb{R} \times \mathbb{R} \rightarrow (0, 1)$.

Algorithm (Skew-symmetric)

- ① Suppose $X_t = x \in \mathbb{R}$.
- ② Sample $\nu \sim N(0, 1)$, let $\xi = \Delta t^{1/2}\sigma(x)\nu$.
- ③ Sample b such that

$$b = \begin{cases} 1, & \text{with probability } p(x, \xi) \\ -1 & \text{with probability } 1 - p(x, \xi). \end{cases}$$



Skew-symmetric scheme for diffusions ($d = 1$)

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t.$$

Fix a step size $\Delta t \in (0, 1)$ and a function $p : \mathbb{R} \times \mathbb{R} \rightarrow (0, 1)$.

Algorithm (Skew-symmetric)

- ① Suppose $X_t = x \in \mathbb{R}$.
- ② Sample $\nu \sim N(0, 1)$, let $\xi = \Delta t^{1/2}\sigma(x)\nu$.
- ③ Sample b such that

$$b = \begin{cases} 1, & \text{with probability } p(x, \xi) \\ -1 & \text{with probability } 1 - p(x, \xi). \end{cases}$$

- ④ Set the next step

$$X_{t+1} = X_t + b \xi$$

Then approximate the path $Y_{\Delta t}, Y_{2\Delta t}, \dots$ with X_1, X_2, \dots

Example

Suppose we want to approximate

$$dX_t = \frac{1}{2} (\log \pi(x))' dt + dW_t$$

Example

Suppose we want to approximate

$$dX_t = \frac{1}{2} (\log \pi(x))' dt + dW_t$$

use F and f the CDF/PDF of the logistic R.V. and set

$$p(x, \xi) = F\left(\xi \frac{\mu(x)}{2f(0) \sigma^2(x)}\right) = \left[1 + \exp\{-\xi (\log \pi(x))'\}\right]^{-1}$$

Example

Suppose we want to approximate

$$dX_t = \frac{1}{2} (\log \pi(x))' dt + dW_t$$

use F and f the CDF/PDF of the logistic R.V. and set

$$p(x, \xi) = F\left(\xi \frac{\mu(x)}{2f(0) \sigma^2(x)}\right) = \left[1 + \exp\{-\xi (\log \pi(x))'\}\right]^{-1}$$

Sanity check: $p(x, \xi)$ is large when the proposed jump ξ and $(\log \pi(x))'$ align.

Assumptions for general dimension

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t,$$

where $\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ symmetric matrix. W a d -dimensional Brownian motion.

Assumption

Assumptions for general dimension

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t,$$

where $\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ symmetric matrix. W a d -dimensional Brownian motion.

Assumption

- For all $x \in \mathbb{R}^d$, $\sigma(x)$ is diagonal.

Assumptions for general dimension

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t,$$

where $\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ symmetric matrix. W a d -dimensional Brownian motion.

Assumption

- For all $x \in \mathbb{R}^d$, $\sigma(x)$ is diagonal.
- There exist $M > N > 0$ such that for all $x \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$,

$$N \leq \sigma_{i,i}(x) \leq M.$$



Assumptions for general dimension

Suppose we are targeting the diffusion

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t,$$

where $\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ symmetric matrix. W a d -dimensional Brownian motion.

Assumption

- For all $x \in \mathbb{R}^d$, $\sigma(x)$ is diagonal.
- There exist $M > N > 0$ such that for all $x \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$,

$$N \leq \sigma_{i,i}(x) \leq M.$$

- There exist $C > 0$, $a > 0$, $R > 0$: for all $\|x\| \geq R$

$$\langle \mu(x), x \rangle \leq -C\|x\|^{-a}.$$



Assumption

For all $i \in \{1, \dots, d\}$ let $p_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow (0, 1)$ a function satisfying

$$p_i(x, 0) = \frac{1}{2},$$

Main Assumptions for general dimension (continued...)

