Nonreversible MCMC from conditional invertible transforms: a complete recipe with convergence guarantees

Alain Durmus[1]

Joint work with: Christophe Andrieu[2], Eric Moulines[3] and Achille Thin[3], Nikita Kotolevskii[4] and Maxim Panov[4]

[1]ENS Paris-Saclay

[2]University of Bristol

[3]Ecole Polytechnique

[4]CDISE, Skolkovo Institute of Science and Technology

January 4, 2021

Setting and problematic

- Consider a target distribution π on $(\mathsf{Z}, \mathcal{Z})$.
- Goal: getting samples with distribution $\approx \pi$.
- For this, the Metropolis-Hastings (MH) recipe is a workhorse.
- Classical recipe to construct a Markov kernel reversible w.r.t. π.
- Reversibility ensures that π is invariant.
- Yet, reversibility is not necessarily desirable when considering performance.

Reversible vs non-reversible MCMC

- Recent interest in the design of non-reversible or irreversible MCMC
- non-reversible MCMC = MCMC defining a Markov kernel which is not reversible but π is still invariant.
- Regarding the design of non-reversible MCMC:
  - do not follow from a general recipe as the classical MH algorithm;
  - we aim at filling this gap.
- More precisely, our goal:
  - Give a general recipe to construct non-reversible MCMC methods targeting π;
  - give simple conditions ensuring their convergence.

The classical Metropolis-Hastings recipe

Algorithm 1: the Metropolis-Hastings algorithm

- First recall the MH method defining a Markov chain $(Z_k)_{k\in\mathbb{N}}$:
  - Input:
    - Initial state Z_0;
    - Proposal kernel Q on Z;
    - Acceptance probability $\alpha : \mathsf{Z}^2 \to [0,1]$;
  - at stage $k+1$:
    - sample a proposal $Y_{k+1} \sim Q(Z_k, \cdot)$;
    - Set $Z_{k+1} = Y_{k+1}$ with probability $\alpha(Z_k, Y_{k+1})$;
    - Set $Z_{k+1} = Z_k$ otherwise.

The classical Metropolis-Hastings recipe: the RWM

Algorithm 2: the random walk Metropolis-Hastings algorithm

- Example: the random walk Metropolis algorithm (RWM) on $\mathsf{Z} = \mathbb{R}^d$:
  - Input:
    - Initial state Z_0;
    - Proposal kernel Q on Z;
    - Example: $Q(z_0, \mathsf{A}) = \int_{\mathsf{A}} \varphi(z_1 - z_0) \mathrm{d}z_1$ where φ is a symmetric density;
    - Acceptance probability $\alpha : \mathsf{Z}^2 \to [0, 1]$;
    - Example: $\alpha(z_0, z_1) = 1 \wedge [\pi(z_1)/\pi(z_0)]$;
  - at stage $k + 1$:
    - sample a proposal $Y_{k+1} \sim Q(Z_k, \cdot)$;
    - Example: $Y_{k+1} = Z_k + G_{k+1}$, $G_{k+1} \sim \varphi$;
    - Set $Z_{k+1} = Y_{k+1}$ with probability $\alpha(Z_k, Y_{k+1})$;
    - Set $Z_{k+1} = Z_k$ otherwise.

The classical Metropolis-Hastings recipe: reversibility condition

- The Markov kernel associated with $(Z_k)_{k\in\mathbb{N}}$ is given for any $z \in \mathsf{Z}$ and $\mathsf{A} \in \mathcal{Z}$,

$$P(z_0, \mathsf{A}) = \int_{\mathsf{A}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) + \delta_{z_0}(\mathsf{A}) \int_{\mathsf{Z}} \{1 - \alpha(z_0, z_1)\} Q(z_0, \mathrm{d}z_1) \ .$$

- General necessary and sufficient conditions on α and Q implying that π is reversible with respect to P given in [Tie98].
- Recall that P is reversible with respect to π if for any $f, g : \mathsf{Z} \to \mathbb{R}_+$, bounded,

$$\int_{\mathsf{Z}^2} f(z_0) g(z_1) \pi(\mathrm{d}z_0) P(z_0, \mathrm{d}z_1) = \int_{\mathsf{Z}^2} f(z_1) g(z_0) \pi(\mathrm{d}z_0) P(z_0, \mathrm{d}z_1) \ .$$

- In probabilistic language: if $Z_0 \sim \pi$, $Z_1 | Z_0 \sim P(Z_0, \cdot)$, then

$$(Z_0, Z_1) \overset{\text{law}}{=} (Z_1, Z_0) \ .$$

The classical Metropolis-Hastings recipe: reversibility condition

- The Markov kernel associated with $(Z_k)_{k\in\mathbb{N}}$ is given for any $z \in \mathsf{Z}$ and $\mathsf{A} \in \mathcal{Z}$,

$$P(z_0, \mathsf{A}) = \int_{\mathsf{A}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) + \delta_{z_0}(\mathsf{A}) \int_{\mathsf{Z}} \{1 - \alpha(z_0, z_1)\} Q(z_0, \mathrm{d}z_1) \; .$$

- General necessary and sufficient conditions on α and Q implying that π is reversible with respect to P given in [Tie98].
- Recall that P is reversible with respect to π if for any $f, g : \mathsf{Z} \to \mathbb{R}_+$, bounded,

$$\int_{\mathsf{Z}^2} f(z_0) g(z_1) \pi(\mathrm{d}z_0) P(z_0, \mathrm{d}z_1) = \int_{\mathsf{Z}^2} f(z_1) g(z_0) \pi(\mathrm{d}z_0) P(z_0, \mathrm{d}z_1) \; .$$

- In measure theoretical language:

Classical Metropolis-Hastings algorithm: reversibility

- Recall that P is reversible with respect to π if for any $f, g : \mathsf{Z} \to \mathbb{R}_+$, bounded,

$$\int_{\mathsf{Z}^2} f(z_0)g(z_1)\pi(\mathrm{d}z_0)P(z_0, \mathrm{d}z_1) = \int_{\mathsf{Z}^2} f(z_1)g(z_0)\pi(\mathrm{d}z_0)P(z_0, \mathrm{d}z_1) \ .$$

- In the case where π and P are absolutely continuous with respect to a common dominating measure μ, denoted by π and p:

$$\pi(z_0)p(z_0, z_1) = \pi(z_1)p(z_1, z_0) \ .$$

- But P does not admit a density in general!
- except in the discrete setting...

Reversibility of MH if $\exists$ common dominating measure

- In the case Q and π have a common dominating measure μ on $(\mathsf{Z}, \mathcal{Z})$.
- Denote by q and π their densities.
- Assume that they are positive.
- Then,

$$P(z_0, \mathsf{A}) = \int_{\mathsf{A}} \alpha(z_0, z_1) q(z_0, z_1) \mathrm{d}\mu(z_1) + \delta_{z_0}(\mathsf{A}) \int_{\mathsf{Z}} \{1 - \alpha(z_0, z_1)\} q(z_0, z_1) \mathrm{d}\mu(z_1) \,.$$

Reversibility of MH if $\exists$ common dominating measure

- Existence of a common dominating measure μ: P is given by

$$
\begin{aligned}
P(z_0, \mathsf{A}) = \int_{\mathsf{A}} \alpha(z_0, z_1) q(z_0, z_1) \mathrm{d}\mu(z_1) \\
+ \delta_{z_0}(\mathsf{A}) \int_{\mathsf{Z}} \{1 - \alpha(z_0, z_1)\} q(z_0, z_1) \mathrm{d}\mu(z_1) \\
= Q_\alpha(z_0, \mathsf{A}) + \delta_{z_0}(\mathsf{A}) \{1 - Q_\alpha(z_0, \mathsf{Z})\} \, .
\end{aligned}
$$

- Q_α is the absolutely continuous part of P:
  - it is a sub-Markovian kernel on $(\mathsf{Z}, \mathcal{Z})$: $Q_\alpha(z_0, \mathsf{Z}) \neq 1$;
  - admitting the transition density:

$$
Q_\alpha(z_0, \mathsf{A}) = \int_{\mathsf{A}} q_\alpha(z_0, z_1) \mathrm{d}\mu(z_1) \, , \quad q_\alpha : (z_0, z_1) \mapsto \alpha(z_0, z_1) q(z_0, z_1) \, .
$$

Reversibility of MH if $\exists$ common dominating measure

- Existence of a common dominating measure μ: P is given by

$$P(z_0, \mathsf{A}) = Q_\alpha(z_0, \mathsf{A}) + \delta_{z_0}(\mathsf{A})\{1 - Q_\alpha(z_0, \mathsf{Z})\} ,$$
$$Q_\alpha(z_0 \mathsf{A}) = \int_{\mathsf{A}} q_\alpha(z_0, z_1)\mathrm{d}\mu(z_1) , \quad q_\alpha : (z_0, z_1) \mapsto \alpha(z_0, z_1)q(z_0, z_1) .$$

