MCMC and non-reversibility

M. Ottobre (Maxwell Institute, Edinburgh)

Joint work with N. Pillai (Harvard), K. Spiliopoulos (Boston)

Durham Symposium, July 2017

Overview

- Markov Chain Monte Carlo (MCMC)

Overview

- Markov Chain Monte Carlo (MCMC)
- Metropolis-Hastings and MALA (Metropolis-Adjusted Langevin Algorithm)

Overview

- Markov Chain Monte Carlo (MCMC)
- Metropolis-Hastings and MALA (Metropolis-Adjusted Langevin Algorithm)
- Reversible vs non-reversible Langevin dynamics

Overview

- Markov Chain Monte Carlo (MCMC)
- Metropolis-Hastings and MALA (Metropolis-Adjusted Langevin Algorithm)
- Reversible vs non-reversible Langevin dynamics
- How to quantify and exploit the advantages of non-reversibility in MCMC

Overview

- Markov Chain Monte Carlo (MCMC)
- Metropolis-Hastings and MALA (Metropolis-Adjusted Langevin Algorithm)
- Reversible vs non-reversible Langevin dynamics
- How to quantify and exploit the advantages of non-reversibility in MCMC
  - Various approaches taken so far

Overview

- Markov Chain Monte Carlo (MCMC)
- Metropolis-Hastings and MALA (Metropolis-Adjusted Langevin Algorithm)
- Reversible vs non-reversible Langevin dynamics
- How to quantify and exploit the advantages of non-reversibility in MCMC
  - Various approaches taken so far
- Non-reversible Hamiltonian Monte Carlo

Overview

- Markov Chain Monte Carlo (MCMC)
- Metropolis-Hastings and MALA (Metropolis-Adjusted Langevin Algorithm)
- Reversible vs non-reversible Langevin dynamics
- How to quantify and exploit the advantages of non-reversibility in MCMC
    - Various approaches taken so far
- Non-reversible Hamiltonian Monte Carlo
- MALA with irreversible proposal (**ipMALA**)

Monte Carlo vs Markov Chain Monte Carlo

- Monte Carlo. Want to compute

$$\mathbb{E}_\pi(f) = \int f(x)d\pi(x)$$

Monte Carlo vs Markov Chain Monte Carlo

- Monte Carlo. Want to compute

$$\mathbb{E}_\pi(f) = \int f(x) d\pi(x)$$

- Use Law of Large numbers: generate i. i. d. samples from π

$$\frac{1}{K}\sum_{j=1}^{K} f(x_j) \stackrel{K\to\infty}{\longrightarrow} \int f(x) d\pi(x), \qquad x_j \sim \pi$$

Monte Carlo vs Markov Chain Monte Carlo

- Monte Carlo. Want to compute

$$\mathbb{E}_\pi(f) = \int f(x) d\pi(x)$$

- Use Law of Large numbers: generate i. i. d. samples from π

$$\frac{1}{K}\sum_{j=1}^{K} f(x_j) \xrightarrow{K\to\infty} \int f(x) d\pi(x), \qquad x_j \sim \pi$$

- MCMC. What if we can't sample directly from π?

MCMC

Step 1. Generate samples from a given target distribution π

MCMC

Step 1. Generate samples from a given target distribution π

- *How?* Construct a Markov Chain x_k that converges to π

MCMC

Step 1. Generate samples from a given target distribution π

- *How?* Construct a Markov Chain x_k that converges to π

Step 2. Calculate integrals of the form

$$\int_{\mathbb{R}^N} f(x) d\pi(x)$$

MCMC

Step 1. Generate samples from a given target distribution π

- *How?* Construct a Markov Chain x_k that converges to π

Step 2. Calculate integrals of the form

$$\int_{\mathbb{R}^N} f(x)d\pi(x)$$

- *How?*: use the Ergodic Theorem

$$\lim_{M\to\infty} \frac{1}{M} \sum_{k=0}^{M} f(x_k) = \int_{\mathbb{R}^N} f(x)d\pi(x)$$

