Unbiasing HMC algorithms for general Hamiltonian functions

Régis SANTET

(CERMICS, École des Ponts & MATHERIALS Team, Inria Paris)

Joint work with T. Lelièvre, G. Stoltz

Hamiltonian Monte Carlo - 1

Target measure: $\pi(\mathrm{d}q) = Z_\pi^{-1} \mathrm{e}^{-V(q)} \mathrm{d}q$

Hamiltonian Monte Carlo - 1

Target measure: $\pi(\mathrm{d}q) = Z_\pi^{-1} \mathrm{e}^{-V(q)} \mathrm{d}q$

Augmented space: Boltzmann–Gibbs measure $\mu(\mathrm{d}q\,\mathrm{d}p) = Z_\mu^{-1} \mathrm{e}^{-H(q,p)} \mathrm{d}q\,\mathrm{d}p$
with $H(q,p) = V(q) + \frac{|p|^2}{2}$

$$\begin{cases} \tilde{p}^0 \sim Z_p^{-1} \mathrm{e}^{-|p|^2/2} \mathrm{d}p \\ (\tilde{q}^1, \tilde{p}^1) = \psi_{\Delta t}(q^0, \tilde{p}^0) \\ \text{If } U^n \leqslant \min\left(1, \mathrm{e}^{-H(\tilde{q}^1,\tilde{p}^1)+H(q^0,\tilde{p}^0)}\right) \\ \quad \text{accept the proposal: } (q^1, p^1) = (\tilde{q}^1, \tilde{p}^1) \\ \quad \text{else reject the proposal: } (q^1, p^1) = (q^0, \tilde{p}^0) \end{cases}$$

where $\psi_{\Delta t} = S \circ \varphi_{\Delta t}$ with $S(q,p) = (q,-p)$ and $\varphi_{\Delta t}$ is one step of the Störmer–Verlet scheme:

$$\begin{cases} \tilde{p}^{1/2} = \tilde{p}^0 - \dfrac{\Delta t}{2} \nabla V(q^0) \\ \tilde{q}^1 = q^0 + \Delta t\, \tilde{p}^{1/2} \\ \tilde{p}^1 = \tilde{p}^{1/2} - \dfrac{\Delta t}{2} \nabla V(\tilde{q}^1) \end{cases}$$

Hamiltonian Monte Carlo - 2

Main result: HMC leaves μ (and therefore π) invariant

Proof: the map $\varphi_{\Delta t} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$
- preserves the Lebesgue measure $\mathrm{d}q\,\mathrm{d}p \leftarrow$ **symplecticity**

$$\forall (q,p) \in \mathbb{R}^d \times \mathbb{R}^d, \qquad \nabla\varphi_{\Delta t}(q,p)^\mathsf{T} J \nabla\varphi_{\Delta t}(q,p) = J, \qquad J = \begin{pmatrix} 0_d & \mathrm{I}_d \\ -\mathrm{I}_d & 0_d \end{pmatrix}$$

Hamiltonian Monte Carlo - 2

Main result: HMC leaves μ (and therefore π) invariant

Proof: the map $\varphi_{\Delta t} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$

- preserves the Lebesgue measure $\mathrm{d}q\,\mathrm{d}p \leftarrow$ **symplecticity**

$$\forall (q,p) \in \mathbb{R}^d \times \mathbb{R}^d, \qquad \nabla\varphi_{\Delta t}(q,p)^\mathsf{T} J \nabla\varphi_{\Delta t}(q,p) = J, \qquad J = \begin{pmatrix} 0_d & \mathrm{I}_d \\ -\mathrm{I}_d & 0_d \end{pmatrix}$$

- is **S-reversible** on the whole configurational space

$$\forall (q,p) \in \mathbb{R}^d \times \mathbb{R}^d, \qquad \psi_{\Delta t} \circ \psi_{\Delta t} = \mathrm{id}_{\mathbb{R}^d \times \mathbb{R}^d}$$

Thus the Metropolis–Hastings ratio is

$$\frac{\mu(\mathrm{d}q'\,\mathrm{d}p')\delta_{\psi_{\Delta t}(q',p')}(\mathrm{d}q\,\mathrm{d}p)}{\mu(\mathrm{d}q\,\mathrm{d}p)\delta_{\psi_{\Delta t}(q,p)}(\mathrm{d}q'\,\mathrm{d}p')} = \mathrm{e}^{-H(q',p')+H(q,p)}$$

Hamiltonian Monte Carlo - 2

Main result: HMC leaves μ (and therefore π) invariant

Proof: the map $\varphi_{\Delta t} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$