- It turns out that P is reversible w.r.t. π if Q_α is "reversible" w.r.t. π:

$$\pi(z_0)q_\alpha(z_0, z_1) = \pi(z_1)q_\alpha(z_1, z_0) .$$

- Then a necessary and sufficient condition on α is

$$\pi(z_0)\alpha(z_0, z_1)q(z_0, z_1) = \pi(z_1)\alpha(z_1, z_0)q(z_1, z_0) .$$

Reversibility of MH if $\exists$ common dominating measure

- Existence of a common dominating measure μ: P is given by

$$P(z_0, \mathsf{A}) = Q_\alpha(z_0, \mathsf{A}) + \delta_{z_0}(\mathsf{A})\{1 - Q_\alpha(z_0, \mathsf{Z})\} .$$

- P is reversible w.r.t. π if

$$\pi(z_0)\alpha(z_0, z_1)q(z_0, z_1) = \pi(z_1)\alpha(z_1, z_0)q(z_1, z_0) .$$

- It is satisfied since π and q are positive if

$$\alpha(z_0, z_1) = \mathrm{a}\left[\frac{\pi(z_1)q(z_1, z_0)}{\pi(z_0)q(z_0, z_1)}\right] ,$$

where $\mathrm{a} : \mathbb{R}_+ \to [0, 1]$ satisfying $\mathrm{a}(0) = 0$,

$$t\mathrm{a}(1/t) = \mathrm{a}(t) .$$

- Examples:

$$\mathrm{a}(t) = \min(1, t) \text{ leads to the classical Metropolis ratio } ,$$
$$\mathrm{a}(t) = t/(1 + t) \text{ leads to the Barker ratio } .$$

Reversibility of MH in the general case

- Going back to the general formulation:

$$P(z_0, \mathsf{A}) = \int_{\mathsf{A}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) + \delta_{z_0}(\mathsf{A}) \int_{\mathsf{Z}} \{1 - \alpha(z_0, z_1)\} Q(z_0, \mathrm{d}z_1) \,.$$

- Question:
  what about the case Q and π do not have a common dominating measure?
- This questions have been addressed in [Tie98].
- The main idea is to construct a symmetric common dominating measure.

Reversibility of the MH algorithm: the general case

- In hindsight, what does the reversibility condition means? what do we need?
- Reversibility condition for P means:

$$\tilde{\nu}_P(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z) P(z, \mathsf{B}) \; ,$$

as a probability measure on Z^2 is equal to

$$\tilde{\nu}_P^s(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} \pi(\mathrm{d}z) P(z, \mathsf{A}) \; .$$

- Note that $\tilde{\nu}_P^s$ is the pushforward measure of $\tilde{\nu}_P$ by $(z_0, z_1) \mapsto (z_1, z_0)$ on Z^2.

Reversibility of the MH algorithm: the general case

- Reversibility for P if Q_α is reversible.
- Reversibility for Q_α means: $\tilde{\nu}_\alpha = \tilde{\nu}_\alpha^s$ where

$$\tilde{\nu}_\alpha(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z_0) \int_{\mathsf{B}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) \, ,$$

$$\tilde{\nu}_\alpha^s(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} \pi(\mathrm{d}z_0) \int_{\mathsf{A}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) \, .$$

- Note that $\tilde{\nu}_\alpha^s$ is the pushforward measure of $\tilde{\nu}_\alpha$ by $(z_0, z_1) \mapsto (z_1, z_0)$ on Z^2.

Reversibility of the MH algorithm: the general case

- Reversibility for Q_α means: $\tilde{\nu}_\alpha = \tilde{\nu}^s_\alpha$ where

$$\tilde{\nu}_\alpha(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z_0) \int_{\mathsf{B}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) \,,$$

$$\tilde{\nu}^s_\alpha(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} \pi(\mathrm{d}z_0) \int_{\mathsf{A}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) \,.$$

- These two measures admit the density $(z_0, z_1) \mapsto \alpha(z_0, z_1)$ with respect to the probability measures $\tilde{\nu}$ on Z^2:

$$\tilde{\nu}_Q(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z_0) \int_{\mathsf{B}} Q(z_0, \mathsf{B}) \,, \quad \tilde{\nu}^s_Q(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} \pi(\mathrm{d}z_0) \int_{\mathsf{A}} Q(z_0, \mathsf{B}) \,.$$

- Note that $\tilde{\nu}^s_Q$ is the pushforward measure of $\tilde{\nu}_Q$ by $(z_0, z_1) \mapsto (z_1, z_0)$ on Z^2.

Reversibility of the MH algorithm: the general case

- We consider the probability measures on Z^2:

$$\tilde{\nu}_Q(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z_0) \int_{\mathsf{B}} Q(z_0, \mathsf{B}) \;, \quad \tilde{\nu}_Q^s(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} \pi(\mathrm{d}z_0) \int_{\mathsf{A}} Q(z_0, \mathsf{B}) \;.$$

- The assumption that Q and π have a common dominating measure:

$$\frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = \pi(z_0) q(z_0, z_1) \;, \quad \frac{\mathrm{d}\tilde{\nu}_Q^s}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = \pi(z_1) q(z_1, z_0) \;.$$

- Reversibility for Q_α, i.e. $\tilde{\nu}_\alpha = \tilde{\nu}_\alpha^s$, is equivalent in that case to:

$$\pi(z_0)\alpha(z_0, z_1) q(z_0, z_1) = \pi(z_1)\alpha(z_1, z_0) q(z_1, z_0) \;.$$

- Choice for α enforces reversibility with respect to π!

Reversibility of the MH algorithm: the general case

- The assumption that Q and π have a common dominating measure:

$$
\frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = h_Q(z_0, z_1) = \pi(z_0) q(z_0, z_1) \;,
$$
$$
\frac{\mathrm{d}\tilde{\nu}_Q^s}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = h_Q^s(z_0, z_1) = \pi(z_1) q(z_1, z_0) \;.
$$

- Previous examples for α

$$
\alpha(z_0, z_1) = \mathrm{a}\left[\frac{\pi(z_1) q(z_1, z_0)}{\pi(z_0) q(z_0, z_1)}\right] \;,
$$

assume also π and q are positive.

- We can identify the ratio of the density:

$$
\frac{\pi(z_1) q(z_1, z_0)}{\pi(z_0) q(z_0, z_1)} = \frac{h_Q^s(z_0, z_1)}{h_Q(z_0, z_1)} \;.
$$

Reversibility of the MH algorithm: the general case

- The assumption that Q and π have a common dominating measure:

$$\frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = \pi(z_0) q(z_0, z_1) \;, \quad \frac{\mathrm{d}\tilde{\nu}_Q^s}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = \pi(z_1) q(z_1, z_0) \;.$$

- Reversibility for Q_α, i.e. $\tilde{\nu}_\alpha = \tilde{\nu}_\alpha^s$, is equivalent in that case to:

$$\pi(z_0)\alpha(z_0, z_1) q(z_0, z_1) = \pi(z_1)\alpha(z_1, z_0) q(z_1, z_0) \;.$$

- Previous examples for α

$$\alpha(z_0, z_1) = \mathrm{a}\left[\frac{\pi(z_1) q(z_1, z_0)}{\pi(z_0) q(z_0, z_1)}\right] \;,$$

assume also π and q are positive.

- π and q positive ensures $\tilde{\nu}_Q$ and $\tilde{\nu}_Q^s$ are equivalent: $\tilde{\nu}_Q \ll \tilde{\nu}_Q^s$, $\tilde{\nu}_Q^s \ll \tilde{\nu}_Q$ and we identify

$$\frac{\pi(z_1) q(z_1, z_0)}{\pi(z_0) q(z_0, z_1)} = \frac{h_Q^s(z_0, z_1)}{h_Q(z_0, z_1)} = \frac{\mathrm{d}\tilde{\nu}_Q^s}{\mathrm{d}\tilde{\nu}_Q}(z_0, z_1) \;.$$

Reversibility of the MH algorithm: the general case

- First question:
  what about the case Q and π do not have a common dominating measure?
  Can we find a common dominating measure for $\tilde{\nu}_Q$ and $\tilde{\nu}_Q^s$ still?
- Second question:
  what about the case Q and π have a common dominating measure but non-negative densities?
  Can we still find some functions α which enforces reversibility?
  In other word: is the condition that $\tilde{\nu}_Q$ and $\tilde{\nu}_Q^s$ equivalent really necessary?