Markov Chain Monte Carlo

- **Goal**: Sample from a given target measure $\pi(x)$ on $\mathbb{R}^N$

Markov Chain Monte Carlo

- **Goal**: Sample from a given target measure $\pi(x)$ on $\mathbb{R}^N$
- **How?** Build a Markov Chain $\{x_k\}_{k\in\mathbb{N}}$ which converges to π

Markov Chain Monte Carlo

- **Goal**: Sample from a given target measure $\pi(x)$ on $\mathbb{R}^N$
- **How?** Build a Markov Chain $\{x_k\}_{k\in\mathbb{N}}$ which converges to π

Metropolis-Hastings Philosophy

Markov Chain Monte Carlo

- **Goal**: Sample from a given target measure $\pi(x)$ on $\mathbb{R}^N$
- **How?** Build a Markov Chain $\{x_k\}_{k\in\mathbb{N}}$ which converges to π

Metropolis-Hastings Philosophy

- Generate a chain $\{x_k\}_{k\in\mathbb{N}}$ satisfying the detailed balance condition with respect to the target measure π

$$\pi(x)p(x,y) = \pi(y)p(y,x) \qquad \textbf{Detailed Balance}$$

Markov Chain Monte Carlo

- **Goal**: Sample from a given target measure $\pi(x)$ on $\mathbb{R}^N$
- **How?** Build a Markov Chain $\{x_k\}_{k\in\mathbb{N}}$ which converges to π

Metropolis-Hastings Philosophy

- Generate a chain $\{x_k\}_{k\in\mathbb{N}}$ satisfying the detailed balance condition with respect to the target measure π

$$\pi(x)p(x,y) = \pi(y)p(y,x) \qquad \text{Detailed Balance}$$

- Detailed Balance (Reversibility) $\Rightarrow$ π is invariant

Markov Chain Monte Carlo

- **Goal**: Sample from a given target measure $\pi(x)$ on $\mathbb{R}^N$
- **How?** Build a Markov Chain $\{x_k\}_{k\in\mathbb{N}}$ which converges to π

Metropolis-Hastings Philosophy

- Generate a chain $\{x_k\}_{k\in\mathbb{N}}$ satisfying the detailed balance condition with respect to the target measure π

$$\pi(x)p(x,y) = \pi(y)p(y,x) \qquad \textbf{Detailed Balance}$$

- Detailed Balance (Reversibility) $\Rightarrow$ π is invariant

Metropolis-Hastings Algorithm

- At step k, chain is in x_k

Metropolis-Hastings Algorithm

- At step k, chain is in x_k

  1. Propose move
  $$y_{k+1} \sim Q(x_k, \cdot)$$

Metropolis-Hastings Algorithm

- At step k, chain is in x_k

  1. Propose move
  $$y_{k+1} \sim Q(x_k, \cdot)$$
  2. Calculate *acceptance probability*
  $$\alpha_k := \alpha(x_k, y_{k+1}) = \min\left\{1, \frac{\pi(y_{k+1})Q(y_{k+1}, x_k)}{\pi(x_k)Q(x_k, y_{k+1})}\right\}$$

Metropolis-Hastings Algorithm

- At step k, chain is in x_k

1. Propose move
$$y_{k+1} \sim Q(x_k, \cdot)$$
2. Calculate *acceptance probability*
$$\alpha_k := \alpha(x_k, y_{k+1}) = \min\left\{1, \frac{\pi(y_{k+1})Q(y_{k+1}, x_k)}{\pi(x_k)Q(x_k, y_{k+1})}\right\}$$
3. Update position
$$x_{k+1} = \begin{cases} y_{k+1} & \text{with probability } \alpha_k \\ x_k & \text{with probability } 1 - \alpha_k \end{cases}$$

Metropolis-Hastings Algorithm

- At step k, chain is in x_k

  1. Propose move

     $$y_{k+1} \sim Q(x_k, \cdot)$$

  2. Calculate *acceptance probability*

     $$\alpha_k := \alpha(x_k, y_{k+1}) = \min\left\{1, \frac{\pi(y_{k+1})Q(y_{k+1}, x_k)}{\pi(x_k)Q(x_k, y_{k+1})}\right\}$$