- preserves the Lebesgue measure $\mathrm{d}q\,\mathrm{d}p \leftarrow$ **symplecticity**

$$\forall (q,p) \in \mathbb{R}^d \times \mathbb{R}^d, \qquad \nabla \varphi_{\Delta t}(q,p)^\mathsf{T} J \nabla \varphi_{\Delta t}(q,p) = J, \qquad J = \begin{pmatrix} 0_d & \mathrm{I}_d \\ -\mathrm{I}_d & 0_d \end{pmatrix}$$

- is **S-reversible** on the whole configurational space

$$\forall (q,p) \in \mathbb{R}^d \times \mathbb{R}^d, \qquad \psi_{\Delta t} \circ \psi_{\Delta t} = \mathrm{id}_{\mathbb{R}^d \times \mathbb{R}^d}$$

Thus the Metropolis–Hastings ratio is

$$\frac{\mu(\mathrm{d}q'\,\mathrm{d}p')\delta_{\psi_{\Delta t}(q',p')}(\mathrm{d}q\,\mathrm{d}p)}{\mu(\mathrm{d}q\,\mathrm{d}p)\delta_{\psi_{\Delta t}(q,p)}(\mathrm{d}q'\,\mathrm{d}p')} = \mathrm{e}^{-H(q',p')+H(q,p)}$$

What if H is another (nonseparable) Hamiltonian function ?

Motivation

When does a nonseparable Hamiltonian function appear ?

- Molecular dynamics written in internal coordinates[1]
- Shadow HMC[2]
- Riemannian Manifold HMC[3]

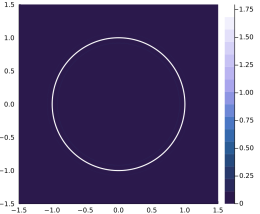
$$H(q,p) = V(q) - \frac{1}{2}\ln\det D(q) + \frac{1}{2}p^\mathsf{T} D(q) p$$

where $D(q)$ is a position-dependent symmetric positive definite matrix

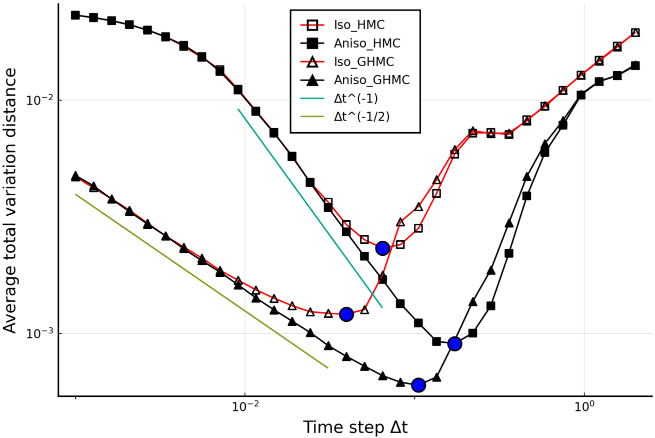
Note that $\int \mathrm{e}^{-H(q,p)} dp \propto \mathrm{e}^{-V} \propto \pi$

[1] Hairer/Lubich/Wanner (2006), Fang *et al* (2014)
[2] Izaguirre/Hampton (2004)
[3] Girolami/Calderhead (2011)

Riemannian Manifold HMC



Target pdf



$$D_{\text{iso}} \equiv (1+\varepsilon)\mathrm{I}_2 \text{ or } D_{\text{aniso}}(x,y) = \varepsilon \mathrm{I}_2 + e_\theta \otimes e_\theta,\ e_\theta = \frac{(-y\ \ x)^\mathsf{T}}{\sqrt{x^2+y^2}},$$

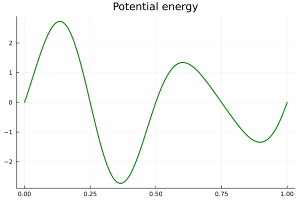
Setting: $V(x,y) = 100(x^2+y^2-1)$, 10^6 iterations of (G)HMC with reversibility checks. TV norm computed with respect to the uniform distribution in the angle θ, $\varepsilon = 0.1$

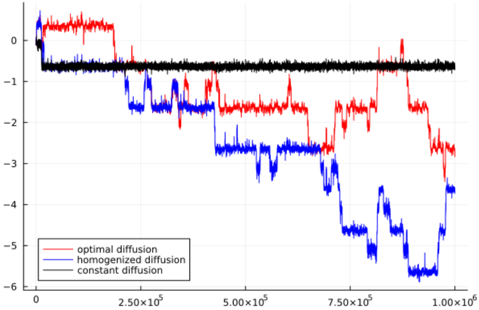
RMHMC and overdamped Langevin

Yields a consistent discretization of the overdamped Langevin dynamics with a position-dependent diffusion[4]: $D(q_t) \in \mathcal{S}_d^{++}$

$$\mathrm{d}q_t = (-D(q_t)\nabla V(q_t) + \mathrm{div} D(q_t))dt + \sqrt{2\beta^{-1}} D(q_t)^{1/2} \mathrm{d}W_t$$

