Metropolized Integrators for SDEs

Nawaf Bou-Rabee

Affiliation: Courant Institute, NYU

Project in collaboration with Eric Vanden-Eijnden

Venue: 2009 Structured Integrators Workshop, Caltech

- Analyze numerical methods for ergodic SDES that:

  (i) ergodic to exact equilibrium distribution of SDE on infinite time intervals; and,
  (ii) strongly converge to solution of SDE on finite time intervals.

- Pure sampling methods can accomplish (i), but do not approximate dynamics

- Integrators can do (ii), but often are divergent on infinite time intervals or ergodic w.r.t different equilibrium distribution

Consider a particle diffusing in $U(x) = x^4/4$ with inverse temperature β

- Overdamped dynamics for this system

$$dY = -Y^3 dt + \sqrt{2\beta^{-1}} dW, \quad Y(0) = x.$$

- $W(t)$ is one-dimensional Brownian motion

- Solution process ergodic with respect to

$$\mu(dx) = Z^{-1} \exp(-\beta x^4/4) dx.$$

Forward-Euler

$$\tilde{X}_{k+1} = \tilde{X}_k - h\tilde{X}_k^3 + \sqrt{2\beta^{-1}}(W(t_{k+1}) - W(t_k)), \quad \tilde{X}_0 = x.$$

- transition density implies irreducibility

$$q_h(x,y) = (4\pi\beta^{-1}h)^{-1/2}\exp\left(-\frac{\left|y - x + hx^3\right|^2}{4\beta^{-1}h}\right)$$

- however, chain is stochastically unstable

- drift is destabilizing in

$$B_h = \{x : |1 - hx^2| > 1\}.$$

- whenever $x \in B_h$ forward euler drift moves particle to higher energy, unlike continuous drift

- ***Metropolis-Hastings*** is a Monte-Carlo method for sampling from a (known) probability distribution.

- Method generates a Markov chain from a given ***proposal chain***.

- Algorithm computes a ***proposal move*** and accepts with probability designed to ensure Metropolized chain samples right distribution.