  3. Update position

     $$x_{k+1} = \begin{cases} y_{k+1} & \text{with probability } \alpha_k \\ x_k & \text{with probability } 1 - \alpha_k \end{cases}$$

- **Whatever the proposal, M-H always creates a reversible chain!**

1953, *Equation of state calculations by fast computing machines*

Figure: Metropolis

Figure: The Tellers

Figure: M. Rosenbluth

1953, *Equation of state calculations by fast computing machines*

Figure: Metropolis

Figure: The Tellers

Figure: M. Rosenbluth

THE purpose of this paper is to describe a general method, suitable for fast electronic computing machines, of calculating the properties of any substance which may be considered as composed of interacting individual molecules.

1953, *Equation of state calculations by fast computing machines*

Figure: Metropolis

Figure: The Tellers

Figure: M. Rosenbluth

THE purpose of this paper is to describe a general method, suitable for fast electronic computing machines, of calculating the properties of any substance which may be considered as composed of interacting individual molecules.

MANIAC = Mathematical Analyzer Numerical Integrator And Calculator

Figure: Ulam

MALA (Metropolis-Adjusted Langevin Algorithm)

- Inspiration: the diffusion process

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t$$

converges to $\pi(x) = e^{-V(x)}$

MALA (Metropolis-Adjusted Langevin Algorithm)

- Inspiration: the diffusion process

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t$$

converges to $\pi(x) = e^{-V(x)}$

- MALA proposal:

$$y = x - \sigma \nabla V(x) + \sqrt{2\sigma}\,\xi, \qquad \xi \sim \mathcal{N}(0, Id_N),\ \sigma = \frac{\ell}{N^\gamma}$$

MALA (Metropolis-Adjusted Langevin Algorithm)

- Inspiration: the diffusion process

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t$$

converges to $\pi(x) = e^{-V(x)}$

- MALA proposal:

$$y = x - \sigma\nabla V(x) + \sqrt{2\sigma}\,\xi, \qquad \xi \sim \mathcal{N}(0, Id_N),\ \sigma = \frac{\ell}{N^\gamma}$$

- MALA algorithm: proposal + accept-reject

MALA (Metropolis-Adjusted Langevin Algorithm)

- Inspiration: the diffusion process

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t$$

converges to $\pi(x) = e^{-V(x)}$

- MALA proposal:

$$y = x - \sigma \nabla V(x) + \sqrt{2\sigma}\,\xi, \qquad \xi \sim \mathcal{N}(0, Id_N),\ \sigma = \frac{\ell}{N^\gamma}$$

- MALA algorithm: proposal + accept-reject

Remark:

MALA (Metropolis-Adjusted Langevin Algorithm)

- Inspiration: the diffusion process

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t$$

converges to $\pi(x) = e^{-V(x)}$

- MALA proposal:

$$y = x - \sigma\nabla V(x) + \sqrt{2\sigma}\,\xi, \qquad \xi \sim \mathcal{N}(0, Id_N),\ \sigma = \frac{\ell}{N^\gamma}$$

- MALA algorithm: proposal + accept-reject

Remark:

- Can think of MALA as a “correct” way of discretizing Langevin dynamics

Non-reversible Langevin

- Langevin (reversible)

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t, \qquad X_t \in \mathbb{R}^N.$$

- X_t is ergodic with invariant measure $\pi(y) = e^{-V(y)}$.

Non-reversible Langevin

- Langevin (reversible)

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t, \qquad X_t \in \mathbb{R}^N .$$

  - X_t is ergodic with invariant measure $\pi(y) = e^{-V(y)}$.