Optimal diffusion, optimal diffusion in the homogenized limit, constant diffusion (normalized in L^2)[5]





Using a position dependent mass/metric/diffusion accelerates the sampling[6]

[4] Lelièvre/RS/Stoltz (2023)
[5] Lelièvre/Pavliotis/Robin/RS/Stoltz (WIP)
[6] Roberts/Stramer (2002), Girolami/Calderhead (2011), Bou-Rabee *et al* (2014)

Nonseparable HMC

Same algorithm except that $\varphi_{\Delta t}$ is one step of the Generalized Störmer–Verlet (GSV) scheme:

$$\begin{cases} \tilde{p}^{1/2} = p^n - \frac{\Delta t}{2} \nabla_q H(q^0, \tilde{p}^{1/2}) \\ q^1 = q^0 + \frac{\Delta t}{2} \Big(\nabla_p H(q^0, \tilde{p}^{1/2}) + \nabla_p H(q^1, \tilde{p}^{1/2}) \Big) \\ p^1 = \tilde{p}^{1/2} - \frac{\Delta t}{2} \nabla_q H(q^1, \tilde{p}^{1/2}) \end{cases}$$

For a nonseparable Hamiltonian function, the scheme is implicit

Nonseparable HMC

Same algorithm except that $\varphi_{\Delta t}$ is one step of the Generalized Störmer–Verlet (GSV) scheme:

$$\begin{cases} \tilde{p}^{1/2} = p^n - \dfrac{\Delta t}{2} \nabla_q H(q^0, \tilde{p}^{1/2}) \\ q^1 = q^0 + \dfrac{\Delta t}{2} \Big(\nabla_p H(q^0, \tilde{p}^{1/2}) + \nabla_p H(q^1, \tilde{p}^{1/2}) \Big) \\ p^1 = \tilde{p}^{1/2} - \dfrac{\Delta t}{2} \nabla_q H(q^1, \tilde{p}^{1/2}) \end{cases}$$

For a nonseparable Hamiltonian function, the scheme is implicit

For sufficiently small Δt, $\varphi_{\Delta t}$ is again symplectic and S-reversible[7]

[7] Hairer/Lubich/Wanner (2006)

Nonseparable HMC

Same algorithm except that $\varphi_{\Delta t}$ is one step of the Generalized Störmer–Verlet (GSV) scheme:

$$\begin{cases} \tilde{p}^{1/2} = p^n - \frac{\Delta t}{2}\nabla_q H(q^0, \tilde{p}^{1/2}) \\ q^1 = q^0 + \frac{\Delta t}{2}\left(\nabla_p H(q^0, \tilde{p}^{1/2}) + \nabla_p H(q^1, \tilde{p}^{1/2})\right) \\ p^1 = \tilde{p}^{1/2} - \frac{\Delta t}{2}\nabla_q H(q^1, \tilde{p}^{1/2}) \end{cases}$$

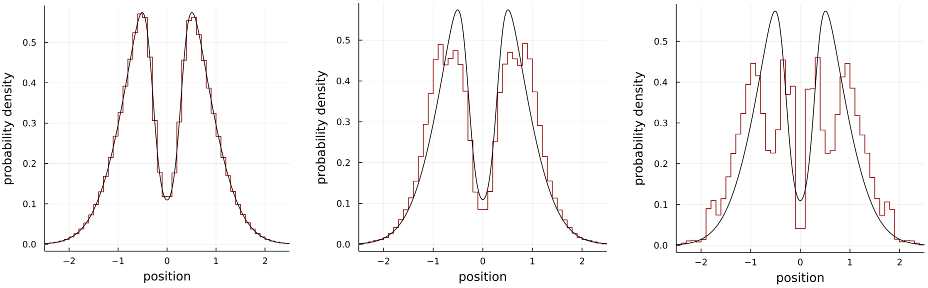
For a nonseparable Hamiltonian function, the scheme is implicit

For sufficiently small Δt, $\varphi_{\Delta t}$ is again symplectic and S-reversible[7]

But what happens for larger Δt ?