Assumption

For all $i \in \{1, \dots, d\}$ let $p_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow (0, 1)$ a function satisfying

$$p_i(x, 0) = \frac{1}{2},$$

and

$$\partial_{\xi_i} p_i(x, 0) = \frac{\mu_i(x)}{2\sigma_{i,i}^2(x)}$$

Skew-symmetric scheme for diffusions (any d).

Fix a step size $\Delta t \in (0, 1)$ and a family of $p_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow (0, 1)$.

Skew-symmetric scheme for diffusions (any d).

Fix a step size $\Delta t \in (0, 1)$ and a family of $p_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow (0, 1)$.

Algorithm (Skew-symmetric)

- ① Suppose $X_t = x \in \mathbb{R}^d$.
- ② Sample $\nu \sim N(0, I_d)$, let $\xi_i = \Delta t^{1/2} \sigma_{i,i}(x) \nu_i$ and $\xi = (\xi_1, \dots, \xi_d)$.

Skew-symmetric scheme for diffusions (any d).

Fix a step size $\Delta t \in (0, 1)$ and a family of $p_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow (0, 1)$.

Algorithm (Skew-symmetric)

- ① Suppose $X_t = x \in \mathbb{R}^d$.
- ② Sample $\nu \sim N(0, I_d)$, let $\xi_i = \Delta t^{1/2} \sigma_{i,i}(x) \nu_i$ and $\xi = (\xi_1, \dots, \xi_d)$.
- ③ For all $i \in \{1, \dots, d\}$, sample b_i such that

$$b_i = \begin{cases} 1, & \text{with probability } p_i(x, \xi) \\ -1 & \text{with probability } 1 - p_i(x, \xi). \end{cases}$$

Skew-symmetric scheme for diffusions (any d).

Fix a step size $\Delta t \in (0, 1)$ and a family of $p_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow (0, 1)$.

Algorithm (Skew-symmetric)

- ① Suppose $X_t = x \in \mathbb{R}^d$.
- ② Sample $\nu \sim N(0, I_d)$, let $\xi_i = \Delta t^{1/2} \sigma_{i,i}(x) \nu_i$ and $\xi = (\xi_1, \dots, \xi_d)$.
- ③ For all $i \in \{1, \dots, d\}$, sample b_i such that

$$b_i = \begin{cases} 1, & \text{with probability } p_i(x, \xi) \\ -1 & \text{with probability } 1 - p_i(x, \xi). \end{cases}$$

- ④ Set the next step $X_{t+1} = (X_{t+1}^1, \dots, X_{t+1}^d)$ such that

$$X_{t+1}^i = X_t^i + b_i \xi_i$$

Then approximate the path $Y_{\Delta t}, Y_{2\Delta t}, \dots$ with X_1, X_2, \dots

Choices for p_i .

Proposition

Let F be the CDF of a R.V. with PDF f which is symmetric with respect to zero and $f(0) \neq 0$. Then

$$p_i(x, \xi) = F\left(\xi_i \frac{\mu_i(x)}{2f(0) \sigma_{i,i}^2(x)}\right)$$

is a valid choice.

Choices for p_i .

Proposition

Let F be the CDF of a R.V. with PDF f which is symmetric with respect to zero and $f(0) \neq 0$. Then

$$p_i(x, \xi) = F\left(\xi_i \frac{\mu_i(x)}{2f(0) \sigma_{i,i}^2(x)}\right)$$

is a valid choice.

Example

If F is the CDF of the logistic function (i.e. $F(y) = [1 + \exp\{-y\}]^{-1}$) then we get

$$p_i(x, \xi) = \frac{1}{1 + \exp\left\{-\xi_i \frac{2\mu_i(x)}{\sigma_{i,i}^2(x)}\right\}}$$

Let's call this the **Unadjusted Barker Algorithm** for diffusions.

