Error estimates on the computation of transport coefficients

Gabriel STOLTZ

gabrie.stoltz@enpc.fr

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

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

NESP 2015, ICTS (Bangalore), November 2015

Motivation

- Transport coefficients can be obtained from
  - nonequilibrium dynamics in the linear response regime
  - integrated correlation functions at equilibrium (Green-Kubo)
  - transient response to displacement from equilibrium

For concreteness: mobility/self-diffusion (direction $F \in \mathbb{R}^d$)

$$F^T D F = \lim_{t \to +\infty} \frac{\mathbb{E}\left[\left(F \cdot (Q_t - Q_0)\right)^2\right]}{2t}$$

What is the numerical error arising from time discretization?

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)

Reminder: Error estimates in Monte Carlo simulations

- General SDE $dx_t = b(x_t)\,dt + \sigma(x_t)\,dW_t$, invariant measure π

- Discretization $x^{n+1} \simeq x_{n\Delta t}$, invariant measure $\pi_{\Delta t}$. For instance,

$$x^{n+1} = x^n + \Delta t\, b(x^n) + \sqrt{\Delta t}\,\sigma(x^n)\,G^n, \qquad G^n \sim \mathcal{G}(0, \mathrm{Id}) \text{ i.i.d.}$$

Error estimates for finite trajectory averages

$$\widehat{A}_{N_{\mathrm{iter}}} = \frac{1}{N_{\mathrm{iter}}} \sum_{n=1}^{N_{\mathrm{iter}}} A(x^n) = \mathbb{E}_\pi(A) + \underbrace{C\Delta t^\alpha}_{\mathrm{bias}} + \underbrace{\frac{\sigma_{A,\Delta t}}{\sqrt{N_{\mathrm{iter}}}}\mathscr{G}}_{\mathrm{statistical\ error}}$$

- Bias $\mathbb{E}_{\pi_{\Delta t}}(A) - \mathbb{E}_\pi(A) \longrightarrow$ **Focus of this talk**

- Variance $\dfrac{\sigma_{A,\Delta t}}{\sqrt{N_{\mathrm{iter}}}} \simeq \sqrt{\dfrac{2\displaystyle\int_0^{+\infty} \mathbb{E}\left[\delta A(x_t)\delta A(x_0)\right]\,dt}{T}}$ where $T = N_{\mathrm{iter}}\Delta t$

Definition of the self-diffusion

Definition of the self-diffusion: Langevin dynamics

- Periodic potential V, Langevin dynamics for $(q,p) \in \mathcal{E} = (L\mathbb{T})^d \times \mathbb{R}^d$

$$\left\{ \begin{aligned} dq_t &= M^{-1} p_t \, dt \\ dp_t &= \Big(-\nabla V(q_t) + \eta F \Big) dt - \gamma M^{-1} p_t \, dt + \sqrt{\frac{2\gamma}{\beta}} \, dW_t \end{aligned} \right.$$

Mobility: averages in a nonequilibrium steady-state

$$\nu_F = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta \left(F^T M^{-1} p \right)}{\eta} = \beta \int_{\mathcal{E}} F^T M^{-1} p \, f_{0,1}(q,p) \, \mu(dq\, dp) = \beta \, F^T D F$$

Effective diffusion computed at equilibrium: $\eta = 0$

Unperiodized displacement $Q_t - Q_0 = \displaystyle\int_0^t M^{-1} p_s \, ds$

$$F^T D F = \int_0^{+\infty} \mathbb{E}_0 \left[\left(F^T M^{-1} p_t \right) \left(F^T M^{-1} p_0 \right) \right] dt$$

- Integrability by exp. convergence to $\mu(dq\, dp) = Z^{-1} \mathrm{e}^{-\beta H(q,p)} \, dq \, dp$

Definition of the self-diffusion: overdamped Langevin

- Dynamics $dq_t = \left(-\nabla V(q_t) + \eta F \right) dt + \sqrt{\frac{2}{\beta}} \, dW_t$
- Invariant measure $\nu(dq) = Z^{-1} \mathrm{e}^{-\beta V(q)} \, dq$