- First question:
  what about the case Q and π do not have a common dominating measure?
  Can we find a common dominating measure for $\tilde{\nu}_Q$ and $\tilde{\nu}^s_Q$ still?
- Answer:
  - $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}^s_Q$;
  - Then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and denote

$$h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q} \,, \qquad h^s_Q = \frac{\mathrm{d}\tilde{\nu}^s_Q}{\mathrm{d}\tilde{\lambda}_Q} \,.$$

  - We can show that $h^s_Q(z_0, z_1) = h_Q(z_1, z_0)$.

Reversibility of the MH algorithm: the general case

- Second question:
  is the condition that $\tilde{\nu}_Q$ and $\tilde{\nu}^s_Q$ equivalent really necessary?
- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}^s_Q$, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and denote

$$h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q}\,, \qquad h^s_Q = \frac{\mathrm{d}\tilde{\nu}^s_Q}{\mathrm{d}\tilde{\lambda}_Q}\,.$$

- Answer:

Proposition 1: (Tierney 98, Proposition 1)

Set

$$A_Q = \left\{ h_Q \times h^s_Q > 0 \right\} \in \mathcal{Z}^{\otimes 2}\,.$$

Then, the restrictions

- $\tilde{\nu}_A(\cdot) = \tilde{\nu}(\cdot \cap A_Q)$ and $\tilde{\nu}^s_A(\cdot) = \tilde{\nu}^s(\cdot \cap A_Q)$ are equivalent;
- $\tilde{\nu}_{A,\mathrm{c}}(\cdot) = \tilde{\nu}(\cdot \cap A^{\mathrm{c}}_Q)$ and $\tilde{\nu}^s_{A,\mathrm{c}}(\cdot) = \tilde{\nu}^s(\cdot \cap A^{\mathrm{c}}_Q)$ are mutually singular.

Reversibility of the MH algorithm: the general case

- Second question:
  is the condition that $\tilde{\nu}_Q$ and $\tilde{\nu}^s_Q$ equivalent really necessary?
- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}^s_Q$, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and denote

$$h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q} , \qquad h^s_Q = \frac{\mathrm{d}\tilde{\nu}^s_Q}{\mathrm{d}\tilde{\lambda}_Q} .$$

- Answer:

Proposition 2: (Tierney 98, Proposition 1)

Set

$$A_Q = \left\{ h_Q \times h^s_Q > 0 \right\} \in \mathcal{Z}^{\otimes 2} .$$

Define, for $(z_0, z_1) \in A_Q$,

$$r_Q(z_0, z_1) = h_Q(z_0, z_1) / h^s_Q(z_0, z_1) .$$

Then, r_Q is a version of the density of $\tilde{\nu}_A$ w.r.t. $\tilde{\nu}^s_A$, *i.e.* $r_Q = \mathrm{d}\tilde{\nu}_A / \mathrm{d}\tilde{\nu}^s_A$.

Reversibility of the MH algorithm: the general case

- Based on the previous results, [Tie98] gives a necessary and sufficient condition on α so that P is π-reversible.
- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}^s_Q$, then, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and denote

$$h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q}\,, \qquad h^s_Q = \frac{\mathrm{d}\tilde{\nu}^s_Q}{\mathrm{d}\tilde{\lambda}_Q}\,.$$

- Set $A_Q = \left\{h_Q \times h^s_Q > 0\right\} \in \mathcal{Z}^{\otimes 2}$ and

$$r_Q(z_0, z_1) = h_Q(z_0, z_1)/h^s_Q(z_0, z_1)\,, (z_0, z_1) \in A_Q\,.$$

Theorem 1: (Tierney 1998, Thereom 2)

The sub-Markovian kernel Q_α is π-reversible if and only if the following conditions hold.

- The function α is zero $\tilde{\nu}_Q$-a.e.on A^c_Q.
- The function α satisfies $\alpha(z_0, z_1) r_Q(z_0, z_1) = \alpha(z_1, z_0)$ $\tilde{\nu}_Q$-a.e.on A_Q.

Reversibility of the MH algorithm: the general case

- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}_Q^s$, then, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and denote

$$h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q} \,, \qquad h_Q^s = \frac{\mathrm{d}\tilde{\nu}_Q^s}{\mathrm{d}\tilde{\lambda}_Q} \,.$$

- Set $A_Q = \left\{ h_Q \times h_Q^s > 0 \right\} \in \mathcal{Z}^{\otimes 2}$ and

$$r_Q(z_0, z_1) = h_Q(z_0, z_1) / h_Q^s(z_0, z_1) \,, (z_0, z_1) \in A_Q \,.$$

- We can then define the Metropolis-Hastings rejection probability by

$$\alpha(z_0, z_1) = \begin{cases} \mathrm{a}\left(\frac{h_Q^s(z_0,z_1)}{h_Q(z_0,z_1)}\right) = \mathrm{a}\left(1/r_Q(z_0, z_1)\right) & h_Q(z_0, z_1) \neq 0 \,, \\ 1 & h_Q(z_0, z_1) = 0 \,, \end{cases}$$

  where $\mathrm{a} : \mathbb{R}_+ \to [0, 1]$ satisfies $\mathrm{a}(0) = 0$, $t\mathrm{a}(1/t) = \mathrm{a}(t)$.
- This choice of α ensures reversibility for Q_α and therefore for the MH kernel P.

Reversibility of the MH kernel: $\exists$ common dominating measure

- If π and Q have densities with respect to μ:

$$\frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = h_Q(z_0, z_1) = \pi(z_0)q(z_0, z_1) \; ,$$
$$\frac{\mathrm{d}\tilde{\nu}^s_Q}{\mathrm{d}\mu^{\otimes 2}}(z_0, z_1) = h^s_Q(z_0, z_1) = \pi(z_1)q(z_1, z_0) \; .$$

and therefore:

$$\alpha(z_0, z_1) = \begin{cases} \mathrm{a}\left[\frac{\pi(z_1)q(z_1,z_0)}{\pi(z_0)q(z_0,z_1)}\right] & \pi(z_0)q(z_0, z_1) \neq 0, \\ 1 & \pi(z_0)q(z_0, z_1) = 0 \; , \end{cases}$$

where $\mathrm{a} : \mathbb{R}_+ \to [0, 1]$ satisfies $\mathrm{a}(0) = 0, \; t\mathrm{a}(1/t) = \mathrm{a}(t)$.

Reversibility of MH kernel: deterministic proposal

- Suppose now that Φ is an involution from Z onto Z: such that $\Phi^{-1} = \Phi$.
- We consider the deterministic proposal kernel

$$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1) \text{: when the current state is } z_0\text{, the proposal is } \Phi(z_0)\ .$$

- Setting introduced in [Tie98, Section 2] and more recently in [Nek+20].

Reversibility of MH kernel: deterministic proposal

- Suppose now that Φ is an involution from Z onto Z: such that $\Phi^{-1} = \Phi$.
- We consider the deterministic proposal kernel

$$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)\text{: when the current state is } z_0\text{, the proposal is } \Phi(z_0) \quad .$$

- In this scenario we have that

$$\tilde{\nu}\big(\mathrm{d}(z_0, z_1)\big) = \pi(\mathrm{d}z_0)\delta_{\Phi(z_0)}(\mathrm{d}z_1) \text{ and } \tilde{\nu}^s\big(\mathrm{d}(z_0, z_1)\big) = \pi(\mathrm{d}z_1)\delta_{\Phi^{-1}(z_1)}(\mathrm{d}z_0) \quad .$$

Reversibility of MH kernel: deterministic proposal

- Suppose now that Φ is an involution from Z onto Z: such that $\Phi^{-1} = \Phi$.
- We consider the deterministic proposal kernel

$$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)\text{: when the current state is } z_0\text{, the proposal is } \Phi(z_0) \;.$$

- The function h_Q is given by

$$h_Q(z_0, z_1) = \mathbb{1}_{\Phi(z_0)}(z_1) k(z_0) \text{ with } \quad k(z) = \frac{\mathrm{d}\pi}{\mathrm{d}\lambda}(z), \quad \lambda = \pi + (\Phi^{-1})_{\#}\pi \;,$$
$$h_Q^s(z_0, z_1) = \mathbb{1}_{\Phi(z_1)}(z_0) k(z_1) = \mathbb{1}_{\Phi(z_0)}(z_1) k(\Phi(z_0)) \;.$$

- Therefore, $\alpha\big(z, \Phi(z)\big) = \bar{\alpha}(z)$ with

$$\bar{\alpha}(z) = \begin{cases} \mathrm{a}\left(\frac{k\left(\Phi(z)\right)}{k(z)}\right) & \text{if } k(z) > 0 \;, \\ 1 & \text{otherwise} \;. \end{cases}$$

Of course, there is no need to define $\alpha(z_0, z_1)$ for $z_1 \neq \Phi(z_0)$.

- Computation of k?

Reversibility of MH kernel: deterministic proposal

- Suppose now that Φ is an involution from Z onto Z: such that $\Phi^{-1} = \Phi$.
- We consider the deterministic proposal kernel

$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)$: when the current state is z_0, the proposal is $\Phi(z_0)$.