[7] Hairer/Lubich/Wanner (2006)

Bias with standard implementation



$\Delta t = 0.18$ $\Delta t = 0.69$ $\Delta t = 1.08$

Black solid line: target pdf. Red histograms: sampled stationary distribution

Setting: 1d double-well potential, $V_{\sigma,h}(q) = q^2 - 1 + \frac{h}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{q^2}{2\sigma^2}\right)$ with $\sigma = 0.2$ and $h = 1$,

diffusion $D(q) = \left(\frac{1.5+\cos(\pi q)}{2}\right)^2$. Sampling with 10^7 iterations of GHMC ($\gamma = 1$), and rejection if a solution is not found for the forward move solved by Newton.

Numerical flow

• Theoretically built using the Implicit Function theorem: GSV numerical scheme is written as

$$\Phi_{\Delta t}(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1) = \begin{pmatrix} q^{1/2} - q^0 - \frac{\Delta t}{2} \nabla_p H(q^0, p^{1/2}) \\ p^{1/2} - p^0 - \frac{\Delta t}{2} \nabla_p H(p^0, p^{1/2}) \\ q^1 - q^{1/2} - \frac{\Delta t}{2} \nabla_p H(q^1, p^{1/2}) \\ p^1 - p^{1/2} - \frac{\Delta t}{2} \nabla_p H(p^1, p^{1/2}) \end{pmatrix}$$

IFT assumption to define $\varphi_{\Delta t}$ on an open set $\mathcal{A}_{\Delta t}$:

$$(q^0, p^0) \in \mathcal{A}_{\Delta t} \Leftrightarrow \nabla_{(q^{1/2}, p^{1/2}, q^1, p^1)} \Phi_{\Delta t}(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1) \text{ is invertible}$$

Numerical flow

• Theoretically built using the Implicit Function theorem: GSV numerical scheme is written as

$$\Phi_{\Delta t}(q^0,p^0,q^{1/2},p^{1/2},q^1,p^1) = \begin{pmatrix} q^{1/2} - q^0 - \frac{\Delta t}{2}\nabla_p H(q^0,p^{1/2}) \\ p^{1/2} - p^0 - \frac{\Delta t}{2}\nabla_p H(p^0,p^{1/2}) \\ q^1 - q^{1/2} - \frac{\Delta t}{2}\nabla_p H(q^1,p^{1/2}) \\ p^1 - p^{1/2} - \frac{\Delta t}{2}\nabla_p H(p^1,p^{1/2}) \end{pmatrix}$$

IFT assumption to define $\varphi_{\Delta t}$ on an open set $\mathcal{A}_{\Delta t}$:

$$(q^0,p^0) \in \mathcal{A}_{\Delta t} \Leftrightarrow \nabla_{(q^{1/2},p^{1/2},q^1,p^1)}\Phi_{\Delta t}(q^0,p^0,q^{1/2},p^{1/2},q^1,p^1) \text{ is invertible}$$

• $\varphi_{\Delta t}$ is $\mathcal{C}^1$ on $\mathcal{A}_{\Delta t}$ and $|\det \nabla\varphi_{\Delta t}| = 1$ (local preservation of the Lebesgue measure)

Numerical flow

• Theoretically built using the Implicit Function theorem: GSV numerical scheme is written as

$$\Phi_{\Delta t}(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1) = \begin{pmatrix} q^{1/2} - q^0 - \frac{\Delta t}{2}\nabla_p H(q^0, p^{1/2}) \\ p^{1/2} - p^0 - \frac{\Delta t}{2}\nabla_p H(p^0, p^{1/2}) \\ q^1 - q^{1/2} - \frac{\Delta t}{2}\nabla_p H(q^1, p^{1/2}) \\ p^1 - p^{1/2} - \frac{\Delta t}{2}\nabla_p H(p^1, p^{1/2}) \end{pmatrix}$$

IFT assumption to define $\varphi_{\Delta t}$ on an open set $\mathcal{A}_{\Delta t}$:

$$(q^0, p^0) \in \mathcal{A}_{\Delta t} \Leftrightarrow \nabla_{(q^{1/2}, p^{1/2}, q^1, p^1)} \Phi_{\Delta t}(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1) \text{ is invertible}$$

• $\varphi_{\Delta t}$ is $\mathcal{C}^1$ on $\mathcal{A}_{\Delta t}$ and $|\det \nabla \varphi_{\Delta t}| = 1$ (local preservation of the Lebesgue measure)

• S-reversibility for $\Phi_{\Delta t}$

$$\Phi_{\Delta t}\left(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1\right) = 0 \Leftrightarrow \Phi_{\Delta t}\left(S(q^1, p^1), S(q^{1/2}, p^{1/2}), S(q^0, p^0)\right) = 0$$