Diffusion approximation

Result (Appropriate semigroup expansion)

Under the assumptions above, for all $f \in C_P^6(\mathbb{R}^d)$, for all $\Delta t \in (0, 1)$ and $x \in \mathbb{R}^d$ we have

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

L : generator of the diffusion. \mathcal{A}_2 : some operator

Diffusion approximation

Result (Appropriate semigroup expansion)

Under the assumptions above, for all $f \in C_P^6(\mathbb{R}^d)$, for all $\Delta t \in (0, 1)$ and $x \in \mathbb{R}^d$ we have

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

L : generator of the diffusion. \mathcal{A}_2 : some operator

Note that

$$\mathbb{E}_x[f(Y_{\Delta t})] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \frac{1}{2}L^2 f(x) + O(\Delta t^3).$$

Diffusion approximation

Result (Appropriate semigroup expansion)

Under the assumptions above, for all $f \in C_P^6(\mathbb{R}^d)$, for all $\Delta t \in (0, 1)$ and $x \in \mathbb{R}^d$ we have

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

L : generator of the diffusion. \mathcal{A}_2 : some operator

Note that

$$\mathbb{E}_x[f(Y_{\Delta t})] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \frac{1}{2}L^2 f(x) + O(\Delta t^3).$$

So $X_1 \approx Y_{\Delta t}$ for small Δt .

Diffusion approximation

Result (Appropriate semigroup expansion)

Under the assumptions above, for all $f \in C_P^6(\mathbb{R}^d)$, for all $\Delta t \in (0, 1)$ and $x \in \mathbb{R}^d$ we have

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

L : generator of the diffusion. \mathcal{A}_2 : some operator

Note that

$$\mathbb{E}_x[f(Y_{\Delta t})] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \frac{1}{2}L^2 f(x) + O(\Delta t^3).$$

So $X_1 \approx Y_{\Delta t}$ for small Δt . Better approximation if $\mathcal{A}_2 = \frac{1}{2}L^2$.

Diffusion approximation

Result (Appropriate semigroup expansion)

Under the assumptions above, for all $f \in C_P^6(\mathbb{R}^d)$, for all $\Delta t \in (0, 1)$ and $x \in \mathbb{R}^d$ we have

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

L : generator of the diffusion. \mathcal{A}_2 : some operator

Note that

$$\mathbb{E}_x[f(Y_{\Delta t})] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \frac{1}{2}L^2 f(x) + O(\Delta t^3).$$

So $X_1 \approx Y_{\Delta t}$ for small Δt . Better approximation if $\mathcal{A}_2 = \frac{1}{2}L^2$.
This is essentially impossible.

Weak convergence

Result (Order 1 convergence)

Fix a time horizon $T > 0$. Under appropriate conditions, then there exists $K : \mathbb{R}^d \rightarrow \mathbb{R}_+$ such that for all $\Delta t \in (0, 1)$, for all $k \leq N = \lfloor \frac{T}{\Delta t} \rfloor$ and for all $f \in C_P^4(\mathbb{R}^d)$ we have

$$|\mathbb{E}_x[f(X_k) - f(Y_{k\Delta t})]| \leq K(x)\Delta t,$$

for all $x \in \mathbb{R}^d$.

Extra Assumptions

Assumption

- For $i \in \{1, \dots, d\}$ each $\mu_i(x)$ satisfies

$$\lim_{\substack{\|x\| \rightarrow \infty \\ x_i < -\|x\|_\infty/2}} \mu_i(x) = \infty, \quad \lim_{\substack{\|x\| \rightarrow \infty \\ x_i > \|x\|_\infty/2}} \mu_i(x) = -\infty.$$

- For all $i \in \{1, \dots, d\}$, $p_i(x, \xi) := g_i(x, \xi_i)$, and each g_i is bounded away from 0 on compact sets. In addition $g_i(x, \xi_i) \rightarrow 1$ as $\xi_i \cdot \mu_i(x) \rightarrow \infty$ and $\rightarrow 0$ as $\xi_i \cdot \mu_i(x) \rightarrow -\infty$.