- A special case of interest is when $\mathsf{Z} = \mathbb{R}^d$ and $\pi(\mathrm{d}z) = \pi(z)\,\mathrm{Leb}_d(\mathrm{d}z)$.
- Here the dominating measure $\tilde{\lambda}$ is given by

$$\tilde{\lambda}(\mathrm{d}z) = \pi + (\Phi^{-1})_{\#}\pi = \{\pi(z) + \pi \circ \Phi(z)\mathrm{Jac}_{\Phi}(z)\}\,\mathrm{Leb}_d(\mathrm{d}z)\ ,$$

where Jac_f denotes the absolute value of the Jacobian determinant of f.

Reversibility of MH kernel: deterministic proposal

- Suppose now that Φ is an involution from Z onto Z: such that $\Phi^{-1} = \Phi$.
- We consider the deterministic proposal kernel

$$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1) \text{: when the current state is } z_0 \text{, the proposal is } \Phi(z_0) \; .$$

- A special case of interest is when $\mathsf{Z} = \mathbb{R}^d$ and $\pi(\mathrm{d}z) = \pi(z)\,\mathrm{Leb}_d(\mathrm{d}z)$.
- Then, the density $k(z)$ is given by

$$k(z) = \frac{\mathrm{d}\pi}{\mathrm{d}\tilde{\lambda}}(z) = \frac{\pi(z)}{\pi(z) + \pi \circ \Phi(z) \mathrm{Jac}_\Phi(z)} \; .$$

- The acceptance ratio $\bar{\alpha}(z)$ takes the simple form

$$\bar{\alpha}(z) = \begin{cases} \mathrm{a}\left(\frac{\pi \circ \Phi(z) \mathrm{Jac}_\Phi(z)}{\pi(z)}\right) \; , & \text{if } \pi(z) \neq 0 \\ 1 & \text{otherwise} \; . \end{cases}$$

1 Introduction

2 The reversible Metropolis-Hastings recipe

- The MH algorithm and reversibility

3 Non-reversible MH algorithms

- Motivating examples
- (π, S)-reversibility and the Generalized MH rule

The standard random walk Metropolis

- For simplicity assume that $Z_d = \{1, \ldots, d\}$ and we are interested in sampling from π.

Algorithm 1: RWM

At stage $k+1$ and state Z_k

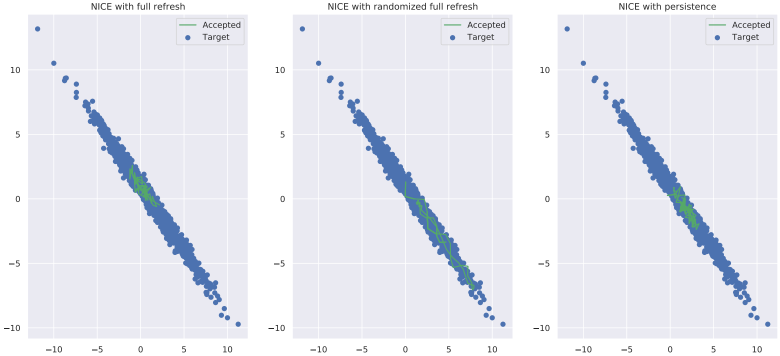
- Sample an increment $V_{k+1} \sim \mathrm{Unif}(\{-1, 1\})$
- Compute

$$\alpha(Z_k, Z_k + V_{k+1}) = \min\left\{1, \frac{\pi(Z_k + V_{k+1})}{\pi(Z_k)}\right\}$$

- Set $Z_{k+1} = Z_k + V_{k+1}$ with probability $\alpha(Z_k, Z_k + V_{k+1})$, otherwise set $Z_{k+1} = Z_k$.

- The chain can "backtrack", which we may want to avoid (we have $\pi(x)P(x,y) = \pi(y)P(y,x)$).

Reversible vs. nonreversible



Reversible **Non-reversible** **Non-reversible**

Motivation–the guided random walk

- Gustafson's guided walk Metropolis method [Gus98; Hor91] addresses the backtracking problem.
- The guided walk Metropolis works for π defined on $\mathsf{Z} = \mathbb{R}^d, \mathbb{Z}$, but $\mathsf{Z} = \{1, \ldots, d\}$ captures the important features required in what follows,
- We introduce the auxiliary variable $v \in \{-1, 1\} = \mathsf{V}$,
- we extend the state space:$\mathsf{E} = \mathsf{Z}_d \times \mathsf{V}$,
- and consider the new target distribution $\mu(x, v) = \pi(x)\frac{1}{2}$.

Motivation–the guided random walk

- We extend the state space $\mathrm{E} = \mathrm{Z}_d \times \mathrm{V} = \mathrm{Z}_d \times \{-1, 1\}$
- We consider $\mu(x, v) = \pi(x)\frac{1}{2}$, and the Markov transition P on E given by:

Algorithm 2: Guided random walk

At stage $k+1$ and state (Z_k, V_k),

- Calculate the acceptance ratio

$$r(Z_k, V_k) = \frac{\mu(Z_k + V_k, V_k)}{\mu(Z_k, V_k)} = \frac{\pi(Z_k + V_k)}{\pi(Z_k)}.$$

- Set $(Z_{k+1}, V_{k+1}) = (Z_k + V_k, V_k)$ with probability $\min\{1, r(Z_k, V_k)\}$, otherwise set $(Z_{k+1}, V_{k+1}) = (Z_k, -V_k)$.

- The process will travel in the same direction until a rejection occurs.
- The process is nonreversible and but satisfies

$$\mu(x, v)P(x, v; y, w) = \mu(y, w)P(y, -w; x, -v).$$

- μ is invariant for P.

Random vs. guided random walk

RWM

At stage $k+1$ and state Z_k

- Sample an increment $V_{k+1} \sim \mathrm{Unif}(\{-1,1\})$
- Compute

$$\alpha(Z_k, Z_k + V_{k+1}) = \min\left\{1, \frac{\pi(Z_k + V_{k+1})}{\pi(Z_k)}\right\}$$

- Set $Z_{k+1} = Z_k + V_{k+1}$ with probability $\alpha(Z_k, Z_k + V_{k+1})$, otherwise set $Z_{k+1} = Z_k$.

Guided random walk

At stage $k+1$ and state (Z_k, V_k),

- Calculate the acceptance ratio

$$r(Z_k, V_k) = \frac{\mu(Z_k + V_k, V_k)}{\mu(Z_k, V_k)} = \frac{\pi(Z_k + V_k)}{\pi(Z_k)}.$$

- Set $(Z_{k+1}, V_{k+1}) = (Z_k + V_k, V_k)$ with probability $\min\{1, r(Z_k, V_k)\}$, otherwise set $(Z_{k+1}, V_{k+1}) = (Z_k, -V_k)$.

- It can be shown that this deterministic behaviour most often leads to performance superior to that of the RW Metropolis in terms of asymptotic variance [AL19].

Random vs. guided random walk

RWM

At stage $k+1$ and state Z_k

- Sample an increment $V_{k+1} \sim \mathrm{Unif}(\{-1,1\})$
- Compute

$$\alpha(Z_k, Z_k + V_{k+1}) = \min\left\{1, \frac{\pi(Z_k + V_{k+1})}{\pi(Z_k)}\right\}$$

- Set $Z_{k+1} = Z_k + V_{k+1}$ with probability $\alpha(Z_k, Z_k + V_{k+1})$, otherwise set $Z_{k+1} = Z_k$.

Guided random walk

At stage $k+1$ and state (Z_k, V_k),

- Calculate the acceptance ratio

$$r(Z_k, V_k) = \frac{\mu(Z_k + V_k, V_k)}{\mu(Z_k, V_k)} = \frac{\pi(Z_k + V_k)}{\pi(Z_k)}.$$

- Set $(Z_{k+1}, V_{k+1}) = (Z_k + V_k, V_k)$ with probability $\min\{1, r(Z_k, V_k)\}$, otherwise set $(Z_{k+1}, V_{k+1}) = (Z_k, -V_k)$.

- However not quantitative, and e.g. Gustafson [Gus98] reports modest improvements (in simulations).

Convergence to equilibrium?