Numerical flow

• Theoretically built using the Implicit Function theorem: GSV numerical scheme is written as

$$\Phi_{\Delta t}(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1) = \begin{pmatrix} q^{1/2} - q^0 - \frac{\Delta t}{2} \nabla_p H(q^0, p^{1/2}) \\ p^{1/2} - p^0 - \frac{\Delta t}{2} \nabla_p H(p^0, p^{1/2}) \\ q^1 - q^{1/2} - \frac{\Delta t}{2} \nabla_p H(q^1, p^{1/2}) \\ p^1 - p^{1/2} - \frac{\Delta t}{2} \nabla_p H(p^1, p^{1/2}) \end{pmatrix}$$

IFT assumption to define $\varphi_{\Delta t}$ on an open set $\mathcal{A}_{\Delta t}$:

$$(q^0, p^0) \in \mathcal{A}_{\Delta t} \Leftrightarrow \nabla_{(q^{1/2}, p^{1/2}, q^1, p^1)} \Phi_{\Delta t}(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1) \text{ is invertible}$$

• $\varphi_{\Delta t}$ is $\mathcal{C}^1$ on $\mathcal{A}_{\Delta t}$ and $|\det \nabla \varphi_{\Delta t}| = 1$ (local preservation of the Lebesgue measure)

• S-reversibility for $\Phi_{\Delta t}$

$$\Phi_{\Delta t}\left(q^0, p^0, q^{1/2}, p^{1/2}, q^1, p^1\right) = 0 \Leftrightarrow \Phi_{\Delta t}\left(S(q^1, p^1), S(q^{1/2}, p^{1/2}), S(q^0, p^0)\right) = 0$$

$\psi_{\Delta t} = S \circ \varphi_{\Delta t}$ is defined on $\mathcal{A}_{\Delta t}$, but $\mathcal{A}_{\Delta t} \neq \mathbb{R}^d \times \mathbb{R}^d$
$\rightarrow$ S-reversibility for $\Phi_{\Delta t} \not\Rightarrow$ S-reversibility for $\varphi_{\Delta t}$ *i.e.* $\psi_{\Delta t} \circ \psi_{\Delta t} = \mathrm{id}$

Reversibility check

• In practice, one numerically builds $\varphi_{\Delta t}$ using Newton's method to find a solution to the implicit problem
$\rightarrow$ Need to extend $\psi_{\Delta t} = S \circ \varphi_{\Delta t}$ to the whole configurational space while still satisfying the two fundamental properties (Lebesgue-preserving, S-reversibility)

Reversibility check

• In practice, one numerically builds $\varphi_{\Delta t}$ using Newton's method to find a solution to the implicit problem
$\rightarrow$ Need to extend $\psi_{\Delta t} = S \circ \varphi_{\Delta t}$ to the whole configurational space while still satisfying the two fundamental properties (Lebesgue-preserving, S-reversibility)

• Numerical flow with reversibility check[8]

$$\psi_{\Delta t}^{\mathrm{rev}} = \psi_{\Delta t} \mathbf{1}_{\mathcal{B}_{\Delta t}} + \mathrm{id} \mathbf{1}_{\mathcal{B}_{\Delta t}^c}$$

where

$$\mathcal{B}_{\Delta t} = \{(q,p) \in \mathcal{A}_{\Delta t} \text{ s.t. } \psi_{\Delta t}(q,p) \in \mathcal{A}_{\Delta t},\ \psi_{\Delta t} \circ \psi_{\Delta t}(q,p) = (q,p)\}$$

[8] Goodman/Holmes-Cerfon/Zappa (2017)

Reversibility check

• In practice, one numerically builds $\varphi_{\Delta t}$ using Newton's method to find a solution to the implicit problem
$\rightarrow$ Need to extend $\psi_{\Delta t} = S \circ \varphi_{\Delta t}$ to the whole configurational space while still satisfying the two fundamental properties (Lebesgue-preserving, S-reversibility)

• Numerical flow with reversibility check[8]

$$\psi_{\Delta t}^{\mathrm{rev}} = \psi_{\Delta t} \mathbf{1}_{\mathcal{B}_{\Delta t}} + \mathrm{id} \mathbf{1}_{\mathcal{B}_{\Delta t}^c}$$

where

$$\mathcal{B}_{\Delta t} = \{(q,p) \in \mathcal{A}_{\Delta t} \text{ s.t. } \psi_{\Delta t}(q,p) \in \mathcal{A}_{\Delta t},\ \psi_{\Delta t} \circ \psi_{\Delta t}(q,p) = (q,p)\}$$

Proposition (Lelièvre/RS/Stoltz (2023))

$\mathcal{B}_{\Delta t}$ is an open set. The map $\psi_{\Delta t}^{\mathrm{rev}}$ is globally well-defined, preserves the Lebesgue measure and is an involution on the whole configurational space.

