Using Metropolis schemes to estimate correlation functions

Gabriel STOLTZ

gabriel.stoltz@enpc.fr

(CERMICS, Ecole des Ponts & MATHERIALS team, INRIA Rocquencourt)

Work also supported by ANR Funding ANR-14-CE23-0012 ("COSMOS")

Meeting GDR ISIS, Telecom (Paris), November 2015

Motivation

- Computation of integrated correlation functions
  - transport coefficients in molecular dynamics
  - variance of time averages for SDEs $\widehat{\varphi}_t = \dfrac{1}{t}\displaystyle\int_0^t \varphi(q_s)\,ds$
- Assume that...
  - the SDE q_t has a unique invariant measure π
  - φ has average 0 with respect to π
  - time discretization with timestep $\Delta t > 0 \to$ invariant measure $\pi_{\Delta t}$

What is the numerical error arising from $\Delta t > 0$?

$$\sigma_\varphi^2 = \lim_{t\to+\infty} t\,\mathbb{E}\left(\widehat{\varphi}_t^2\right) = 2\int_0^{+\infty} \mathbb{E}\Big(\varphi(q_t)\varphi(q_0)\Big)\,dt$$

- Can be extended to the estimation of $\displaystyle\int_0^{+\infty} \mathbb{E}\Big(\varphi(x_t)\psi(x_0)\Big)\,dt$

Metropolize the discretization of the dynamics...?

- **Pros**
  - Removes the bias on the invariant measure
  - Stabilizes the discretization for non-globally Lipschitz drifts
- **Cons**
  - Scaling of the rejection rate with the dimension
  - Cannot be used for non-reversible dynamics...
  - ... or worse: nonequilibrium systems for which the invariant measure is unknown!
- An early reference in the physics literature... "SmartMC" = MALA!

P. J. Rossky, J. D. Doll, and H. L. Friedman, Brownian dynamics as smart Monte Carlo simulation, *J. Chem. Phys.* (1978)

Error estimates for MALA (1)

- Potential energy function V, invariant measure $\nu(dq) = Z^{-1}\,\mathrm{e}^{-\beta V(q)}\,dq$

Proposal move (recall $\nabla V = (\partial_{q_1} V, \dots, \partial_{q_d} V)$, dimension d)

$$\widetilde{q}^{n+1} = \Phi_{\Delta t}(q^n, G^n) = q^n - \beta \Delta t\, \nabla V(q^n) + \sqrt{2\Delta t}\, G^n$$

- **Acceptance rate**: Metropolis-Hastings criterion

$$A_{\Delta t}\left(q^n, \widetilde{q}^{n+1}\right) = \min\left(\frac{\mathrm{e}^{-\beta V(\widetilde{q}^{n+1})}\, T_{\Delta t}(\widetilde{q}^{n+1}, q^n)}{\mathrm{e}^{-\beta V(q^n)}\, T_{\Delta t}(q^n, \widetilde{q}^{n+1})}, 1\right),$$

where $T_{\Delta t}(q, q') = \left(\dfrac{1}{4\pi\Delta t}\right)^{d/2} \exp\left(-\dfrac{|q' - q + \beta\Delta t\,\nabla V(q)|^2}{4\Delta t}\right)$

Markov chain encoded by a transition function

$$q^{n+1} = \Psi_{\Delta t}(q^n, G^n, U^n) = q^n + \mathbf{1}_{U^n \leqslant A_{\Delta t}(q^n, \Phi_{\Delta t}(q^n, G^n))}\left(\Phi_{\Delta t}\left(q^n, G^n\right) - q^n\right)$$

Error estimates for MALA (2)

- Numerical scheme = Markov chain characterized by transition operator

$$P_{\Delta t}\varphi(q) = \mathbb{E}\Big(\varphi\left(q^{n+1}\right) \,\Big|\, q^n = q\Big)$$