- We extend the state space $\mathsf{E} = \mathsf{Z}_d \times \mathsf{V} = \mathsf{Z}_d \times \{-1, 1\}$
- We consider $\pi(x) = d^{-1}$ and $\mu(x, v) = (2d)^{-1}$, and P on E: given by:

Algorithm 3: Diaconis-Neal-Holmes algorithm

At stage $k+1$ and state (Z_k, V_k),

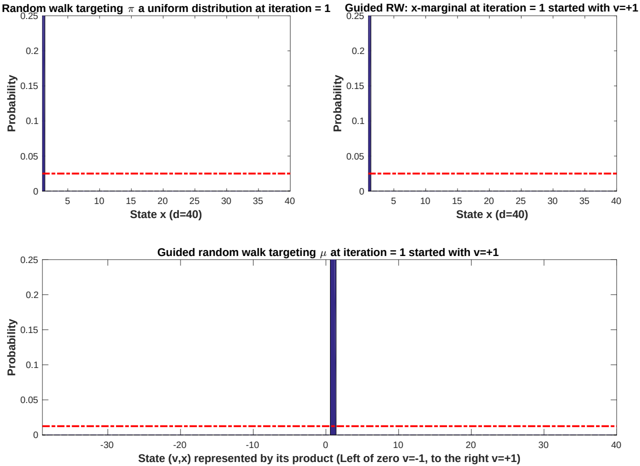
- Calculate the acceptance ratio

$$r(Z_k, V_k) = \frac{\pi(Z_k + V_k)}{\pi(Z_k)} = \begin{cases} 0 & \text{if } (Z_k, V_k) = (1, -1) \text{ or } (Z_k, V_k) = (d, +1) \\ 1 & \text{otherwise} \end{cases} .$$

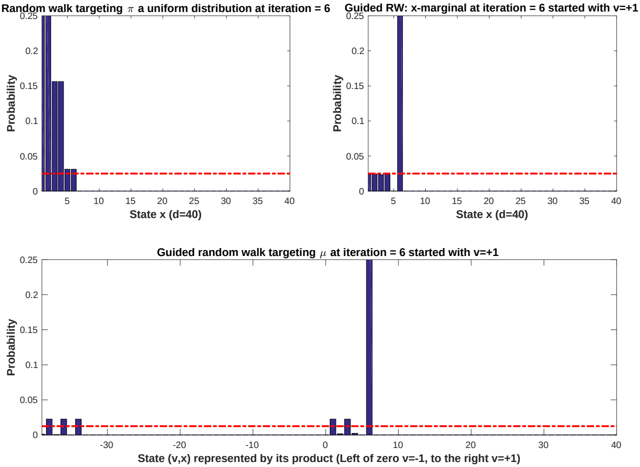
- Set $(Z_{k+1}, V'_{k+1}) = (Z_k + V_k, V_k)$ with probability $\min\{1, r(Z_k, V_k)\}$, otherwise set $(Z_{k+1}, V'_{k+1}) = (Z_k, -V_k)$.
- With probability $\theta \in [0, 1]$ set $(Z_{k+1}, V_{k+1}) = (Z_{k+1}, -V'_{k+1})$.

- Questions:
  - does it converge faster than its reversible counterpart i.e. the random walk?
  - Is there an optimal θ?
  - Dependency on d, the cardinal of Z_d?

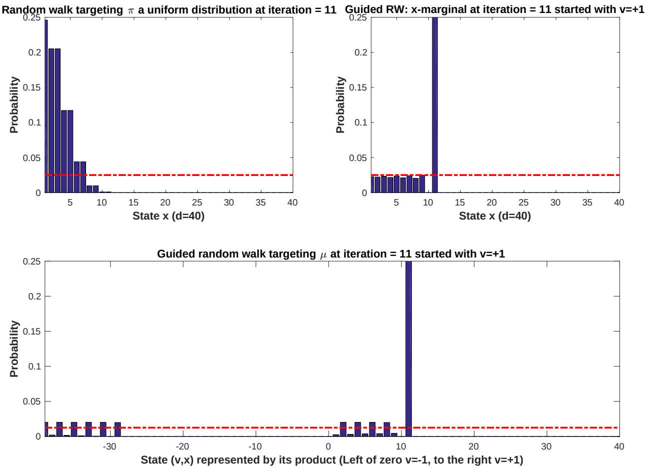
First question : reversible vs. nonreversible?