Metropolized Forward-Euler

- given (X_k, h)
- compute proposal move

$$X_{k+1}^* = X_k - hX_k^3 + \sqrt{2\beta^{-1}}(W(t_{k+1}) - W(t_k))$$

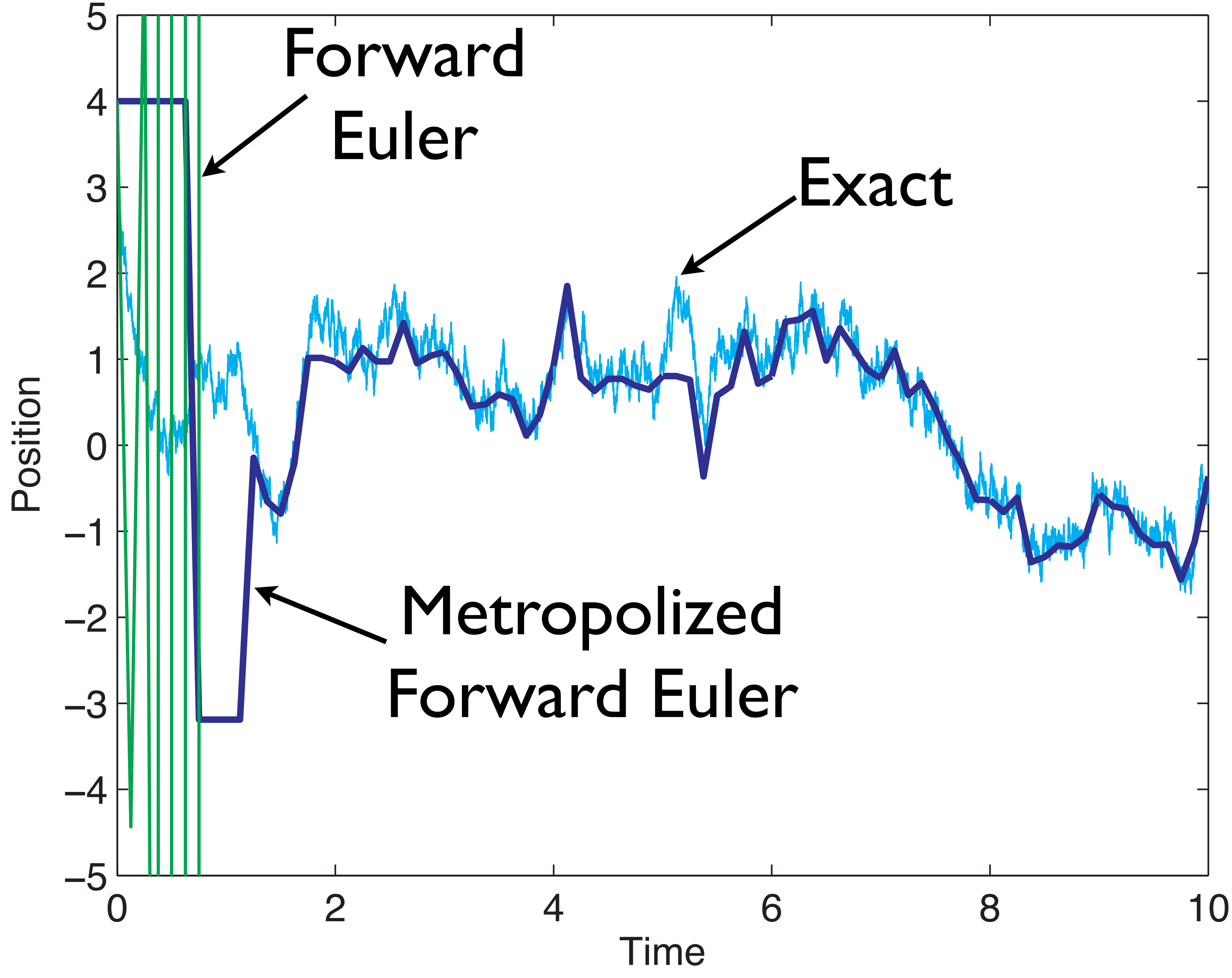
- accept with probability

$$\alpha_h(x,y) = 1 \wedge \frac{q_h(y,x)\pi(y)}{q_h(x,y)\pi(x)}.$$

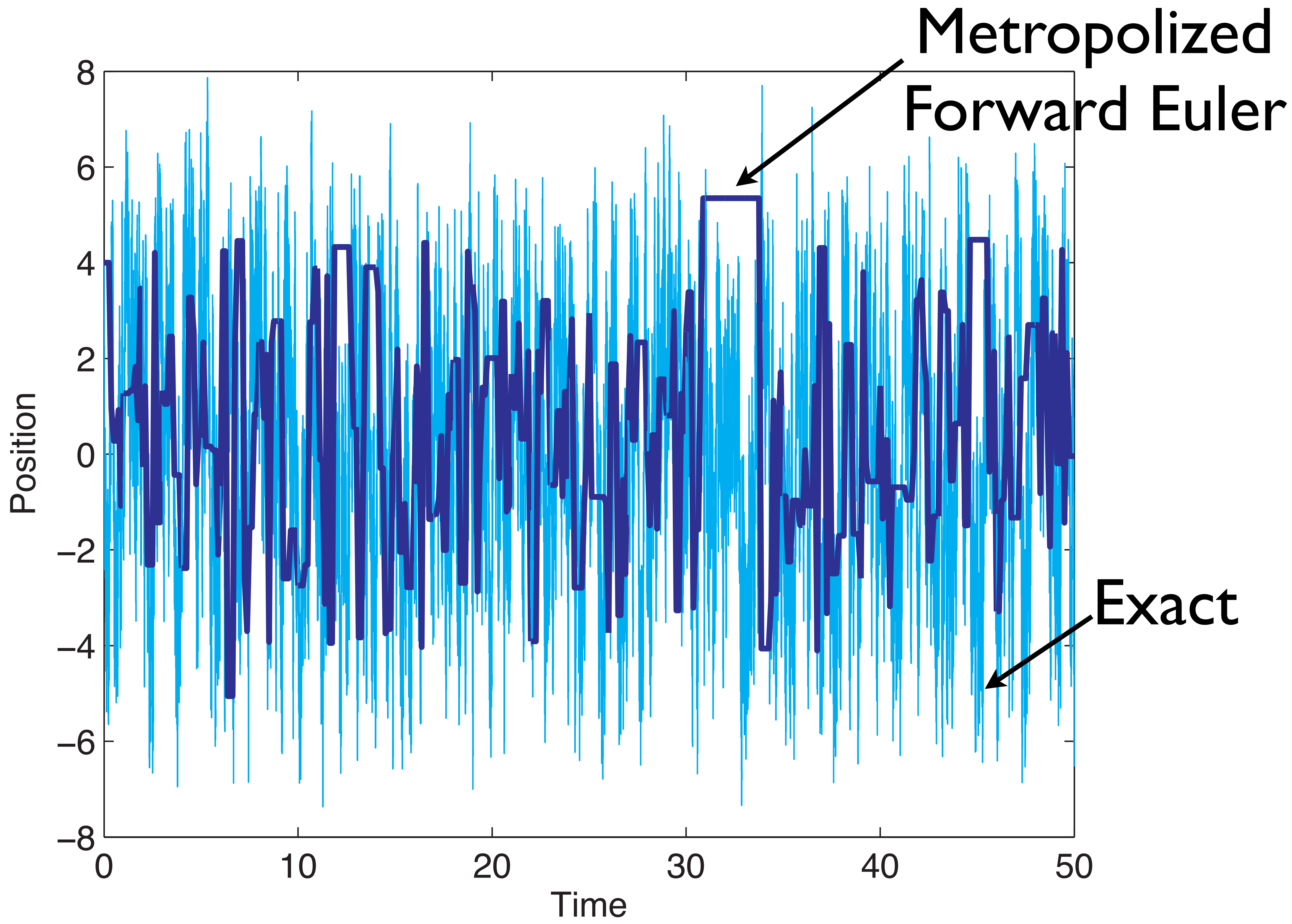
- in other words $\zeta_k \sim U(0,1)$

$$X_{k+1} = \begin{cases} X_{k+1}^* & \text{if } \ \zeta_k < \alpha_h(X_k, X_{k+1}^*) \\ X_k & \text{otherwise} \end{cases}$$

- straightforward to classify as an ergodic chain



Forward
Euler
Exact
Metropolized
Forward Euler
Position
Time
5
4
3
2
1
0
−1
−2
−3
−4
−5
0
2
4
6
8
10


Metropolized
Forward Euler
Exact
Position
Time
8
6
4
2
0
−2
−4
−6
−8
0
10
20
30
40
50

What effect do rejections have on dynamics?

Will answer in more general setting

- overdamped dynamics (nonglobally lipschitz case)

