Unbiased sampling of HMC schemes for non separable Hamiltonian systems

RÉGIS SANTET, GABRIEL STOLTZ, TONY LELIÈVRE

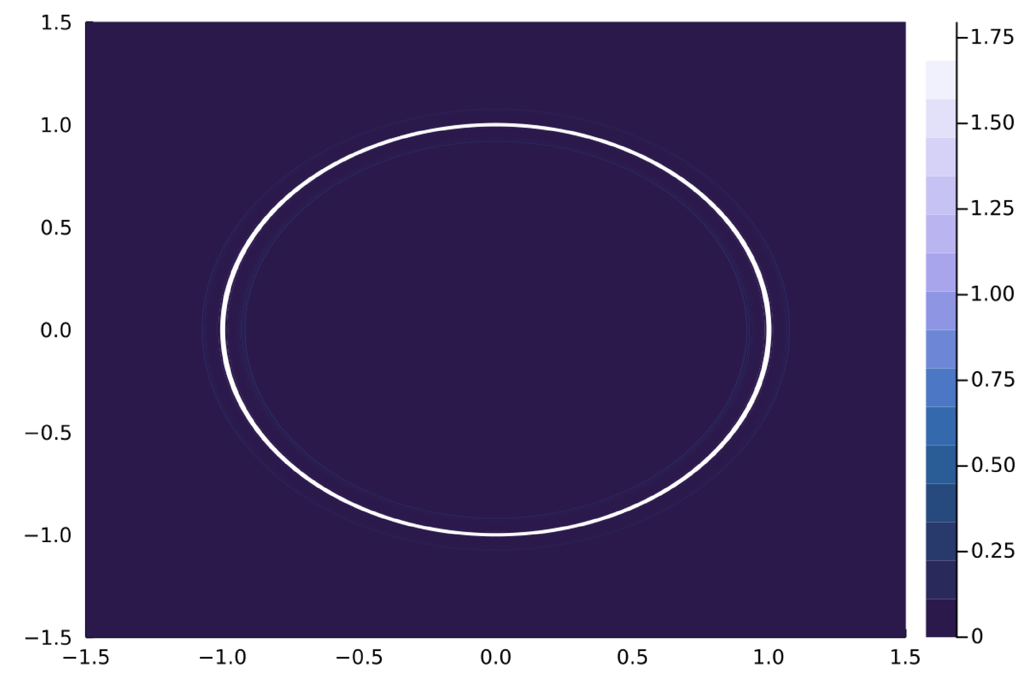
École des Ponts ParisTech — MATHerials — Inria

DIFFUSION DEPENDENT OVERDAMPED LANGEVIN DYNAMICS

Using a position dependent diffusion coefficient[1] D: favoring exploration in **anisotropic** or **metastable** potential landscapes $\Longrightarrow$ faster convergence to steady-state.

$$\boxed{\mathrm{d}q_t = \left(-D(q_t)\nabla V(q_t) + \frac{\operatorname{div} D(q_t)}{\beta}\right)\mathrm{d}t + \sqrt{2\beta^{-1}D(q_t)}\,\mathrm{d}W_t} \qquad (1)$$

- **Aim**: Unbiased estimation of $\mathbb{E}(\varphi) = \int_\Omega \varphi(x)\,\pi(\mathrm{d}x), \quad \pi = \mathrm{e}^{-\beta V}, \quad \beta^{-1} = k_B T.$