- Reference continuous dynamics $dq_t = -\beta \nabla V(q_t)\,dt + \sqrt{2}\,dW_t$
  - leaves ν invariant
  - generator $\mathcal{L} = -\beta \nabla V(q)^T \nabla + \Delta$ (where $\Delta = \partial_{q_1}^2 + \dots \partial_{q_N}^2$)
  - recall that $\dfrac{d}{dt}\mathbb{E}\Big(\varphi(q_t)\Big) = \mathbb{E}\Big(\mathcal{L}\varphi(q_t)\Big)$

Δt-expansion of the evolution operator

$$P_{\Delta t}\varphi = \varphi + \Delta t\, \mathcal{A}_1\varphi + \Delta t^2 \mathcal{A}_2\varphi + \dots + \Delta t^{p+1}\mathcal{A}_{p+1}\varphi + \Delta t^{p+2} r_{\varphi,\Delta t}$$

- Weak order p when $\displaystyle\sup_{0 \leqslant n \leqslant T/\Delta t} \Big|\mathbb{E}\left[\varphi\left(x^n\right)\right] - \mathbb{E}\left[\varphi\left(x_{n\Delta t}\right)\right]\Big| \leqslant C\Delta t^p$

- Satisfied if $\mathcal{A}_k = \dfrac{\mathcal{L}^k}{k!}$ for all $1 \leqslant k \leqslant p$

Example: Euler-Maruyama, weak order 1 (dimension 1)

- Scheme $q^{n+1} = \Phi_{\Delta t}(q^n, G^n) = q^n - \beta \Delta t \, V'(q^n) + \sqrt{2\Delta t}\, G^n$
- Note that $P_{\Delta t}\varphi(q) = \mathbb{E}_G\left[\varphi\big(\Phi_{\Delta t}(q,G)\big)\right]$
- Technical tool: Taylor expansion

$$\varphi(q+\delta) = \varphi(q) + \delta\varphi'(q) + \frac{1}{2}\delta^2\varphi''(q) + \frac{\delta^3}{6}\varphi^{(3)}(q) + \dots$$

- Replace δ with $\sqrt{2\Delta t}\, G - \beta\Delta t\, V'(q)$ and gather in powers of Δt

$$\begin{aligned}\varphi\big(\Phi_{\Delta t}(q,G)\big) &= \varphi(q) + \sqrt{2\Delta t}\, G\varphi'(q) \\ &\quad + \Delta t\Big(G^2\varphi(q)'' - \beta V'(q)\varphi'(q)\Big) + \dots\end{aligned}$$

- Taking expectations w.r.t. G leads to

$$P_{\Delta t}\varphi(q) = \varphi(q) + \Delta t \underbrace{\Big(\varphi''(q) - \beta V'(q)\varphi'(q)\Big)}_{=\mathcal{L}\varphi(q)} + \mathrm{O}(\Delta t^2)$$

Error estimates for MALA (3)

- For MALA, it can be shown that

$$P_{\Delta t}\varphi = \varphi + \Delta t\,\mathcal{L}\varphi + \Delta t^2 \mathcal{T}\varphi + \Delta t^{5/2} r_{\varphi,\Delta t}$$

(Fractional power of Δt is a signature of Metropolis...)

- An important ingredient is that the rejection rate is of order $\Delta t^{3/2}$

$$\mathbb{E}_G \left| A_{\Delta t}\left(q, q - \beta\Delta t\,\nabla V(q) + \sqrt{2\Delta t}\,G\right)\right] - 1 + \Delta t^{3/2}\overline{\xi}(q)\Big|^p \leqslant C_p \Delta t^{2p}$$