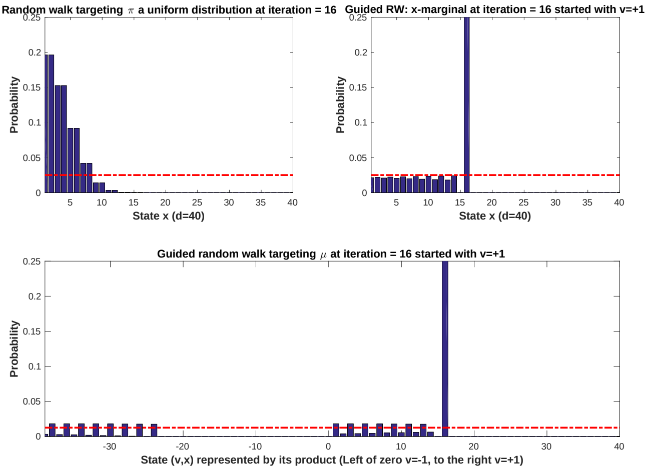
Reversible vs. nonreversible



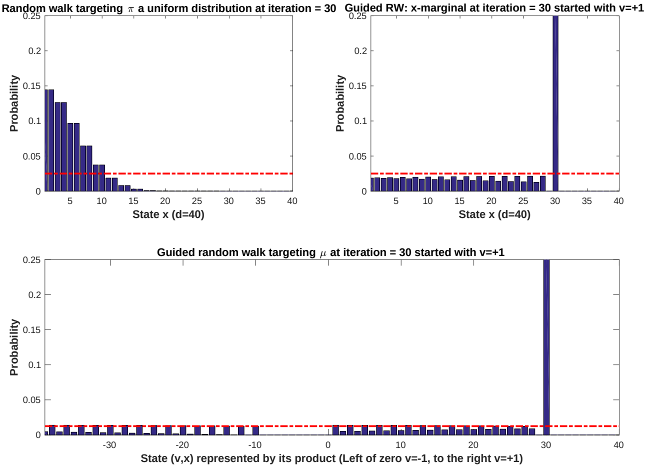
Reversible vs. nonreversible



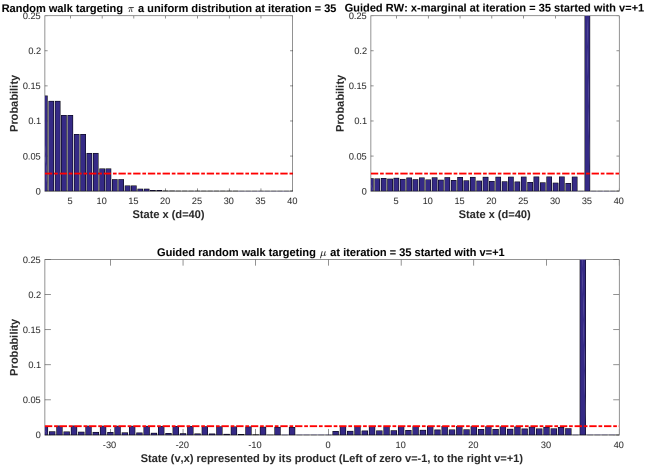
Reversible vs. nonreversible



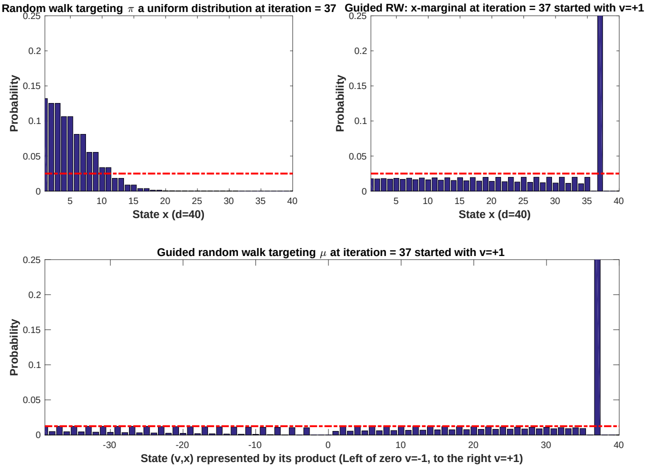
Reversible vs. nonreversible



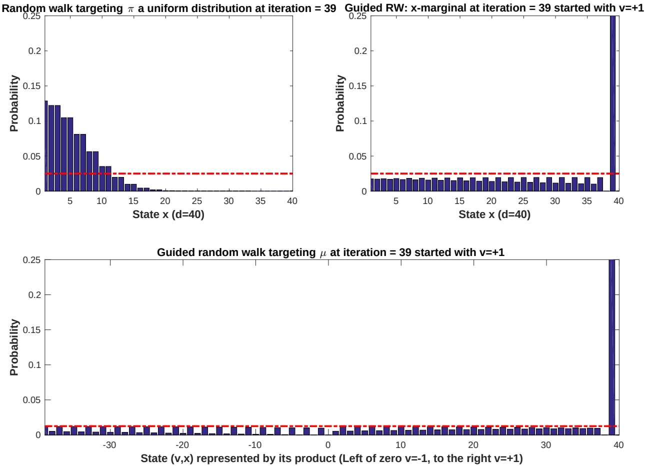
Reversible vs. nonreversible



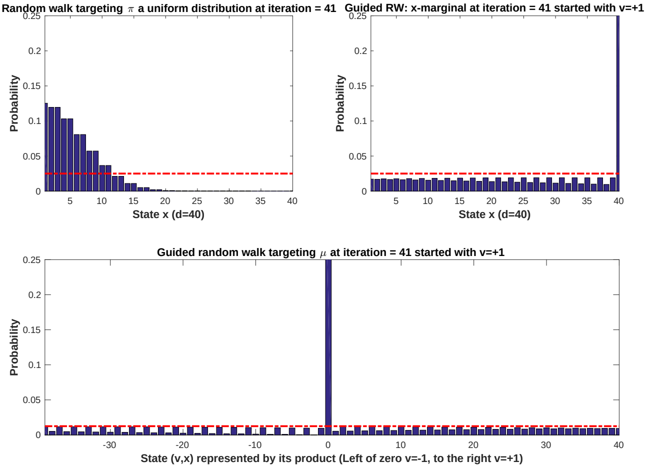
Reversible vs. nonreversible



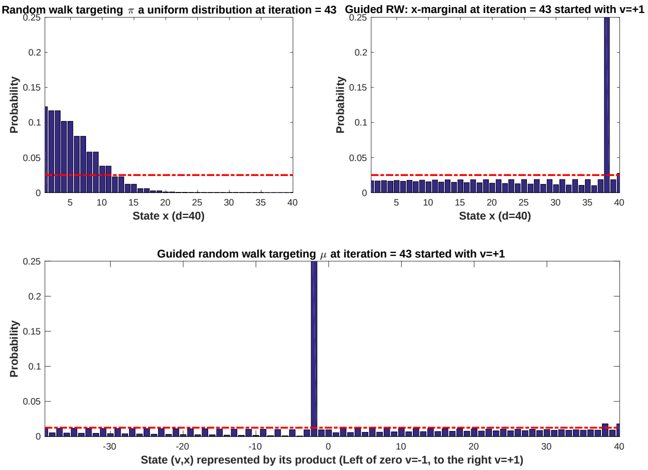
Reversible vs. nonreversible



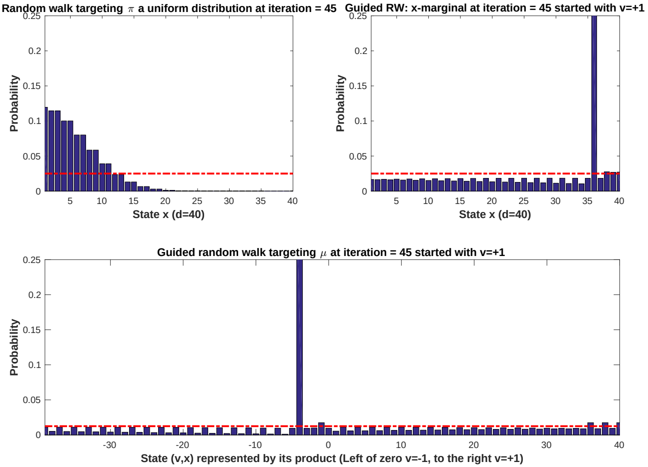
Reversible vs. nonreversible



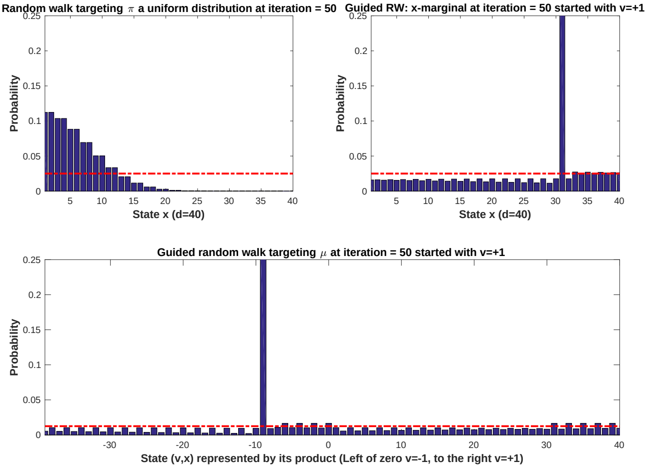
Reversible vs. nonreversible



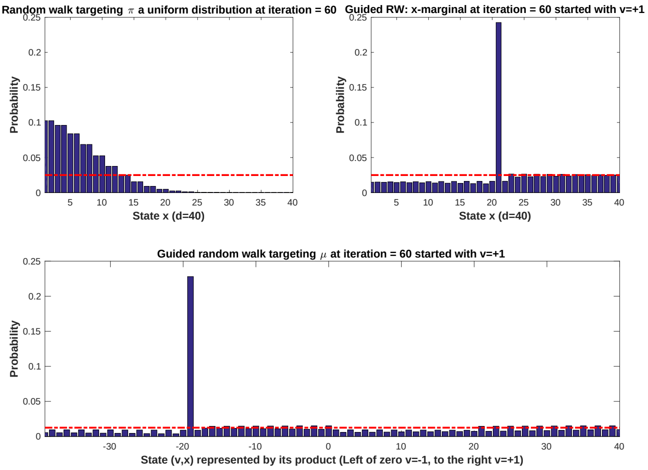
Reversible vs. nonreversible



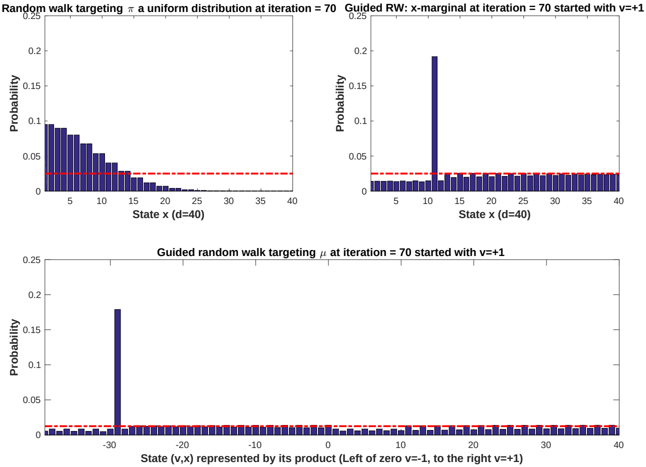
Reversible vs. nonreversible



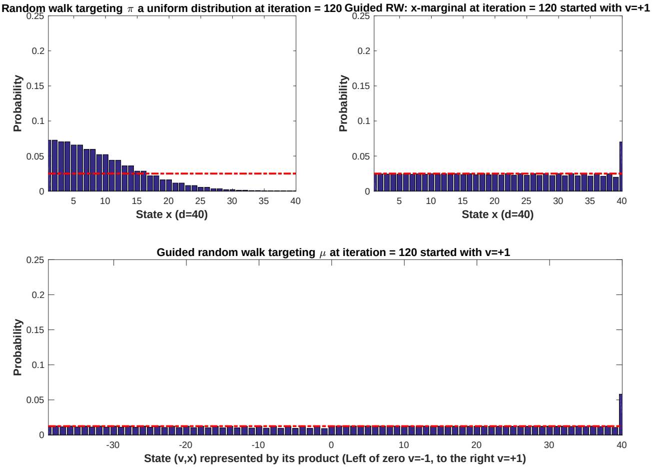
Reversible vs. nonreversible



Reversible vs. nonreversible



Reversible vs. nonreversible



(π, S)-reversibility

- It has been shown in [AL19] that many nonreversible Markov kernels fall under the same common framework of (π, S)-reversibility
- It encompasses the modified (or skew) detailed balance conditions.
- The notion of (π, S)-reversibility is based on the existence of an involution $s : \mathsf{Z} \to \mathsf{Z}$: $s \circ s = \mathrm{Id}$.
- Define the corresponding kernel $S(z, A) = \mathbb{1}_A(s(z)) = \delta_{s(z)}(A)$.

Definition 1: (π, S)-reversibility

P is (π, S)-reversible if it satisfies the condition,

$$\pi(\mathrm{d}z_0)P(z_0, \mathrm{d}z_1) = s_{\#}\pi(\mathrm{d}z_1)SPS(z_1, \mathrm{d}z_0) \; .$$

- In particular, if $s_{\#}\pi = \pi$ and P is a Markov kernel, then π is invariant for P.