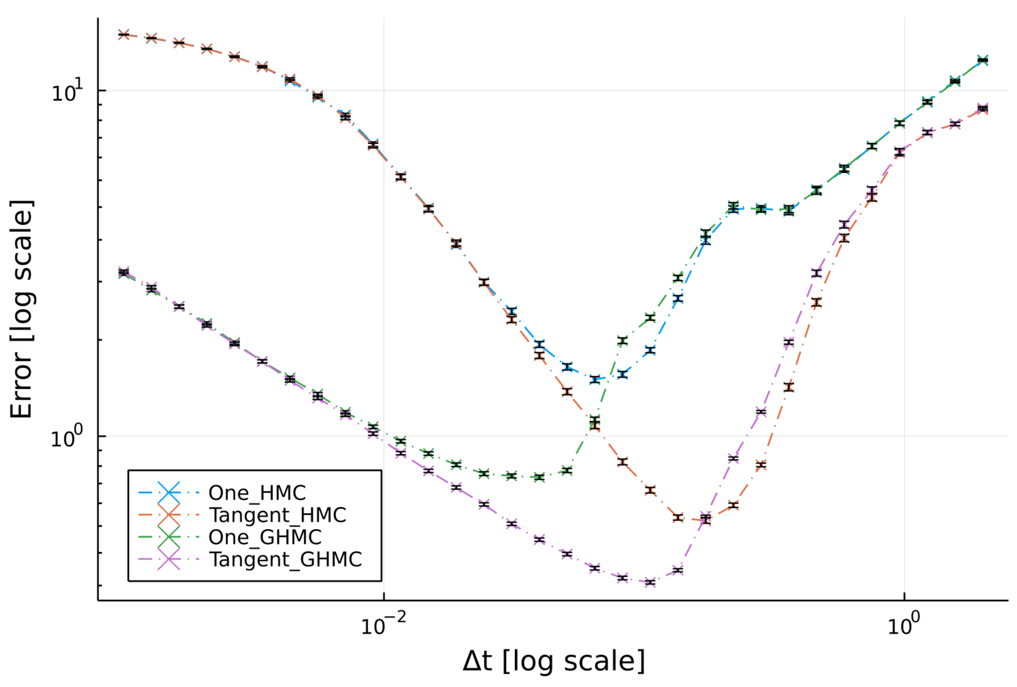
$$V(x,y) = 100(x^2+y^2-1)^2$$

Anisotropic diffusion coefficient

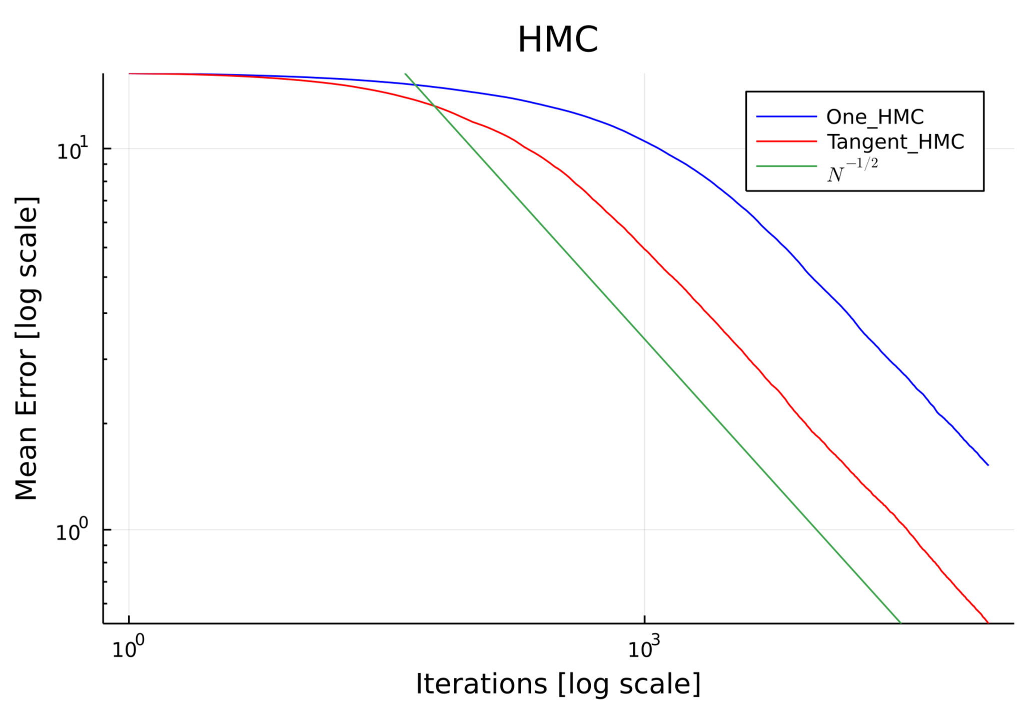
$$D_{\text{Tangent}}(q) = \varepsilon \mathrm{I}_2 + \tilde{q}\tilde{q}^\mathsf{T}/\|q\|^2, \ \tilde{q} = (-y\ x)^\mathsf{T}$$

Isotropic diffusion coefficient

$$D_{\text{One}} \equiv (1+\varepsilon)\mathrm{I}_2, \ \ \varepsilon = 0.1$$



Error between empirical angle distribution and uniform distribution on $[0, 2\pi]$ after 10^5 iterations.