- For compact position spaces, geometric ergodicity can be proved

Error estimates on integrated correlation functions

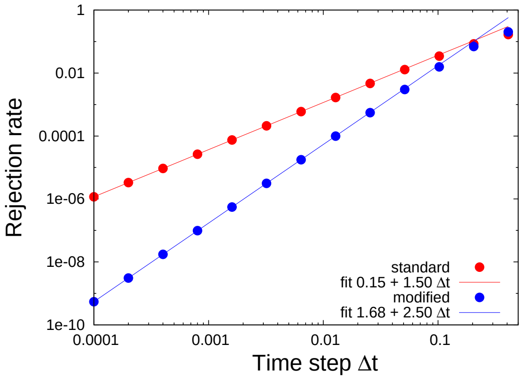
$$\int_0^{+\infty} \mathbb{E}\Big(\varphi(q_t)\varphi(q_0)\Big)\,dt = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t}\Big(\varphi(q^n)\varphi(q^0)\Big) + \mathrm{O}(\Delta t)$$

The error is determined by weak type expansions

Lower the rejection rate?

- Modifying the scheme to lower the rejection rate (1D expressions)
  - modified drift $-\beta V'(q) + \dfrac{\beta \Delta t}{6}\left(V^{(3)} - \beta V'' V'\right)(q)$
  - modified diffusion $\mathrm{Id} + \dfrac{\beta \Delta t}{3} V''(q)$

Rejection rate of order $\Delta t^{5/2}$ but **weak order unchanged**!



Modify the proposal functions

- Midpoint scheme: implicit hence more expensive...

$$\widetilde{q}^{n+1} = q^n - \beta \Delta t \, \nabla V \left(\frac{\widetilde{q}^{n+1} + q^n}{2} \right) + \sqrt{2\Delta t} \, G^n$$

- Hybrid Monte Carlo-like scheme

$$\widetilde{q}^{n+1} = q^n - \beta \Delta t \, \nabla V \left(q^n + \frac{\sqrt{2\Delta t}}{2} \, G^n \right) + \sqrt{2\Delta t} \, G^n$$

Can be reformulated as (using $h = \sqrt{2\beta\Delta t}$)

$$p^n = \beta^{-1/2} G^n, \qquad \begin{cases} q^{n+1/2} = q^n + \dfrac{h}{2} \, p^n, \\ p^{n+1} = p^n - h \, \nabla V \left(q^{n+1/2} \right), \\ \widetilde{q}^{n+1} = q^{n+1/2} + \dfrac{h}{2} \, p^{n+1}. \end{cases}$$

Reversible structure: allows to compute the Metropolis ratio in terms of some extended energy difference $H(q,p) = V(q) + p^2/2$

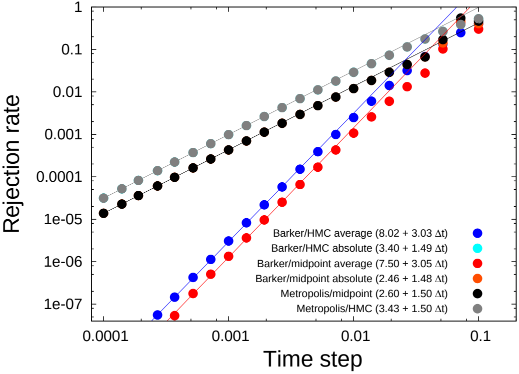
Modify the acceptance criterion

- **Metropolis criterion** $A_{\Delta t}^{\mathrm{MH}}(q^n, \widetilde{q}^{n+1}) = \min\left(1, \mathrm{e}^{-\alpha_{\Delta t}(q^n,\widetilde{q}^{n+1})}\right)$

$\rightarrow$ rejection rate $1 - \mathrm{O}(\Delta t^{3/2})$

- **Barker rule** $A_{\Delta t}^{\mathrm{Barker}}(q^n, \widetilde{q}^{n+1}) = \dfrac{\mathrm{e}^{-\alpha_{\Delta t}(q^n,\widetilde{q}^{n+1})}}{1 + \mathrm{e}^{-\alpha_{\Delta t}(q^n,\widetilde{q}^{n+1})}}$

$\rightarrow$ rejection rate $1/2 + \mathrm{O}(\Delta t^3)$ in average, $1/2 + \mathrm{O}(\Delta t^{3/2})$ in absolute value