Mobility: averages in a nonequilibrium steady-state

$$\nu_F = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta \left(-F^T \nabla V(q) \right)}{\eta} = \beta \int_{\mathcal{E}} -F^T \nabla V(q) \, g_{0,1}(q) \, \nu(dq) = \frac{1}{\beta} F^T \left(\mathrm{Id} - D \right) F$$

Effective diffusion computed at equilibrium: $\eta = 0$

Unperiodized displacement $Q_t - Q_0 = -\int_0^t \nabla V(q_s) \, ds + \sqrt{\frac{2}{\beta}} \, W_t$

$$F^T D F = |F|^2 - \beta^2 \int_0^{+\infty} \mathbb{E}_0 \left[\left(F^T \nabla V(q_t) \right) \left(F^T \nabla V(q_0) \right) \right] dt$$

General definition of transport coefficients

- Ergodic stochastic dynamics $dX_t = b(X_t)\,dt + \sigma\,dW_t$
  - Invariant measure $\pi(dx)$
  - Generator $\mathcal{L} = b \cdot \nabla + \dfrac{\sigma^2}{2}\Delta$
- Drift perturbation $b + \eta\widetilde{b} \longrightarrow$ perturbation of generator $\widetilde{\mathcal{L}} = \widetilde{b} \cdot \nabla$

Linear response and Green-Kubo type formulas

With adjoints taken on $L^2(\pi)$ and when $\mathbb{E}_0(R) = 0$,

$$\alpha = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta(R)}{\eta} = -\int_{\mathcal{E}} \left[\mathcal{L}^{-1}R\right]\left[\widetilde{\mathcal{L}}^*\mathbf{1}\right] d\pi = \int_0^{+\infty} \mathbb{E}_0\Big(R(x_t)S(x_0)\Big)\,dt$$

- Conjugated response function $S = \widetilde{\mathcal{L}}^*\mathbf{1}$
- Relies on the equality $-\mathcal{L}^{-1} = \displaystyle\int_0^{+\infty} \mathrm{e}^{t\mathcal{L}}\,dt$ (functional analysis)

Error estimates on α resulting from finite $\Delta t > 0$?

Error estimates for equilibrium dynamics

Weak type expansions

- Numerical scheme = Markov chain characterized by evolution operator
$$P_{\Delta t}\varphi(x) = \mathbb{E}\Big(\varphi\left(x^{n+1}\right) \Big| x^n = x\Big)$$
where (x^n) is an approximation of $(x_{n\Delta t})$

- (Infinitely) Many possibilities! Numerical analysis allows to discriminate

Δt-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 $\mathcal{A}_k = \mathcal{L}^k/k!$ for $1 \leqslant k \leqslant p$, namely
$$\sup_{0\leqslant n\leqslant T/\Delta t} \Big| \mathbb{E}\left[\varphi\left(x^n\right)\right] - \mathbb{E}\left[\varphi\left(x_{n\Delta t}\right)\right] \Big| \leqslant C\Delta t^p$$

- Ergodicity of the numerical scheme with invariant measure $\pi_{\Delta t}$
$$\frac{1}{N_{\rm iter}} \sum_{n=1}^{N_{\rm iter}} A(x^n) \xrightarrow[N_{\rm iter}\to+\infty]{} \int_{\mathcal{X}} A(x)\, \pi_{\Delta t}(dx)$$

Example: Euler-Maruyama, weak order 1

- Scheme $x^{n+1} = \Phi_{\Delta t}(x^n, G^n) = x^n + \Delta t\, b(x^n) + \sqrt{\Delta t}\, \sigma(x^n)\, G^n$
- Note that $P_{\Delta t}\varphi(x) = \mathbb{E}_G\left[\varphi\big(\Phi_{\Delta t}(x, G)\big)\right]$
- Technical tool: Taylor expansion

$$\varphi(x+\delta) = \varphi(x) + \delta^T \nabla\varphi(x) + \frac{1}{2}\delta^T \nabla^2\varphi(x)\delta + \frac{1}{6}D^3\varphi(x) : \delta^{\otimes 3} + \dots$$