Mean error for optimal time steps in HMC case.

- **Issue**: **Rejection rates scales as $\mathcal{O}\left(\Delta t^{1/2}\right)$** for Euler–Maruyama + Metropolis–Hastings.

BETTER UNBIASED NUMERICAL SAMPLING ?

- **Solution**: **Riemann Manifold (Generalized) Hamiltonian Monte Carlo**[2,3] scheme based on **Langevin dynamics** integration.

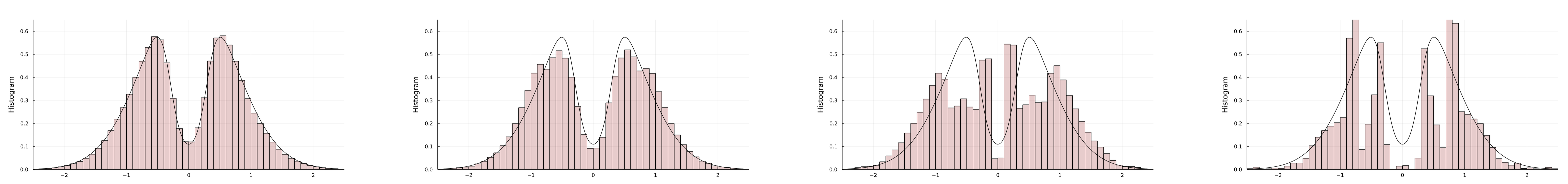
$$\begin{cases} \mathrm{d}q_t = \nabla_p H(q_t,p_t)\,\mathrm{d}t \\ \mathrm{d}p_t = -\nabla_q H(q_t,p_t)\,\mathrm{d}t - \gamma\nabla_p H(q_t,p_t)\,\mathrm{d}t + \sqrt{\dfrac{2\gamma}{\beta}}\,\mathrm{d}W_t \end{cases}$$

- **Rejection rates scale as $\mathcal{O}(\Delta t^3)$**. But need to have a time-reversible and volume-preserving numerical integrator[5].
- **Generalized Störmer–Verlet** [GSV]: implicit symplectic time-reversible integrator[4]

$$\begin{cases} p^{n+\frac12} = p^n - \dfrac{\Delta t}{2}\nabla_q H\left(q^n, p^{n+\frac12}\right) \\ q^{n+1} = q^n + \dfrac{\Delta t}{2}\left(\nabla_p H\left(q^n, p^{n+\frac12}\right) + \nabla_p H\left(q^{n+1}, p^{n+\frac12}\right)\right) \\ p^{n+1} = p^{n+\frac12} - \dfrac{\Delta t}{2}\nabla_q H\left(q^{n+1}, p^{n+\frac12}\right) \end{cases}$$

Proposition. *GSV is first order weakly consistent with the O.L. dynamics with multiplicative noise* (1) *when using the Hamiltonian* $H(q,p) = V(q) - \frac{1}{2}\ln(\det(D(q))) + \frac{1}{2}p^\mathsf{T}D(q)p$.

Sampling results for a double-well confining potential $V(q) = q^2 - 1 + K\mathrm{e}^{-q^2/(2\sigma)}$, oscillating diffusion coefficient $D(q) = \left(\frac{1+\cos(\pi q)}{2}\right)^2$.



$\Delta t = 0.18$ $\quad\Delta t = 0.69$ $\quad\Delta t = 0.86$ $\quad\Delta t = 1.68$

- **Unbiased** sampling with Metropolis–Hastings scheme: naive approach is **biased** !