Results on integrated correlation functions

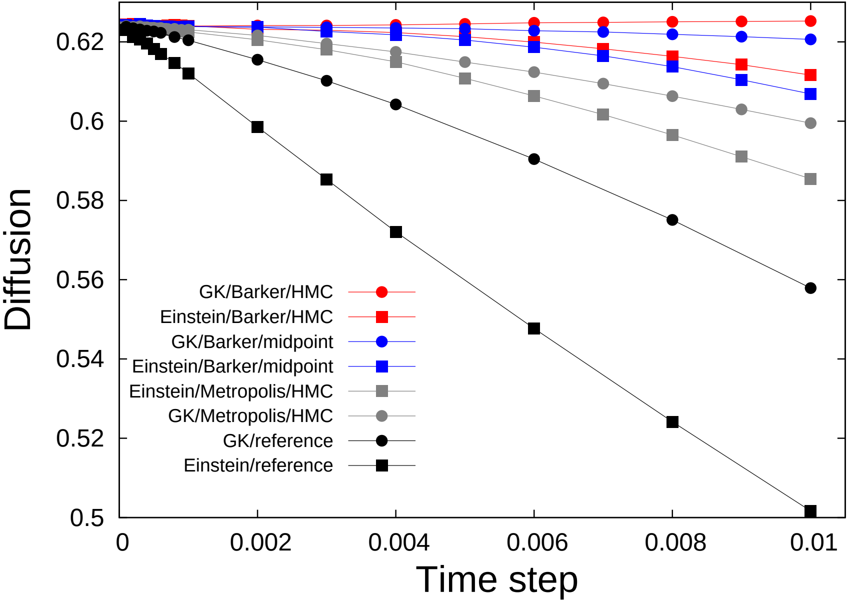
Improved Green-Kubo formulas

Set $a = 1/2$ and $\alpha = 2$ for Barker, and $a = 1$ and $\alpha = 3/2$ for Metropolis-Hastings. Then,

$$\int_0^{+\infty} \mathbb{E}\Big[\psi(q_t)\varphi(q_0)\Big]\,dt = \Delta t\left(a\sum_{n=0}^{+\infty}\mathbb{E}_{\Delta t}\left[\psi(q^n)\varphi(q^0)\right] - \frac{\mathbb{E}_\nu(\psi\varphi)}{2}\right) + \mathrm{O}(\Delta t^\alpha)$$

- **Some comments...**
  - reduces to trapezoidal rule for Metropolis (but error $\Delta t^{3/2}$)
  - time renormalization by a factor 2 for Barker
  - statistical error increased by factor 2 for Barker, but reduced bias
  - no fractional powers of Δt when Barker is used
- Key ingredient in the proof: $\dfrac{P_{\Delta t} - \mathrm{Id}}{\Delta t}\varphi = a\left(\mathcal{L}\varphi + \dfrac{\Delta t}{2}\mathcal{L}^2\varphi\right) + \mathrm{O}(\Delta t^\alpha)$
- **Numerical illustration** for 1D system with $V(q) = \cos(2\pi q)$ and $\beta = 1$

Results on integrated correlations $\varphi = \psi = V'$



Conclusion and perspectives

- Numerical analysis of integrated correlation functions → bias

- Extension to dynamics with multiplicative noise

$$dq_t = \Big(-\beta M(q_t)\nabla V(q_t) + \mathrm{div}(M)(q_t)\Big) dt + \sqrt{2} M^{1/2}(q_t)\, dW_t$$

- Many open issues when the invariant measure is not known explicitely...

→ Nonequilibrium systems in molecular dynamics

References

B. Leimkuhler, Ch. Matthews and G. Stoltz, The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics, *IMA J. Numer. Anal.* (2015)

M. Fathi, A.-A. Homman and G. Stoltz, Error analysis of the transport properties of Metropolized schemes, *ESAIM Proc.* (2015)

M. Fathi and G. Stoltz, Improving dynamical properties of stabilized discretizations of overdamped Langevin dynamics, *arXiv* **1505.04905** (2015)