- Replace δ with $\sqrt{\Delta t}\,\sigma(x)\, G + \Delta t\, b(x)$ and gather in powers of Δt

$$\begin{aligned}\varphi\big(\Phi_{\Delta t}(x, G)\big) &= \varphi(x) + \sqrt{\Delta t}\,\sigma(x)\, G \cdot \nabla\varphi(x) \\ &\quad + \Delta t \left(\frac{\sigma(x)^2}{2} G^T \left[\nabla^2\varphi(x)\right] G + b(x)\cdot\nabla\varphi(x)\right) + \dots\end{aligned}$$

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

$$P_{\Delta t}\varphi(x) = \varphi(x) + \Delta t \underbrace{\left(\frac{\sigma(x)^2}{2}\Delta\varphi(x) + b(x)\cdot\nabla\varphi(x)\right)}_{=\mathcal{L}\varphi(x)} + \mathrm{O}(\Delta t^2)$$

Error estimates on the invariant measure

- Assumptions on the operators in the weak-type expansion
  - invariance of π by $\mathcal{A}_k$ for $1 \leqslant k \leqslant p$, namely $\displaystyle\int_{\mathcal{X}} \mathcal{A}_k \varphi \, d\pi = 0$
  - $\displaystyle\int_{\mathcal{X}} \mathcal{A}_{p+1} \varphi \, d\pi = \int_{\mathcal{X}} g_{p+1} \varphi \, d\pi$ (*i.e.* $g_{p+1} = \mathcal{A}_{p+1}^* \mathbf{1}$)

Error estimates on $\pi_{\Delta t}$

$$\int_{\mathcal{X}} \varphi \, d\pi_{\Delta t} = \int_{\mathcal{X}} \varphi \Big(1 + \Delta t^p f_{p+1} \Big) d\pi + \Delta t^{p+1} R_{\varphi, \Delta t}$$

- In fact, $f_{p+1} = -\left(\mathcal{A}_1^*\right)^{-1} g_{p+1}$
  - when $\mathcal{A}_1 = \mathcal{L}$, the first order correction can be estimated by some integrated correlation function as $\displaystyle\int_0^{+\infty} \mathbb{E}\Big(\varphi(x_t) g_{p+1}(x_0) \Big) \, dt$
  - in general, first order term can be removed by Romberg extrapolation
- Error on invariant measure can be (much) smaller than the weak error

Sketch of proof (1)

Step 1: Establish the error estimate for $\varphi \in \mathrm{Ran}(P_{\Delta t} - \mathrm{Id})$

- Idea: $\pi_{\Delta t} = \pi(1 + \Delta t^p f_{p+1} + \dots)$
- by definition of $\pi_{\Delta t}$

$$\int_{\mathcal{X}} \left[\left(\frac{P_{\Delta t} - \mathrm{Id}}{\Delta t} \right) \psi \right] d\pi_{\Delta t} = 0$$

- compare to first order correction to the invariant measure

$$\begin{aligned}
&\int_{\mathcal{X}} \left[\left(\frac{P_{\Delta t} - \mathrm{Id}}{\Delta t} \right) \psi \right] (1 + \Delta t^p f_{p+1}) \, d\pi \\
&\qquad = \Delta t^p \int_{\mathcal{X}} \Big(\mathcal{A}_{p+1}\psi + (\mathcal{A}_1\psi) f_{p+1} \Big) d\pi + \mathrm{O}\left(\Delta t^{p+1}\right) \\
&\qquad = \Delta t^p \int_{\mathcal{X}} \Big(g_{p+1} + \mathcal{A}_1^* f_{p+1} \Big) \psi \, d\pi + \mathrm{O}\left(\Delta t^{p+1}\right)
\end{aligned}$$