- Non-reversible Langevin

$$dZ_t = -\nabla V(Z_t)dt + \gamma(Z_t)dt + \sqrt{2}dW_t, \qquad \text{with } \nabla \cdot \left(\gamma(z)e^{-V(z)}\right) = 0 .$$

Non-reversible Langevin

- Langevin (reversible)

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t, \qquad X_t \in \mathbb{R}^N.$$

  - X_t is ergodic with invariant measure $\pi(y) = e^{-V(y)}$.

- Non-reversible Langevin

$$dZ_t = -\nabla V(Z_t)dt + \gamma(Z_t)dt + \sqrt{2}dW_t, \qquad \text{with } \nabla \cdot \left(\gamma(z)e^{-V(z)}\right) = 0.$$

  - Invariant measure is still the same

Second Order Langevin

- For $(q(t), p(t)) \in \mathbb{R}^2$

$$
\begin{aligned}
dq &= p\,dt \\
dp &= -\partial_q V(q)dt - p\,dt + \sqrt{2}\,dW_t
\end{aligned}
$$

Second Order Langevin

- For $(q(t), p(t)) \in \mathbb{R}^2$

$$
\begin{aligned}
dq &= p\,dt \\
dp &= -\partial_q V(q) dt - p\,dt + \sqrt{2}\,dW_t
\end{aligned}
$$

- Admits $\mu(q,p) = e^{-p^2/2} e^{-V(q)} = \mathcal{N}(0,1) \times \pi(q)$ as unique invariant measure

Second Order Langevin

- For $(q(t), p(t)) \in \mathbb{R}^2$

$$dq = p\,dt$$
$$dp = -\partial_q V(q)dt - p\,dt + \sqrt{2}\,dW_t$$

- Admits $\mu(q,p) = e^{-p^2/2}e^{-V(q)} = \mathcal{N}(0,1) \times \pi(q)$ as unique invariant measure
- It is *ergodic, irreversible, hypoelliptic* and *hypocoercive*.

Second Order Langevin

- For $(q(t), p(t)) \in \mathbb{R}^2$

$$dq = p\,dt$$
$$dp = -\partial_q V(q) dt - p\,dt + \sqrt{2}\,dW_t$$

- Admits $\mu(q, p) = e^{-p^2/2} e^{-V(q)} = \mathcal{N}(0,1) \times \pi(q)$ as unique invariant measure
- It is *ergodic, irreversible, hypoelliptic* and *hypocoercive*.
- Decomposition of the dynamics in L^2_μ

$$\mathcal{L} = B - A^* A$$

$$\swarrow \qquad \qquad \searrow$$

antisymmetric — *symmetric*

conservative (deterministic) — *dissipative (stochastic)*

Reversible vs Non - Reversible

Advantages

Reversible vs Non - Reversible

Advantages

- Faster convergence to Equilibrium (Huang, Sheu, Pavliotis, Nier, Lelievre, Duncan, O., Stoltz, Olla, Iacobucci)

Reversible vs Non - Reversible

Advantages

- Faster convergence to Equilibrium (Huang, Sheu, Pavliotis, Nier, Lelievre, Duncan, O., Stoltz, Olla, Iacobucci)
- Reduction of asymptotic variance (some of above + Rey-Bellet & Spiliopoulos)

Reversible vs Non - Reversible

Advantages

- Faster convergence to Equilibrium (Huang, Sheu, Pavliotis, Nier, Lelievre, Duncan, O., Stoltz, Olla, Iacobucci)
- Reduction of asymptotic variance (some of above + Rey-Bellet & Spiliopoulos)

Problems

Reversible vs Non - Reversible

Advantages

- Faster convergence to Equilibrium (Huang, Sheu, Pavliotis, Nier, Lelievre, Duncan, O., Stoltz, Olla, Iacobucci)
- Reduction of asymptotic variance (some of above + Rey-Bellet & Spiliopoulos)

Problems

- Discretization

Reversible vs Non - Reversible

Advantages

- Faster convergence to Equilibrium (Huang, Sheu, Pavliotis, Nier, Lelievre, Duncan, O., Stoltz, Olla, Iacobucci)
- Reduction of asymptotic variance (some of above + Rey-Bellet & Spiliopoulos)