Long-term convergence

The skew-symmetric process X_1, X_2, \dots is not an exact sample from $Y_{\Delta t}, Y_{2\Delta t}, \dots$

Long-term convergence

The skew-symmetric process X_1, X_2, \dots is not an exact sample from $Y_{\Delta t}, Y_{2\Delta t}, \dots$. If $(Y_t)_{t \geq 0}$ has invariant measure π , $(X_n)_{n \geq 0}$ will not converge there.

Long-term convergence

The skew-symmetric process X_1, X_2, \dots is not an exact sample from $Y_{\Delta t}, Y_{2\Delta t}, \dots$. If $(Y_t)_{t \geq 0}$ has invariant measure π , $(X_n)_{n \geq 0}$ will not converge there. However:

The skew-symmetric process X_1, X_2, \dots is not an exact sample from $Y_{\Delta t}, Y_{2\Delta t}, \dots$. If $(Y_t)_{t \geq 0}$ has invariant measure π , $(X_n)_{n \geq 0}$ will not converge there. However:

Result

Under appropriate assumptions, the Markov chain X_1, X_2 has an invariant measure $\pi_{\Delta t}$ with finite moments of all orders.

The skew-symmetric process X_1, X_2, \dots is not an exact sample from $Y_{\Delta t}, Y_{2\Delta t}, \dots$. If $(Y_t)_{t \geq 0}$ has invariant measure π , $(X_n)_{n \geq 0}$ will not converge there. However:

Result

Under appropriate assumptions, the Markov chain X_1, X_2 has an invariant measure $\pi_{\Delta t}$ with finite moments of all orders.

Furthermore, there exists $\lambda > 0$ and $V: \mathbb{R}^d \rightarrow \mathbb{R}_+$ such that for any $n \in \mathbb{N}$ and all $x \in \mathbb{R}^d$

$$\|\mathbb{P}_x(X_n \in \cdot) - \pi_{\Delta t}\|_{TV} \leq V(x) \exp\{-\lambda n\}$$

Assumption

- *The diffusion Y_t admits an invariant measure π and converges there.*
- *The measure π has all its moments finite.*
- *The associated probability density $\pi \in C^4$, $\pi(x) > 0$ for all $x \in \mathbb{R}^d$ and $\log \pi \in C_P^4(\mathbb{R}^d)$.*

Result

Under all previous assumptions, there exists $L > 0$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}$, with $\int g d\pi = 0$, such that for all $\Delta t < 1$, and for all $f \in C_P^\infty(\mathbb{R}^d)$, there exists $|R_{\Delta t, f}| \leq L$ and

$$\int f d\pi_{\Delta t} = \int f d\pi + \Delta t \int f g d\pi + \Delta t^2 R_{\Delta t, f}.$$

Simulations (Ornstein-Uhlenbeck)

