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
    \[ \hat{\varphi}_t = \frac{1}{t} \int_0^t \varphi(q_s) \, ds \]

- Assume that...
  - the SDE \( q_t \) has a unique invariant measure \( \pi \)
  - \( \varphi \) has average 0 with respect to \( \pi \)
  - time discretization with timestep \( \Delta t > 0 \) → invariant measure \( \pi \Delta t \)

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

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

- Can be extended to the estimation of
  \[ \int_0^{+\infty} \mathbb{E}\left( \varphi(x_t)\psi(x_0) \right) \, dt \]
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!
Error estimates for MALA (1)

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

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

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

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

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

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

Markov chain encoded by a transition function

\[
q^{n+1} = \Psi_{\Delta t}(q^n, G^n, U^n) = q^n + 1_{U^n \leq A_{\Delta t}(q^n, \Phi_{\Delta t}(q^n, G^n))} \left( \Phi_{\Delta t} (q^n, G^n) - q^n \right)
\]
Error estimates for MALA (2)

• Numerical scheme = Markov chain characterized by transition operator
  \[ P_{\Delta t}\varphi(q) = \mathbb{E}\left(\varphi(q^{n+1}) \mid q^n = q\right) \]

• Reference continuous dynamics \[ dq_t = -\beta \nabla V(q_t) \, dt + \sqrt{2} \, dW_t \]
  leaves \( \nu \) invariant
  generator \( \mathcal{L} = -\beta \nabla V(q)^T \nabla + \Delta \) (where \( \Delta = \partial^2_{q_1} + \ldots + \partial^2_{q_N} \))
  recall that \[ \frac{d}{dt} \mathbb{E}\left(\varphi(q_t)\right) = \mathbb{E}\left(\mathcal{L}\varphi(q_t)\right) \]

**\( \Delta t \)-expansion of the evolution operator**

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

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

• Satisfied if \( \mathcal{A}_k = \frac{\mathcal{L}^k}{k!} \) for all \( 1 \leq k \leq p \)
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 \( P_{\Delta t} \phi(q) = E_G [\phi(\Phi_{\Delta t}(q, G))] \)

- Technical tool: Taylor expansion
  \[
  \phi(q + \delta) = \phi(q) + \delta \phi'(q) + \frac{1}{2} \delta^2 \phi''(q) + \frac{\delta^3}{6} \phi^{(3)}(q) + \ldots
  \]

- Replace \( \delta \) with \( \sqrt{2\Delta t} \, G - \beta \Delta t \, V'(q) \) and gather in powers of \( \Delta t \)
  \[
  \phi(\Phi_{\Delta t}(q, G)) = \phi(q) + \sqrt{2\Delta t} \, G \phi'(q) + \Delta t \left( G^2 \phi''(q) - \beta V'(q) \phi'(q) \right) + \ldots
  \]

- Taking expectations w.r.t. \( G \) leads to
  \[
  P_{\Delta t} \phi(q) = \phi(q) + \Delta t \left( \phi''(q) - \beta V'(q) \phi'(q) \right) + O(\Delta t^2) = L \phi(q)
  \]
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 \( \Delta t \) is a signature of Metropolis...)

• An important ingredient is that the rejection rate is of order \( \Delta t^{3/2} \)

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

• For compact position spaces, geometric ergodicity can be proved

Error estimates on integrated correlation functions

\[ \int_{0}^{+\infty} \mathbb{E} \left( \varphi(q_t) \varphi(q_0) \right) dt = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} \left( \varphi(q^n) \varphi(q^0) \right) + O(\Delta t) \]

The error is determined by weak type expansions
Lower the rejection rate?

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

Rejection rate of order \(\Delta t^{5/2}\) but **weak order unchanged**!
Modify the proposal functions

- **Midpoint scheme**: implicit hence more expensive...

  \[ \tilde{q}^{n+1} = q^n - \beta \Delta t \nabla V \left( \frac{\tilde{q}^{n+1} + q^n}{2} \right) + \sqrt{2\Delta t} G^n \]

- **Hybrid Monte Carlo-like** scheme

  \[ \tilde{q}^{n+1} = q^n - \beta \Delta t \nabla V \left( q^n + \frac{\sqrt{2\Delta t}}{2} G^n \right) + \sqrt{2\Delta t} G^n \]

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

\[
\begin{align*}
  q^{n+1/2} &= q^n + \frac{h}{2} p^n, \\
p^n &= \beta^{-1/2} G^n, \\
p^{n+1} &= p^n - h \nabla V \left( q^{n+1/2} \right), \\
\tilde{q}^{n+1} &= q^{n+1/2} + \frac{h}{2} p^{n+1}.
\end{align*}
\]

**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**
  \[ A^\text{MH}_{\Delta t}(q^n, \tilde{q}^{n+1}) = \min\left(1, e^{-\alpha \Delta t(q^n, \tilde{q}^{n+1})}\right) \]
  \[ \rightarrow \text{rejection rate } 1 - O(\Delta t^{3/2}) \]

- **Barker rule**
  \[ A^\text{Barker}_{\Delta t}(q^n, \tilde{q}^{n+1}) = \frac{e^{-\alpha \Delta t(q^n, \tilde{q}^{n+1})}}{1 + e^{-\alpha \Delta t(q^n, \tilde{q}^{n+1})}} \]
  \[ \rightarrow \text{rejection rate } 1/2 + O(\Delta t^3) \text{ in average, } 1/2 + O(\Delta t^{3/2}) \text{ in absolute value} \]
Improved Green-Kubo formulas

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

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

- 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 $\Delta t$ when Barker is used

- Key ingredient in the proof: 
  $$
  \frac{P_{\Delta t} - Id}{\Delta t} \varphi = a \left( L\varphi + \frac{\Delta t}{2} L^2\varphi \right) + O(\Delta t^\alpha)
  $$

- Numerical illustration for 1D system with $V(q) = \cos(2\pi q)$ and $\beta = 1$
Results on integrated correlations $\phi = \psi = V'$
Conclusion and perspectives

- Numerical analysis of integrated correlation functions → bias
- Extension to dynamics with multiplicative noise
  \[ dq_t = \left( -\beta M(q_t)\nabla V(q_t) + \text{div}(M)(q_t) \right) dt + \sqrt{2}M^{1/2}(q_t) \, dW_t \]
- Many open issues when the invariant measure is not known explicitly...
  → Nonequilibrium systems in molecular dynamics

References