Problems

- Discretization
  1. Keep invariant measure
  2. Preserve non-reversibility

Reversible vs Non - Reversible

Advantages

- Faster convergence to Equilibrium (Huang, Sheu, Pavliotis, Nier, Lelievre, Duncan, O., Stoltz, Olla, Iacobucci)
- Reduction of asymptotic variance (some of above + Rey-Bellet & Spiliopoulos)

Problems

- Discretization
  1. Keep invariant measure
  2. Preserve non-reversibility
- Non-reversible processes are, in general, harder to study

...various sources of complication

...various sources of complication

1. Convergence criteria based on Lyapunov functions don't do a good job

...various sources of complication

1. Convergence criteria based on Lyapunov functions don't do a good job
2. • Reversible: geometric ergodicity $\Leftrightarrow L^2_\pi -$ spectral gap

...various sources of complication

1. Convergence criteria based on Lyapunov functions don't do a good job
2. • Reversible: geometric ergodicity $\Leftrightarrow L^2_\pi -$ spectral gap
   • In general : geometric ergodicity $\Leftrightarrow L^\infty_V -$ spectral gap (Kontoyannis & Meyn)

...various sources of complication

1. Convergence criteria based on Lyapunov functions don't do a good job
2. • Reversible: geometric ergodicity $\Leftrightarrow L^2_\pi -$ spectral gap
   • In general : geometric ergodicity $\Leftrightarrow L^\infty_V -$ spectral gap (Kontoyannis & Meyn)
3. Example

...various sources of complication

1. Convergence criteria based on Lyapunov functions don't do a good job
2. • Reversible: geometric ergodicity $\Leftrightarrow L^2_\pi -$ spectral gap
   • In general : geometric ergodicity $\Leftrightarrow L^\infty_V -$ spectral gap (Kontoyannis & Meyn)
3. Example

$$dX_t = \delta dt + dW_t \quad \text{on } S^1$$
$$\mathcal{L}_\delta = \Delta + \delta \nabla$$

...various sources of complication

1. Convergence criteria based on Lyapunov functions don't do a good job
2. • Reversible: geometric ergodicity $\Leftrightarrow L^2_\pi -$ spectral gap
   • In general : geometric ergodicity $\Leftrightarrow L^\infty_V -$ spectral gap (Kontoyannis & Meyn)
3. Example

$$dX_t = \delta dt + dW_t \quad \text{on } S^1$$
$$\mathcal{L}_\delta = \Delta + \delta \nabla$$

Eigenvalues $\rightsquigarrow \quad \lambda_n = -n^2 + in\delta$

Asymptotic variance $\rightsquigarrow \quad \sigma^2(\delta) = \int_0^\infty \langle e^{t\mathcal{L}} f, f \rangle_{L^2} dt = \sum_{n=1}^\infty \frac{2\,|c_n|^2}{n^2 + \delta^2}$

Approaches taken so far

- Produce non- reversible algorithm (abandon M-H framework)

  1. Discretize non-reversible dynamics in a way that the discretization is still reversible - **Non-reversible Hamiltonian Monte Carlo** (Horowitz, Stuart, Pinski, O., Pillai)
  2. Piecewise linear algorithms, Bouncy Particle and Zig-Zag (Bierkens, Roberts, Vollmer, Doucet, Monmarche)
  3. Event chain algorithm (W. Krauth et al, related to work of Diaconis)
  4. General irreversible samplers (Chen et al, Poncet)

Approaches taken so far

- Produce non- reversible algorithm (abandon M-H framework)

  1. Discretize non-reversible dynamics in a way that the discretization is still reversible - **Non-reversible Hamiltonian Monte Carlo** (Horowitz, Stuart, Pinski, O., Pillai)
  2. Piecewise linear algorithms, Bouncy Particle and Zig-Zag (Bierkens, Roberts, Vollmer, Doucet, Monmarche)
  3. Event chain algorithm (W. Krauth et al, related to work of Diaconis)
  4. General irreversible samplers (Chen et al, Poncet)

- Observe that bias is much smaller compared to gain in speed of convergence - "just" simulate (Pavliotis, Duncan, Spiliopoulos, Zygalakis)
  - Design appropriate splitting skemes

(above list not exhaustive)

ipMALA

Question: is it good to use non-reversible proposals within Metropolis-Hastings?

ipMALA

Question: is it good to use non-reversible proposals within Metropolis-Hastings?