IMPLICIT INTEGRATION

- Need to check for **forward/backward convergence** of implicit method and **reversibility check** when using Newton's/fixed-point iteration methods[6,7]

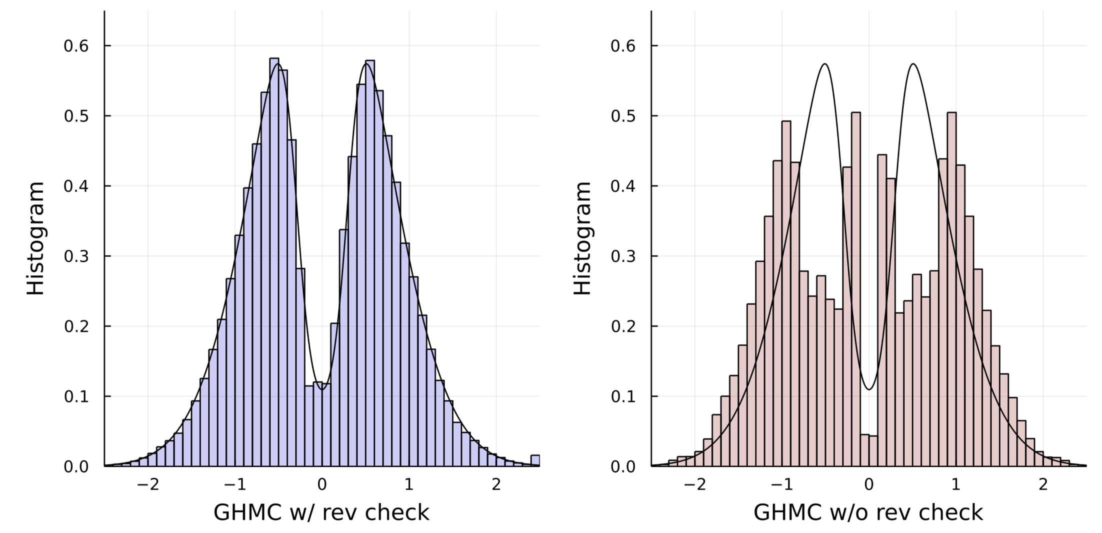
- Starting from q_0, given a time step Δt,

1. Simulate $p_0 \sim \mathcal{N}(0, D(q_0)^{-1})$,
2. Integrate the Hamiltonian dynamics using GSV during Δt
   - if this *forward* integration does not converge, stay in place: return $(q_1, p_1) = (q_0, p_0)$
   - if it converges to $(q_\star, p_\star)$, integrate the Hamiltonian dynamics starting from $(q_\star, -p_\star)$:
     - if this *backward* integration does not converge, stay in place
     - else if the result differs from $(q_0, -p_0)$, stay in place
3. Apply the M–H procedure between (q_0, p_0) and $(q_\star, p_\star)$ and return the position component.

GHMC can be recovered by integrating an Ornstein-Ulhenbeck process for the first step.

- **There are 4 ways to reject the proposal**: no forward/backward convergence, no numerical reversibility, M–H ratio computation.

- Even for large time steps, **unbiased sampling of the configuration space.**



$\Delta t = 0.86$

REFERENCES

[1] N. Bou-Rabee, A. Donev, and E. Vanden-Eijnden. "Metropolis integration schemes for self-adjoint diffusions". In: *Multiscale Modeling & Simulation* (2014).

[2] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. "Hybrid Monte Carlo". In: *Physics Letters B* (1987).

[3] M. Girolami and B. Calderhead. "Riemann Manifold Langevin and Hamiltonian Monte Carlo methods". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* (2011).

[4] E. Hairer, G. Wanner, and C. Lubich. "Symplectic integration of hamiltonian systems". In: *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. 2006.

[5] T. Lelièvre, M. Rousset, and G. Stoltz. *Free Energy Computations: A Mathematical Perspective*. Imperial College Press, 2010.

[6] T. Lelièvre, M. Rousset, and G. Stoltz. "Hybrid Monte Carlo methods for sampling probability measures on submanifolds". In: *Numerische Mathematik* (2019).

[7] E. Zappa, M. Holmes-Cerfon, and J. Goodman. "Monte Carlo on manifolds: sampling densities and integrating functions". In: *Communications on Pure and Applied Mathematics* (2018).