As a corollary, HMC implemented with $\psi_{\Delta t}^{\mathrm{rev}}$ preserves the measure μ.

[8] Goodman/Holmes-Cerfon/Zappa (2017)

HMC with reversibility check

$H(q,p)$ general Hamiltonian function

$$\begin{cases} \tilde{p}^0 \sim Z_p^{-1} \mathrm{e}^{-H(q^0,p)} \mathrm{d}p \\ (\tilde{q}^1, \tilde{p}^1) = \psi_{\Delta t}^{\mathrm{rev}}(q^0, \tilde{p}^0) \\ \text{If } U^n \leqslant \min\left(1, \mathrm{e}^{-H(\tilde{q}^1,\tilde{p}^1)+H(q^0,\tilde{p}^0)}\right) \\ \quad \text{accept the proposal: } (q^1,p^1) = (\tilde{q}^1, \tilde{p}^1) \\ \quad \text{else reject the proposal: } (q^1,p^1) = (q^0, \tilde{p}^0) \end{cases}$$

HMC with reversibility check

$H(q,p)$ general Hamiltonian function

$$\begin{cases} \tilde{p}^0 \sim Z_p^{-1} \mathrm{e}^{-H(q^0,p)} \mathrm{d}p \\ (\tilde{q}^1, \tilde{p}^1) = \psi_{\Delta t}^{\mathrm{rev}}(q^0, \tilde{p}^0) \\ \text{If } U^n \leqslant \min\left(1, \mathrm{e}^{-H(\tilde{q}^1,\tilde{p}^1)+H(q^0,\tilde{p}^0)}\right) \\ \quad \text{accept the proposal: } (q^1, p^1) = (\tilde{q}^1, \tilde{p}^1) \\ \quad \text{else reject the proposal: } (q^1, p^1) = (q^0, \tilde{p}^0) \end{cases}$$

$\leftarrow \mathcal{N}(0, D(q^0)^{-1})$ for RMHMC

HMC with reversibility check

$H(q,p)$ general Hamiltonian function

$$\begin{cases} \tilde{p}^0 \sim Z_p^{-1} \mathrm{e}^{-H(q^0,p)} \mathrm{d}p \\ (\tilde{q}^1, \tilde{p}^1) = \psi^{\mathrm{rev}}_{\Delta t}(q^0, \tilde{p}^0) \\ \text{If } U^n \leqslant \min\left(1, \mathrm{e}^{-H(\tilde{q}^1,\tilde{p}^1)+H(q^0,\tilde{p}^0)}\right) \\ \quad \text{accept the proposal: } (q^1,p^1) = (\tilde{q}^1,\tilde{p}^1) \\ \quad \text{else reject the proposal: } (q^1,p^1) = (q^0,\tilde{p}^0) \end{cases}$$

$\leftarrow \mathcal{N}(0, D(q^0)^{-1})$ for RMHMC

Generalized HMC[9]: partial refreshment of the momenta, using a Strang splitting scheme of the Langevin dynamics

$$\begin{cases} \mathrm{d}q_t = \nabla_p H(q_t,p_t)\mathrm{d}t \\ \mathrm{d}p_t = -\nabla_q H(q_t,p_t)\mathrm{d}t - \gamma \nabla_p H(q_t,p_t)\mathrm{d}t + \sqrt{2\gamma}\mathrm{d}W_t \end{cases}$$

[9] Horowitz (1991)

GHMC with reversibility check

$H(q,p)$ general Hamiltonian function

$$\begin{cases} \tilde{p}^0 \sim T^{\mathrm{FD}}_{\Delta t/2}(q^0, p^0; \mathrm{d}p) \\ (\tilde{q}^1, \tilde{p}^1) = \psi^{\mathrm{rev}}_{\Delta t}(q^0, \tilde{p}^0) \\ \text{If } U^n \leqslant \min\left(1, \mathrm{e}^{-H(\tilde{q}^1,\tilde{p}^1)+H(q^0,\tilde{p}^0)}\right) \\ \quad \text{accept the proposal: } (q^1, \bar{p}^1) = (\tilde{q}^1, \tilde{p}^1) \\ \quad \text{else reject the proposal: } (q^1, \bar{p}^1) = (q^0, \tilde{p}^0) \\ \hat{p}^1 = -\bar{p}^1 \\ p^1 \sim T^{\mathrm{FD}}_{\Delta t/2}(q^1, \hat{p}^1; \mathrm{d}p) \end{cases}$$

where $T^{\mathrm{FD}}_{\Delta t/2}(q^0, p^0; \mathrm{d}p)$ is a consistent discretization of the fluctuation dissipation dynamics over a time interval of length $\Delta t/2$:

$$\begin{cases} \mathrm{d}p_t = -\gamma \nabla_p H(q^0, p_t)\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}W_t \\ p_0 = p^0 \end{cases}$$

which leaves μ invariant (mid-point, MALA, ...)