Suggests $f_{p+1} = -\left(\mathcal{A}_1^*\right)^{-1} g_{p+1}$

Sketch of proof (2)

Step 2: Define an approximate inverse

- Issue: derivatives of $(\mathrm{Id} - P_{\Delta t})^{-1}\varphi$ are not controlled
- Consider $\left(\Pi \dfrac{P_{\Delta t} - \mathrm{Id}}{\Delta t} \Pi\right) Q_{\Delta t}\psi = \psi + \Delta t^{p+1} \widetilde{r}_{\psi,\Delta t}$ where

$$\Pi\varphi = \varphi - \int_{\mathcal{X}} \varphi \, d\pi$$

- Idea of the construction: truncate the formal series expression

$$(A + \Delta t\, B)^{-1} = A^{-1} - \Delta t\, A^{-1} B A^{-1} + \Delta t^2\, A^{-1} B A^{-1} B A^{-1} + \dots$$

- Here, set $A = \Pi \mathcal{A}_1 \Pi$ and $B = \Pi \mathcal{A}_2 \Pi + \dots + \Delta t^{p-1} \Pi \mathcal{A}_{p+1} \Pi$
- $Q_{\Delta t}$ stabilizes spaces of smooth functions

Sketch of proof (3)

Step 3: Conclusion

- Invariance of $\pi_{\Delta t}$ rewritten as

$$\int_{\mathcal{X}} \left[\Pi \left(\frac{P_{\Delta t} - \mathrm{Id}}{\Delta t} \right) \Pi \psi \right] d\pi_{\Delta t} = -\frac{1}{\Delta t} \int_{\mathcal{X}} (P_{\Delta t} \psi) \, d\pi,$$

- On the other hand,

$$\int_{\mathcal{X}} \left[\Pi \left(\frac{P_{\Delta t} - \mathrm{Id}}{\Delta t} \right) \Pi \psi \right] (1 + \Delta t^p f_{p+1}) \, d\pi = -\frac{1}{\Delta t} \int_{\mathcal{X}} (P_{\Delta t} \psi) \, d\pi + \mathrm{O}\left(\Delta t^{p+1}\right)$$

- Finally replace ψ by $Q_{\Delta t}\varphi$, and gather in $R_{\varphi,\Delta t}$ all the higher order terms

Error estimates
on the linear response
of nonequilibrium dynamics

Examples of splitting schemes for Langevin dynamics (1)

- Example: Langevin dynamics, discretized using a splitting strategy

$$A = M^{-1}p\cdot\nabla_q, \quad B_\eta = \Big(-\nabla V(q) + \eta F\Big)\cdot\nabla_p, \quad C = -M^{-1}p\cdot\nabla_p + \frac{1}{\beta}\Delta_p$$

- First order splitting schemes: Trotter splitting

$$P_{\Delta t}^{ZYX} = \mathrm{e}^{\Delta t Z}\mathrm{e}^{\Delta t Y}\mathrm{e}^{\Delta t X} \simeq \mathrm{e}^{\Delta t \mathcal{A}}$$

- Second order schemes: Strang splitting

$$P_{\Delta t}^{ZYXYZ} = \mathrm{e}^{\Delta t Z/2}\mathrm{e}^{\Delta t Y/2}\mathrm{e}^{\Delta t X}\mathrm{e}^{\Delta t Y/2}\mathrm{e}^{\Delta t Z/2}$$

- Other category: Geometric Langevin algorithms, *e.g.* $P_{\Delta t}^{\gamma C, A, B_\eta, A}$

Examples of splitting schemes for Langevin dynamics (2)

- $P_{\Delta t}^{B_\eta,A,\gamma C}$ corresponds to $\begin{cases} \widetilde{p}^{n+1} = p^n + \Big(-\nabla V(q^n) + \eta F \Big) \Delta t, \\ q^{n+1} = q^n + \Delta t \, M^{-1} \widetilde{p}^{n+1}, \\ p^{n+1} = \alpha_{\Delta t} \widetilde{p}^{n+1} + \sqrt{\dfrac{1-\alpha_{\Delta t}^2}{\beta} M} \, G^n \end{cases}$

where G^n are i.i.d. Gaussian and $\alpha_{\Delta t} = \exp(-\gamma M^{-1} \Delta t)$