(π, S)-reversibility

- It has been shown in [AL19] that many nonreversible Markov kernels fall under the same common framework of (π, S)-reversibility
- It encompasses the modified (or skew) detailed balance conditions.
- The notion of (π, S)-reversibility is based on the existence of an involution $s : \mathsf{Z} \to \mathsf{Z}$: $s \circ s = \mathrm{Id}$.
- Define the corresponding kernel $S(z, A) = \mathbb{1}_A(s(z)) = \delta_{s(z)}(A)$.

Definition 2: (π, S)-reversibility

P is (π, S)-reversible if it satisfies the condition,

$$\pi(\mathrm{d}z_0)P(z_0, \mathrm{d}z_1) = s_{\#}\pi(\mathrm{d}z_1)SPS(z_1, \mathrm{d}z_0)\ .$$

- In the previous example, $s : (x, v) \mapsto (x, -v)$, and $s_{\#}\mu = \mu$.

(π, S)-reversibility

- It has been shown in [AL19] that many nonreversible Markov kernels fall under the same common framework of (π, S)-reversibility
- It encompasses the modified (or skew) detailed balance conditions.
- The notion of (π, S)-reversibility is based on the existence of an involution $s : \mathsf{Z} \to \mathsf{Z}$: $s \circ s = \mathrm{Id}$.
- Define the corresponding kernel $S(z, A) = \mathbb{1}_A(s(z)) = \delta_{s(z)}(A)$.

Definition 3: (π, S)-reversibility

P is (π, S)-reversible if it satisfies the condition,

$$\pi(\mathrm{d}z_0)P(z_0, \mathrm{d}z_1) = s_{\#}\pi(\mathrm{d}z_1)SPS(z_1, \mathrm{d}z_0) \ .$$

- We assume that the condition $s_{\#}\pi = \pi$ is in force.

(π, S)-reversibility

- It has been shown in [AL19] that many nonreversible Markov kernels fall under the same common framework of (π, S)-reversibility
- It encompasses the modified (or skew) detailed balance conditions.
- The notion of (π, S)-reversibility is based on the existence of an involution $s : \mathsf{Z} \to \mathsf{Z}$: $s \circ s = \mathrm{Id}$.
- Define the corresponding kernel $S(z, A) = \mathbb{1}_A(s(z)) = \delta_{s(z)}(A)$.

Definition 4: (π, S)-reversibility

P is (π, S)-reversible if it satisfies the condition,

$$\pi(\mathrm{d}z_0) P(z_0, \mathrm{d}z_1) = s_{\#}\pi(\mathrm{d}z_1) SPS(z_1, \mathrm{d}z_0)\ .$$

- Note that for $s = \mathrm{Id}$ we recover the standard detailed balance condition.

The generalized Metropolis-Hastings recipe

Algorithm 4: the generalized Metropolis-Hastings (GMH) algorithm

- Input:
  - Initial state Z_0;
  - Proposal kernel Q on Z;
  - Acceptance probability $\alpha : \mathsf{Z}^2 \to [0, 1]$;
- at stage $k + 1$:
  - sample a proposal $Y_{k+1} \sim Q(Z_k, \cdot)$;
  - Set $Z_{k+1} = Y_{k+1}$ with probability $\alpha(Z_k, Y_{k+1})$;
  - Set $Z_{k+1} = s(Z_k)$ otherwise.

(π, S)-reversibility for the GMH kernel

- The Markov kernel associated with $(Z_k)_{k\in\mathbb{N}}$ is given for any $z \in \mathsf{Z}$ and $\mathsf{A} \in \mathcal{Z}$,

$$P(z_0, \mathsf{A}) = \int_{\mathsf{A}} \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) + \delta_{s(z_0)}(\mathsf{A}) \int_{\mathsf{Z}} \{1 - \alpha(z_0, z_1)\} Q(z_0, \mathrm{d}z_1) \; .$$

- General necessary and sufficient conditions on α and Q implying that P is (π, S)-reversible?
- As in the reversible case, we consider

$$Q_\alpha(z_0, \mathrm{d}z_1) = \alpha(z_0, z_1) Q(z_0, \mathrm{d}z_1) \; .$$

- As in the reversible case, P is (π, S)-reversible if Q_α is:

$$\pi(\mathrm{d}z) Q_\alpha(z, \mathrm{d}z') = s_\# \pi(\mathrm{d}z') S Q_\alpha S(z', \mathrm{d}z) \; .$$

- We establish sufficient and necessary conditions for Q_α to be (π, S)-reversible.

The generalized MH rule

- Recall that we considered in the reversible case the probability measures $\tilde{\nu}$ on Z^2:

$$\tilde{\nu}_Q(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z_0) \int_{\mathsf{B}} Q(z_0, \mathsf{B}) , \quad \tilde{\nu}_Q^s(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} \pi(\mathrm{d}z_0) \int_{\mathsf{A}} Q(z_0, \mathsf{B}) .$$

- Note that $\tilde{\nu}_Q^s$ is the pushforward measure of $\tilde{\nu}_Q$ by $(z_0, z_1) \mapsto (z_1, z_0)$ on Z^2.
- Here, we consider $\tilde{\nu}_Q$ still
- but instead,

$$\tilde{\nu}_Q^s \text{ is the pushforward measure of } \tilde{\nu}_Q \text{ by } F_s : (z_0, z_1) \mapsto (s(z_1), s(z_0)) \text{ on } \mathsf{Z}^2 . \quad (1)$$

- We now generalize [Tie94, Proposition 1 and Theorem 2] as follows:
  - Particular restrictions of $\tilde{\nu}_Q$ and $\tilde{\nu}_Q^s$ are equivalent.
  - Necessary and sufficient conditions depending on these restrictions.

The generalized MH recipe

- Consider $\tilde{\nu}$ on Z^2:

$$\tilde{\nu}_Q(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z_0) \int_{\mathsf{B}} Q(z_0, \mathsf{B}) \,.$$

- $\tilde{\nu}_Q^s$ is the pushforward measure of $\tilde{\nu}_Q$ by $F_s : (z_0, z_1) \mapsto (s(z_1), s(z_0))$ on Z^2.
- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}_Q^s$, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q^s \ll \tilde{\lambda}_Q$ and denote $h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q}$,

Proposition 1: Thin et al 2020

Set

$$A_Q = \{h_Q \times h_Q \circ F_s > 0\} \in \mathcal{Z}^{\otimes 2} \,.$$

Then, the restrictions

- $\tilde{\nu}_A(\cdot) = \tilde{\nu}(\cdot \cap A_Q)$ and $\tilde{\nu}_A^s(\cdot) = \tilde{\nu}^s(\cdot \cap A_Q)$ are equivalent;
- $\tilde{\nu}_{A,\mathrm{c}}(\cdot) = \tilde{\nu}(\cdot \cap A_Q^{\mathrm{c}})$ and $\tilde{\nu}_{A,\mathrm{c}}^s(\cdot) = \tilde{\nu}^s(\cdot \cap A_Q^{\mathrm{c}})$ are mutually singular.

The generalized MH recipe

- Consider $\tilde{\nu}$ on Z^2:

$$\tilde{\nu}_Q(\mathsf{A} \times \mathsf{B}) = \int_{\mathsf{A}} \pi(\mathrm{d}z_0) \int_{\mathsf{B}} Q(z_0, \mathsf{B}) \, .$$

- $\tilde{\nu}_Q^s$ is the pushforward measure of $\tilde{\nu}_Q$ by $F_s : (z_0, z_1) \mapsto (s(z_1), s(z_0))$ on Z^2.
- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}_Q^s$, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q^s \ll \tilde{\lambda}_Q$ and denote $h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q}$,

Proposition 2: Thin et al 2020

Set

$$A_Q = \{h_Q \times h_Q \circ F_s > 0\} \in \mathcal{Z}^{\otimes 2} \, .$$

Define, for $(z_0, z_1) \in A_Q$,

$$r_Q(z_0, z_1) = h_Q(z_0, z_1) / h_Q \circ F_s(z_0, z_1) \, .$$

Then, r_Q is a version of the density of $\tilde{\nu}_A$ w.r.t. $\tilde{\nu}_A^s$, *i.e.* $r = \mathrm{d}\tilde{\nu}_A / \mathrm{d}\tilde{\nu}_A^s$.

The generalized MH recipe

- Based on the previous results, we can give a necessary and sufficient condition on α so that P is (π, S)-reversible.
- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}_Q^s$, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q^s \ll \tilde{\lambda}_Q$ and denote $h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q}$.
- Set $A_Q = \{h_Q \times h_Q \circ F_s > 0\} \in \mathcal{Z}^{\otimes 2}$,

$$r_Q(z_0, z_1) = h_Q(z_0, z_1) / h_Q \circ F_s(z_0, z_1) \,, (z_0, z_1) \in A_Q \,.$$

Theorem 1: Thin et al 2020

The sub-Markovian kernel Q_α is (π, S)-reversible if and only if the following conditions hold.