Criterion (to compare with MALA): number of steps taken, in stationarity, to explore target measure

ipMALA

Question: is it good to use non-reversible proposals within Metropolis-Hastings?

Criterion (to compare with MALA): number of steps taken, in stationarity, to explore target measure

- Consider non-reversibe Langevin in $\mathbb{R}^N$

$$dX_t = -\nabla V(X_t)dt + \gamma(X_t)dt + \sqrt{2}dW_t, \qquad \text{with } \nabla \cdot (\gamma(z)e^{-V(z)}) = 0\,.$$

ipMALA

Question: is it good to use non-reversible proposals within Metropolis-Hastings?

Criterion (to compare with MALA): number of steps taken, in stationarity, to explore target measure

- Consider non-reversibe Langevin in $\mathbb{R}^N$

$$dX_t = -\nabla V(X_t)dt + \gamma(X_t)dt + \sqrt{2}dW_t, \qquad \text{with } \nabla \cdot (\gamma(z)e^{-V(z)}) = 0\,.$$

- Choose $\gamma(X_t) = S\nabla V(X_t)$, S antisymmetric matrix

$$dX_t = -\nabla V(X_t)dt + S\nabla V(X_t)dt + \sqrt{2}dW_t$$

ipMALA

Question: is it good to use non-reversible proposals within Metropolis-Hastings?

Criterion (to compare with MALA): number of steps taken, in stationarity, to explore target measure

- Consider non-reversibe Langevin in $\mathbb{R}^N$

$$dX_t = -\nabla V(X_t)dt + \gamma(X_t)dt + \sqrt{2}dW_t, \qquad \text{with } \nabla \cdot \left(\gamma(z)e^{-V(z)}\right) = 0\,.$$

- Choose $\gamma(X_t) = S\nabla V(X_t)$, S antisymmetric matrix

$$dX_t = -\nabla V(X_t)dt + S\nabla V(X_t)dt + \sqrt{2}dW_t$$

- Suppose we want to sample from a Gaussian

$$\pi(x) \propto e^{-\sum_{i=1}^{N}\left|x^i\right|^2/\lambda_i^2} \qquad x = (x^1, \dots, x^N)$$

that is,

$$\pi(x) \sim \mathcal{N}(0, C_N), \qquad C_N = diag\{\lambda_1, \dots, \lambda_N\}\,.$$

ipMALA

- Non reversible Langevin to sample from $\pi(x) \sim \mathcal{N}(0, C_N)$

$$dX_t = -(C_N)^{-1} X_t \, dt + S_N (C_N)^{-1} X_t + \sqrt{2} dW_t, \qquad X_t \in \mathbb{R}^N.$$

ipMALA

- Non reversible Langevin to sample from $\pi(x) \sim \mathcal{N}(0, C_N)$

$$dX_t = -(C_N)^{-1} X_t \, dt + S_N (C_N)^{-1} X_t + \sqrt{2} dW_t, \qquad X_t \in \mathbb{R}^N.$$

- Rescale and obtain

$$dX_t = \left[-\frac{1}{2} X_t + C_N S_N X_t \right] dt + (C_N)^{1/2} dW_t$$

ipMALA

- Non reversible Langevin to sample from $\pi(x) \sim \mathcal{N}(0, C_N)$

$$dX_t = -(C_N)^{-1} X_t \, dt + S_N (C_N)^{-1} X_t + \sqrt{2} dW_t, \qquad X_t \in \mathbb{R}^N.$$