- $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$ for $\begin{cases} \widetilde{p}^{n+1/2} = \alpha_{\Delta t/2} p^n + \sqrt{\dfrac{1-\alpha_{\Delta t}}{\beta} M} \, G^n, \\ p^{n+1/2} = \widetilde{p}^{n+1/2} + \dfrac{\Delta t}{2} \Big(-\nabla V(q^n) + \eta F \Big), \\ q^{n+1} = q^n + \Delta t \, M^{-1} p^{n+1/2}, \\ \widetilde{p}^{n+1} = p^{n+1/2} + \dfrac{\Delta t}{2} \Big(-\nabla V(q^{n+1}) + \eta F \Big), \\ p^{n+1} = \alpha_{\Delta t/2} \widetilde{p}^{n+1} + \sqrt{\dfrac{1-\alpha_{\Delta t}}{\beta} M} \, G^{n+1/2} \end{cases}$

Error estimates on linear response

Error estimates for nonequilibrium dynamics

There exists a function $f_{\alpha,1,\gamma} \in H^1(\mu)$ such that

$$\int_{\mathcal{E}} \psi \, d\mu_{\gamma,\eta,\Delta t} = \int_{\mathcal{E}} \psi \Big(1 + \eta f_{0,1,\gamma} + \Delta t^{\alpha} f_{\alpha,0,\gamma} + \eta \Delta t^{\alpha} f_{\alpha,1,\gamma} \Big) d\mu + r_{\psi,\gamma,\eta,\Delta t},$$

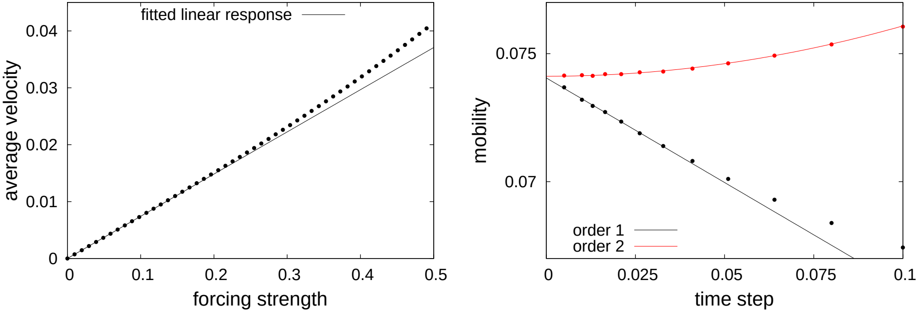
where the remainder is compatible with linear response

$$|r_{\psi,\gamma,\eta,\Delta t}| \leqslant K(\eta^2 + \Delta t^{\alpha+1}), \qquad |r_{\psi,\gamma,\eta,\Delta t} - r_{\psi,\gamma,0,\Delta t}| \leqslant K\eta(\eta + \Delta t^{\alpha+1})$$

- Corollary: error estimates on the numerically computed mobility

$$\begin{aligned} \nu_{F,\gamma,\Delta t} &= \lim_{\eta \to 0} \frac{1}{\eta} \left(\int_{\mathcal{E}} F^T M^{-1} p \, \mu_{\gamma,\eta,\Delta t}(dq\, dp) - \int_{\mathcal{E}} F^T M^{-1} p \, \mu_{\gamma,0,\Delta t}(dq\, dp) \right) \\ &= \nu_{F,\gamma} + \Delta t^{\alpha} \int_{\mathcal{E}} F^T M^{-1} p \, f_{\alpha,1,\gamma} \, d\mu + \Delta t^{\alpha+1} r_{\gamma,\Delta t} \end{aligned}$$

- Results in the overdamped limit



Left: Linear response of the average velocity as a function of η for the scheme associated with $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$ and $\Delta t = 0.01, \gamma = 1$.