$$dY_t = -Y_t dt + \sqrt{2} dW_t.$$

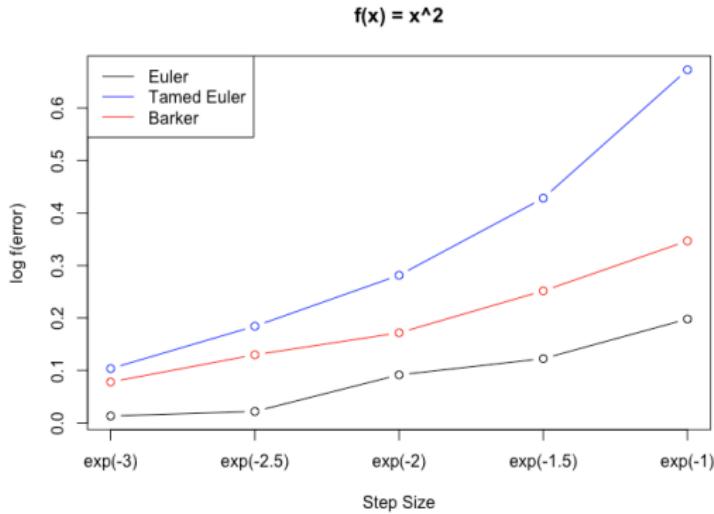


Figure: Absolute errors of simulated Ornstein-Uhlenbeck trajectories

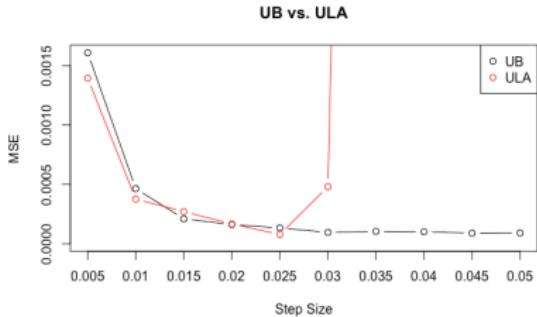
Simulations (Poisson Random Effect)

The Poisson random effects model is of the following form:

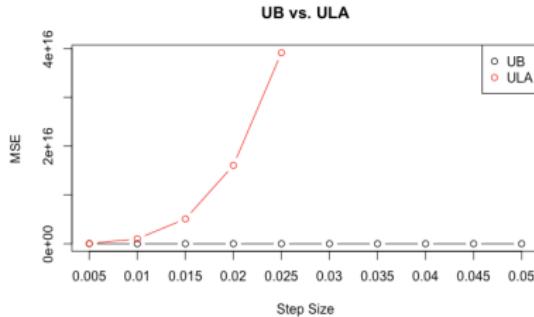
$$\begin{aligned}y_{ij} | \eta_i &\stackrel{\text{indep}}{\sim} \text{Poi}(e^{\eta_i}) \quad j = 1, \dots, 5 \\ \eta_i | \mu &\stackrel{\text{indep}}{\sim} \text{N}(\mu, 1) \quad i = 1, \dots, 50 \\ \mu &\sim \text{N}(0, 100).\end{aligned}$$

Parameter vector $x = (\mu, \eta_1, \dots, \eta_{50})$. The data generate a posterior π on \mathbb{R}^{51} . Approximate the Langevin diffusion
 $dY_t = \nabla \log \pi(Y_t) dt + \sqrt{2} dW_t$

Simulations (Poisson Random Effect)



(a) Equilibrium sampling



(b) Warm Start

Figure: Mean squared error comparisons for the Poisson random effects example.

Simulations (Soft spheres in an anharmonic trap)

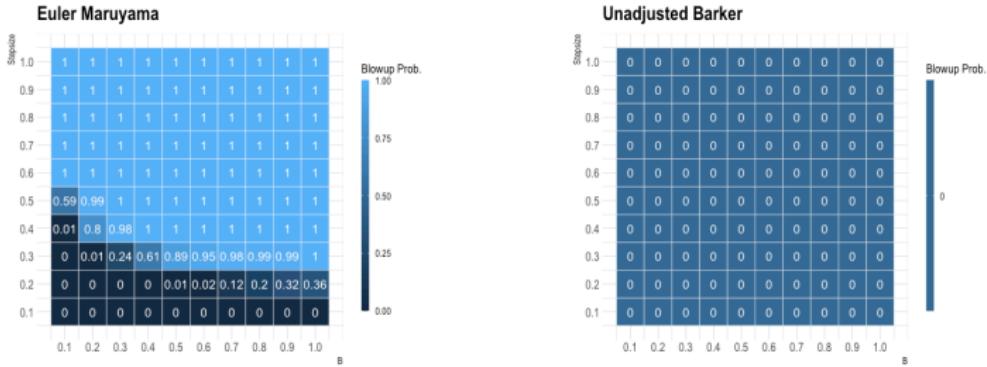
$N = 50$ soft spheres in \mathbb{R}^2 . Sphere i moves along

$$dX_t^{(i)} = -4BX_t^{(i)}\|X_t^{(i)}\|^2 dt + \frac{30}{Nr^2} \sum_{j=1}^N (X_t^{(i)} - X_t^{(j)}) e^{-|X_t^{(i)} - X_t^{(j)}|^2/2r^2} dt \\ + \frac{1}{\sqrt{2}} dW_t$$

for $i = 1, \dots, 50$. Radius of spheres $r = 0.15$.