GHMC with reversibility check

$H(q,p)$ general Hamiltonian function

$$\begin{cases} \tilde{p}^0 \sim T^{\mathrm{FD}}_{\Delta t/2}(q^0, p^0; \mathrm{d}p) \\ (\tilde{q}^1, \tilde{p}^1) = \psi^{\mathrm{rev}}_{\Delta t}(q^0, \tilde{p}^0) \\ \text{If } U^n \leqslant \min\left(1, \mathrm{e}^{-H(\tilde{q}^1,\tilde{p}^1)+H(q^0,\tilde{p}^0)}\right) \\ \quad \text{accept the proposal: } (q^1, \bar{p}^1) = (\tilde{q}^1, \tilde{p}^1) \\ \quad \text{else reject the proposal: } (q^1, \bar{p}^1) = (q^0, \tilde{p}^0) \\ \hat{p}^1 = -\bar{p}^1 \\ p^1 \sim T^{\mathrm{FD}}_{\Delta t/2}(q^1, \hat{p}^1; \mathrm{d}p) \end{cases}$$

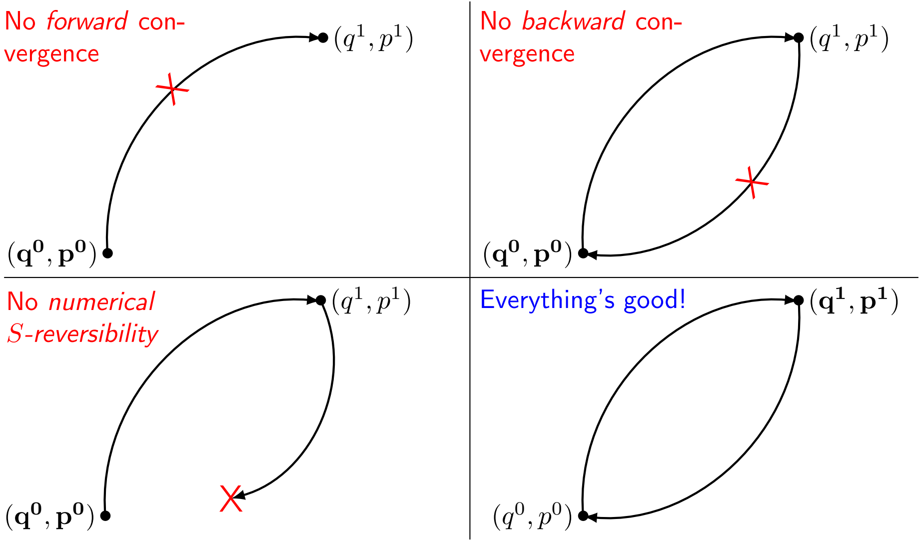
$\leftarrow$ 3 more possible rejections !

where $T^{\mathrm{FD}}_{\Delta t/2}(q^0, p^0; \mathrm{d}p)$ is a consistent discretization of the fluctuation dissipation dynamics over a time interval of length $\Delta t/2$:

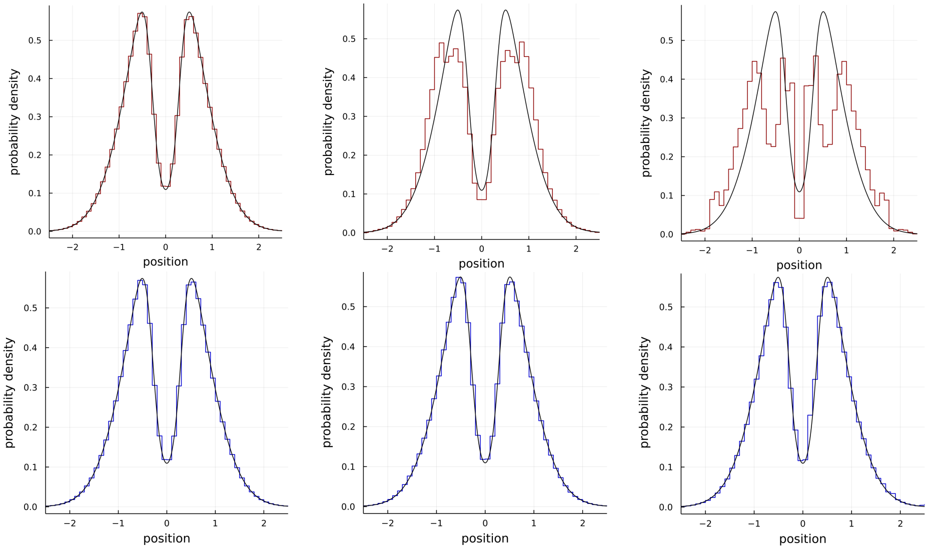
$$\begin{cases} \mathrm{d}p_t = -\gamma \nabla_p H(q^0, p_t)\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}W_t \\ p_0 = p^0 \end{cases}$$

which leaves μ invariant (mid-point, MALA, ...)