- The function α is zero $\tilde{\nu}_Q$-a.e.on A_Q^{c}.
- The function α satisfies $\alpha(z_0, z_1) r_Q(z_0, z_1) = \alpha(s(z'), s(z))$ $\tilde{\nu}_Q$-a.e.on A_Q.

The generalized MH recipe

- $\tilde{\lambda}_Q = \tilde{\nu}_Q + \tilde{\nu}_Q^s$, then, then, $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and $\tilde{\nu}_Q \ll \tilde{\lambda}_Q$ and denote $h_Q = \frac{\mathrm{d}\tilde{\nu}_Q}{\mathrm{d}\tilde{\lambda}_Q}$.
- Set $A_Q = \{h_Q \times h_Q \circ F_s > 0\} \in \mathcal{Z}^{\otimes 2}$ and

$$r_Q(z_0, z_1) = h_Q(z_0, z_1)/h_Q \circ F_s\big(z_0, z_1\big)\ , (z_0, z_1) \in A_Q\ .$$

- We can then define the Metropolis-Hastings rejection probability by

$$\alpha(z_0, z_1) = \begin{cases} \mathrm{a}\left(\frac{h_Q \circ F_s(z_0, z_1)}{h_Q(z_0, z_1)}\right) = \mathrm{a}\left(1/r_Q(z_0, z_1)\right) & h_Q(z_0, z_1) \neq 0\ , \\ 1 & h_Q(z_0, z_1) = 0\ , \end{cases}$$

  where $\mathrm{a} : \mathbb{R}_+ \to [0, 1]$ satisfies $\mathrm{a}(0) = 0,\ t\mathrm{a}(1/t) = \mathrm{a}(t)$.
- This choice of α ensures (π, S)-reversibility for Q_α and therefore for the MH kernel P.

the generalized MH kernel: $\exists$ common dominating measure

- If π and Q have densities with respect to μ:

$$h_Q(z_0, z_1) = \pi(z_0) q(z_0, z_1) \ ,$$
$$h_Q \circ F_s(z_0, z_1) = \pi(s(z_1)) q(s(z_1), s(z_0)) \ .$$

and therefore:

$$\alpha(z_0, z_1) = \begin{cases} \mathrm{a}\left[\frac{\pi(s(z_1))q(s(z_1),s(z_0))}{\pi(z_0)q(z_0,z_1)}\right] & \pi(z_0)q(z_0,z_1) \neq 0, \\ 1 & \pi(z_0)q(z_0,z_1) = 0 \ , \end{cases}$$

where $\mathrm{a} : \mathbb{R}_+ \to [0,1]$ satisfies $\mathrm{a}(0) = 0$, $t\mathrm{a}(1/t) = \mathrm{a}(t)$.

The generalized MH kernel: deterministic proposal

- Suppose now that $\Phi : \mathsf{Z} \to \mathsf{Z}$ satisfying

$$\Phi^{-1} = s \circ \Phi \circ s \, .$$

- $s = \mathrm{Id}$, Φ is an involution.
- We consider the deterministic proposal kernel

$$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)\text{: when the current state is } z_0\text{, the proposal is } \Phi(z_0) \quad .$$

The generalized MH kernel: deterministic proposal

- Suppose now that $\Phi : \mathsf{Z} \to \mathsf{Z}$ satisfying

$$\Phi^{-1} = s \circ \Phi \circ s \ .$$

- We consider the deterministic proposal kernel

$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)$: when the current state is z_0, the proposal is $\Phi(z_0)$.

- In this scenario we have that

$\tilde{\nu}\big(\mathrm{d}(z_0, z_1)\big) = \pi(\mathrm{d}z_0)\delta_{\Phi(z_0)}(\mathrm{d}z_1)$ and $\tilde{\nu}^s\big(\mathrm{d}(z_0, z_1)\big) = \pi(\mathrm{d}z_1)\delta_{\Phi^{-1}(z_1)}(\mathrm{d}z_0)$.

The generalized MH kernel: deterministic proposal

- Suppose now that $\Phi : \mathsf{Z} \to \mathsf{Z}$ satisfying

$$\Phi^{-1} = s \circ \Phi \circ s \, .$$

- We consider the deterministic proposal kernel

$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)$: when the current state is z_0, the proposal is $\Phi(z_0)$.

- The function h_Q is given by

$$h_Q(z_0, z_1) = \mathbb{1}_{\Phi(z_0)}(z_1) k(z_0) \text{ with } \quad k(z) = \frac{\mathrm{d}\pi}{\mathrm{d}\lambda}(z), \quad \lambda = \pi + (\Phi^{-1})_{\#}\pi \, ,$$
$$h_Q^s(z_0, z_1) = \mathbb{1}_{\Phi^{-1}(s(z_1))}(s(z_0)) k(s(z_1)) = \mathbb{1}_{\Phi(z_0)}(z_1) k[\Phi(s(z_0))] \, .$$

- Therefore, $\alpha\big(z, \Phi(z)\big) = \bar{\alpha}(z)$ with

$$\bar{\alpha}(z) = \begin{cases} \mathrm{a}\left(\frac{k\left(\Phi(z)\right)}{k(z)}\right) & \text{if } k(z) > 0 \, , \\ 1 & \text{otherwise} \quad . \end{cases}$$

Of course, there is no need to define $\alpha(z_0, z_1)$ for $z_1 \neq \Phi(z_0)$.

- Computation of k?

The generalized MH kernel: deterministic proposal

- Suppose now that $\Phi : \mathsf{Z} \to \mathsf{Z}$ satisfying

$$\Phi^{-1} = s \circ \Phi \circ s \; .$$

- We consider the deterministic proposal kernel

$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)$: when the current state is z_0, the proposal is $\Phi(z_0)$.

- A special case of interest is when $\mathsf{Z} = \mathbb{R}^d$ and $\pi(\mathrm{d}z) = \pi(z)\,\mathrm{Leb}_d(\mathrm{d}z)$.
- Here the dominating measure $\tilde{\lambda}$ is given by

$$\tilde{\lambda}(\mathrm{d}z) = \pi + (\Phi^{-1})_{\#}\pi = \{\pi(z) + \pi \circ \Phi(z)\mathrm{Jac}_{\Phi}(z)\}\,\mathrm{Leb}_d(\mathrm{d}z) \; ,$$

where Jac_f denotes the absolute value of the Jacobian determinant of f.

The generalized MH kernel: deterministic proposal

- Suppose now that $\Phi : \mathsf{Z} \to \mathsf{Z}$ satisfying

$$\Phi^{-1} = s \circ \Phi \circ s \,.$$

- We consider the deterministic proposal kernel

$Q(z_0, \mathrm{d}z_1) = \delta_{\Phi(z_0)}(\mathrm{d}z_1)$: when the current state is z_0, the proposal is $\Phi(z_0)$.

- A special case of interest is when $\mathsf{Z} = \mathbb{R}^d$ and $\pi(\mathrm{d}z) = \pi(z)\,\mathrm{Leb}_d(\mathrm{d}z)$.
- Then, the density $k(z)$ is given by

$$k(z) = \frac{\mathrm{d}\pi}{\mathrm{d}\tilde{\lambda}}(z) = \frac{\pi(z)}{\pi(z) + \pi \circ \Phi(z) \mathrm{Jac}_{\Phi}(z)} \,.$$

- The acceptance ratio $\bar{\alpha}(z)$ takes the simple form

$$\bar{\alpha}(z) = \begin{cases} \mathrm{a}\left(\frac{\pi \circ \Phi(z) \mathrm{Jac}_{\Phi}(z)}{\pi(z)}\right) \,, & \text{if } \pi(z) \neq 0 \\ 1 & \text{otherwise} \,. \end{cases}$$

Thank you for your attention. Any questions ?

Bibliography I

Christophe Andrieu and Samuel Livingstone. "Peskun-Tierney ordering for Markov chain and process Monte Carlo: beyond the reversible scenario". In: *Accepted for publication in Ann. Statist.* (2019).

Paul Gustafson. "A guided walk Metropolis algorithm". In: *Statistics and computing* 8.4 (1998), pp. 357–364.

A. M. Horowitz. "A generalized guided Monte Carlo algorithm". In: *Physics Letters B* 268.2 (1991), pp. 247–252.

Kirill Neklyudov, Max Welling, Evgenii Egorov, and Dmitry Vetrov. "Involutive MCMC: a Unifying Framework". In: *arXiv preprint arXiv:2006.16653* (2020).

L. Tierney. "Markov chains for exploring posterior disiributions (with discussion)". In: *Ann. Statist.* 22.4 (1994), pp. 1701–1762.

L. Tierney. "A note on Metropolis-Hastings kernels for general state spaces". In: *Ann. Appl. Probab.* 8 (1998), pp. 1–9.