Want to understand the influence of B and step-size Δt on the algorithm.

Simulations (Soft spheres in an anharmonic trap)



(a) Euler-Maruyama

(b) Unadjusted Barker

Figure: Heat map of numerical explosion probability for soft spheres model.

References

-  Max Hird, Samuel Livingstone, and Giacomo Zanella.
A fresh take on ‘barker dynamics’ for mcmc.
In Alexander Keller, editor, *Monte Carlo and Quasi-Monte Carlo Methods*, pages 169–184, Cham, 2022. Springer International Publishing.
-  Samuel Livingstone and Giacomo Zanella.
The Barker Proposal: Combining Robustness and Efficiency in Gradient-Based MCMC.
Journal of the Royal Statistical Society Series B: Statistical Methodology, 84(2):496–523, 01 2022.
-  G. Vasdekis R. Zhang S. Livingstone, N. Nusken.
On skew-symmetric numerical schemes for stochastic differential equations with non-globally Lipschitz drift: an unadjusted Barker algorithm.
-  Jure Vogrinc, Samuel Livingstone, and Giacomo Zanella.
Optimal design of the Barker proposal and other locally

Thank you for your attention!

"Proof" of diffusion approximation ($d = 1$)

$$\begin{aligned}\mathbb{E}_x[f(X_1) - f(x)] &= \mathbb{E}_\nu \left[\mathbb{E}_b \left[f(x + b\Delta t^{1/2} \sigma(x)\nu) - f(x) \right] \right] \\ &= \mathbb{E}_\nu \left[\left(f(x + \Delta t^{1/2} \sigma(x)\nu) - f(x) \right) p(x, \Delta t^{1/2} \sigma(x)\nu) \right] \\ &\quad + \mathbb{E}_\nu \left[\left(f(x - \Delta t^{1/2} \sigma(x)\nu) - f(x) \right) \left(1 - p(x, \Delta t^{1/2} \sigma(x)\nu) \right) \right]\end{aligned}$$

Operator \mathcal{A}_2

$$\begin{aligned}\mathcal{A}_2 f(x) = & \sum_{i,k=1}^d \partial_i f(x) \partial_i^\xi \partial_k^\xi \partial_k^\xi p_i(x, 0) \|\sigma_{i,i}(x)\|^2 \|\sigma_{k,k}(x)\|^2 \\ & + \sum_{i \neq k} \partial_i \partial_k f(x) \mu_i(x) \mu_k(x) \\ & + 4 \sum_{i \neq k} \partial_i \partial_k f(x) \partial_k^\xi p_i(x, 0) \partial_i^\xi p_k(x, 0) \|\sigma_{i,i}(x)\|^2 \|\sigma_{k,k}(x)\|^2 \\ & + \frac{1}{2} \sum_{i,k=1}^d \partial_i \partial_k^2 f(x) \mu_i(x) \|\sigma_{k,k}(x)\|^2 \\ & + \frac{1}{8} \sum_{i,k=1}^d \partial_i^2 \partial_k^2 f(x) \|\sigma_{i,i}(x)\|^2 \|\sigma_{k,k}(x)\|^2,\end{aligned}$$

Choice of p_i

Recall that

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

Choice of p_i

Recall that

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

The bias of $\int f d\pi_{\Delta t}$ can be made $O(\Delta t^2)$ if we enforce

$$\int \mathcal{A}_2 f d\pi = 0$$

"for many f ".

Choice of p_i

Recall that

$$\mathbb{E}_x[f(X_1)] = f(x) + \Delta t \cdot Lf(x) + \Delta t^2 \cdot \mathcal{A}_2 f(x) + O(\Delta t^3).$$

The bias of $\int f d\pi_{\Delta t}$ can be made $O(\Delta t^2)$ if we enforce

$$\int \mathcal{A}_2 f d\pi = 0$$

"for many f ".

This enforces a condition on the third derivative $\partial_{\xi_i}^3 p_i(x, 0)$.

[HLZ22], [LZ22] , [VLZ22], [SL]