- Rescale and obtain

$$dX_t = \left[-\frac{1}{2} X_t + C_N S_N X_t \right] dt + (C_N)^{1/2} dW_t$$

- Use a time- step Euler discretization of the above as M-H proposal

$$y^N_{k+1} = x^N_k - \frac{1}{2} \sigma^2_N x^N_k + \sigma^\alpha_N C_N S_N x^N_k + \sigma_N (C^N)^{1/2} z^N_{k+1}$$

where

$$\sigma_N = \frac{\ell}{N^\gamma}, \qquad \ell, \gamma, \alpha > 0$$

$$y_{k+1}^N = x_k^N - \frac{1}{2}\sigma_N^2 x_k^N + \sigma_N^\alpha C_N S_N x_k^N + \sigma_N (C_N)^{1/2} z_{k+1}^N, \quad \sigma_N = \frac{\ell}{N^\gamma}$$

$$y_{k+1}^N = x_k^N - \frac{1}{2}\sigma_N^2 x_k^N + \sigma_N^\alpha C_N S_N x_k^N + \sigma_N (C_N)^{1/2} z_{k+1}^N, \quad \sigma_N = \frac{\ell}{N^\gamma}$$

- Consider continuous interpolant of the chain

$$x^{(N)}(t) = (N^{\zeta\gamma} t - k) x_{k+1}^N + (k+1 - N^{\zeta\gamma} t) x_k^N, \qquad \frac{k}{N^{\zeta\gamma}} \le t < \frac{k+1}{N^{\zeta\gamma}},$$

$$\zeta = \alpha \quad \text{if } \alpha < 2 \qquad \text{and} \qquad \zeta = 2 \quad \text{if } \alpha \geq 2.$$

$$y_{k+1}^N = x_k^N - \frac{1}{2}\sigma_N^2 x_k^N + \sigma_N^\alpha C_N S_N x_k^N + \sigma_N (C_N)^{1/2} z_{k+1}^N, \quad \sigma_N = \frac{\ell}{N^\gamma}$$

- Consider continuous interpolant of the chain

$$x^{(N)}(t) = (N^{\zeta\gamma} t - k) x_{k+1}^N + (k + 1 - N^{\zeta\gamma} t) x_k^N, \qquad \frac{k}{N^{\zeta\gamma}} \le t < \frac{k+1}{N^{\zeta\gamma}},$$

$$\zeta = \alpha \quad \text{if } \alpha < 2 \qquad \text{and} \qquad \zeta = 2 \quad \text{if } \alpha \ge 2.$$

i) Diffusive regime when $\alpha \ge 2$ $\longrightarrow$ SDE limit – cost is $O(N^{2\gamma})$

$$dX_t = -\frac{\ell^2}{2} h_1 X_t \, dt + h_2 \tilde{S} x \, dt + 2\sqrt{h_1} dW_t$$

ii) Fluid regime $\alpha < 2$ $\longrightarrow$ ODE limit – cost is $O(N^{\gamma\alpha})$ – Potential for improvement

$$dX_t = \bar{h} \tilde{S} x \, dt$$

[1] Chii-Ruey Hwang, Shu-Yin Hwang-Ma, and Shuenn-Jyi Sheu. *Accelerating diffusions.* (2005)

[2] M.O., N. S. Pillai, F. J. Pinski, A.M. Stuart. *A function space HMC algorithm with second order Langevin diffusion limit.* Bernoulli, 2016.

[3] L. Rey-Bellet, K. Spiliopoulos. *Irreversible Langevin samplers and variance reduction: a large deviations approach.* (2015)

[4] A. Bouchard-Cote, A. Doucet, S. Vollmer. *The Bouncy Particle Sampler.* (2017)

[5] J. Bierkens, P. Fearnhead, G. Roberts. *The Zig-Zag Process and super-efficient sampling*. (2016)

[6] A. Duncan, T. Lelievre, G. Pavliotis. *Variance Reduction using non-reversible Langevin Samplers* (2015)

[7] M.O., N. Pillai, K. Spiliopoulos. *Optimal Scaling of the MALA algorithm with irreversible proposals for Gaussian targets* (2017)