Right: Scaling of the mobility $\nu_{F,\gamma,\Delta t}$ for the first order scheme $P_{\Delta t}^{A, B_\eta, \gamma C}$ and the second order scheme $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$.

Error estimates on Green-Kubo formulas

Error estimates on Green-Kubo formulas (1)

- Error of order α on invariant measure: $\displaystyle \int_{\mathcal{X}} \psi \, d\pi_{\Delta t} = \int_{\mathcal{X}} \psi \, d\pi + \mathrm{O}(\Delta t^{\alpha})$
- Expansion of the evolution operator ($p+1 \geqslant \alpha$ and $\mathcal{A}_1 = \mathcal{L}$)

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

Ergodicity of the numerical scheme

$$\forall n \in \mathbb{N}, \qquad \|P_{\Delta t}^n\|_{\mathcal{B}(L^\infty_{\mathcal{K}_s,\Delta t})} \leqslant C_s \mathrm{e}^{-\lambda_s n \Delta t}$$

where $\mathcal{K}_s$ is a Lyapunov function ($1+|p|^{2s}$ for Langevin) and

$$L^\infty_{\mathcal{K}_s,\Delta t} = \left\{ \frac{\varphi}{\mathcal{K}_s} \in L^\infty(\mathcal{X}), \ \int_{\mathcal{X}} \varphi \, d\pi_{\Delta t} = 0 \right\}$$

- Proof: Lyapunov condition + uniform-in-Δt minorization condition[1]

[1] M. Hairer and J. Mattingly, *Progr. Probab.* (2011)

Error estimates on Green-Kubo formulas (2)

Error estimates on integrated correlation functions

Observables φ, ψ with average 0 w.r.t. invariant measure π

$$\int_0^{+\infty} \mathbb{E}\Big(\psi(x_t)\varphi(x_0)\Big)dt = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t}\left(\widetilde{\psi}_{\Delta t,\alpha}\left(x^n\right)\varphi\left(x^0\right)\right) + \Delta t^{\alpha} r_{\Delta t}^{\psi,\varphi},$$

where $\mathbb{E}_{\Delta t}$ denotes expectations w.r.t. initial conditions $x_0 \sim \pi_{\Delta t}$ and over all realizations of the Markov chain (x^n), and

$$\widetilde{\psi}_{\Delta t,\alpha} = \psi_{\Delta t,\alpha} - \int_{\mathcal{X}} \psi_{\Delta t,\alpha}\, d\pi_{\Delta t}$$

with $\psi_{\Delta t,\alpha} = \Big(\mathrm{Id} + \Delta t\, \mathcal{A}_2\mathcal{L}^{-1} + \cdots + \Delta t^{\alpha-1}\mathcal{A}_\alpha\mathcal{L}^{-1}\Big)\psi$

- Useful when $\mathcal{A}_k\mathcal{L}^{-1}$ can be computed, *e.g.* $\mathcal{A}_k = a_k\mathcal{L}^k$
- Reduces to trapezoidal rule for second order schemes

Extension to Metropolized overdamped Langevin dynamics

- Superimpose Metropolis-Hastings correction to discretization of SDE

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

$\longrightarrow$ no bias on invariant measure / stabilization for singular potentials
$\longrightarrow$ error estimates on the diffusion of order Δt

- HMC-like scheme $\widetilde{q}^{n+1} = q^n + \Delta t \, \nabla V \left(q^n + \sqrt{\frac{\Delta t}{2\beta}} \, G^n \right) + \sqrt{\frac{2\Delta t}{\beta}} \, G^n$
    - Error of order $\Delta t^{3/2}$ when the Metropolis-Hastings rule is used
    - Reduced to Δt^2 when a Barker rule is used (replace $\min(1, r)$ by $r/(r+1)$)
    - Requires some time renormalization since rejection rate $\simeq 1/2$
    - Trade-off between increased variance (factor 2) and reduced bias

- Extension to diffusions with multiplicative noise

Numerical results