Reversibility check: rejections



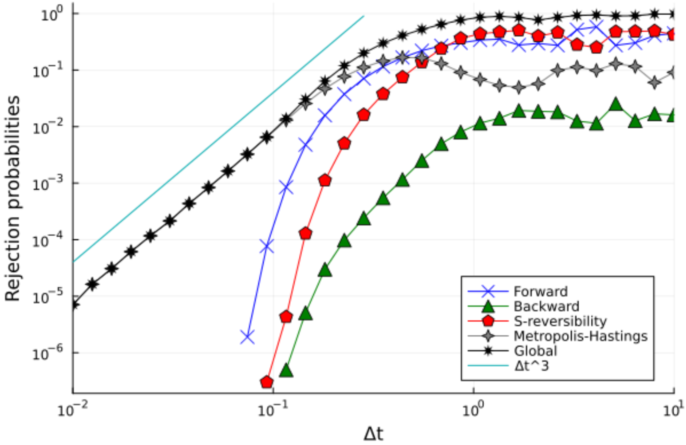
Unbiased sampling with reversibility checks



$\Delta t = 0.18$ $\Delta t = 0.69$ $\Delta t = 1.08$

First row: without reversibility checks; Second row: with reversibility checks

Rejection probability w.r.t. time step



Third picture in last slide: $\Delta t \approx 1.1$

Discussions

Preprint available: `https://arxiv.org/abs/2303.15918`
Lelièvre/RS/Stoltz (2023)

• Construction of the numerical flows (using Newton's method), proofs that $\mathcal{B}_{\Delta t}$ is nonempty under reasonable assumptions (and small enough time steps)

• Other numerical integrators (Implicit Midpoint Rule)

• Other numerical schemes (Generalized Hamiltonian Monte Carlo, Generalized Hybrid MALA[10])

• Sample code available at `https://github.com/rsantet/RMHMC`

[10] Poncet (2017)

Discussions

Preprint available: `https://arxiv.org/abs/2303.15918`
Lelièvre/RS/Stoltz (2023)

• Construction of the numerical flows (using Newton's method), proofs that $\mathcal{B}_{\Delta t}$ is nonempty under reasonable assumptions (and small enough time steps)

• Other numerical integrators (Implicit Midpoint Rule)

• Other numerical schemes (Generalized Hamiltonian Monte Carlo, Generalized Hybrid MALA[10])

• Sample code available at `https://github.com/rsantet/RMHMC`

Thank you !

[10] Poncet (2017)

HMC: canonical measure preservation

Theorem

HMC algorithm preserves the probability measure

$$\mu = \exp\left(-H(q,p)\right)/Z_\mu \, \mathrm{d}q \, \mathrm{d}p$$

Proof $T_{\Delta t}((q,p), \mathrm{d}q' \, \mathrm{d}p') = r_{\Delta t}\delta_{\varphi_{\Delta t}(q,p)}(\mathrm{d}q' \, \mathrm{d}p') + (1 - r_{\Delta t}(q,p))\delta_{(q,p)}(\mathrm{d}q' \, \mathrm{d}p')$

If $f : \mathbb{R}^d \to \mathbb{R}^d \to \mathbb{R}$ measurable & bounded, [x=(q,p), S(q,p)=(q,-p)]

$$\begin{aligned}
\int r_{\Delta t}(x) f(\varphi_{\Delta t}(x))\mu(\mathrm{d}x) \quad &= \int r_{\Delta t}(\varphi_{\Delta t}^{-1}(y)) f(y) \frac{\mathrm{e}^{-\beta\left[H\circ\varphi_{\Delta t}^{-1}\right](y)}}{Z_\mu}\mathrm{d}y \\
[|\nabla\varphi_{\Delta t}| = 1] \quad &= \int r_{\Delta t}((S\circ\varphi_{\Delta t})(z)) f(z) \frac{\mathrm{e}^{-\beta[H\circ S\circ\varphi_{\Delta t}](z)}}{Z_\mu}\mathrm{d}z \\
[S\circ\varphi_{\Delta t}\circ S = \varphi_{\Delta t}^{-1}] \quad &= \int r_{\Delta t}(z) f(z)\mu(\mathrm{d}z)
\end{aligned}$$