$$d\boldsymbol{Y} = -\nabla U(\boldsymbol{Y})dt + \sqrt{2\beta^{-1}}d\boldsymbol{W}$$

- solution process preserves

$$\pi(\boldsymbol{x}) = Z^{-1}\exp(-\beta U(\boldsymbol{x}))$$

- **U**nadjusted **L**angevin **A**lgorithm (**ULA**)

$$\tilde{\boldsymbol{X}}_{k+1} = \tilde{\boldsymbol{X}}_k - h\nabla U(\tilde{\boldsymbol{X}}_k) + \sqrt{2\beta^{-1}}(\boldsymbol{W}(t_{k+1}) - \boldsymbol{W}(t_k))$$

- discretization is stochastically unstable

Metropolis-Adjusted Langevin Algorithm (MALA)

- given: $(\boldsymbol{X}_k, h)$
- compute a proposal move with ULA

$$\boldsymbol{X}^*_{k+1} = \boldsymbol{X}_k - h\nabla U(\boldsymbol{X}_k) + \sqrt{2\beta^{-1}}(\boldsymbol{W}(t_{k+1}) - \boldsymbol{W}(t_k))$$

- accept or rejection with probability

$$\alpha_h(\boldsymbol{x}, \boldsymbol{y}) = 1 \wedge \frac{q_h(\boldsymbol{y}, \boldsymbol{x})\pi(\boldsymbol{y})}{q_h(\boldsymbol{x}, \boldsymbol{y})\pi(\boldsymbol{x})}.$$

- in sum,

$$\boldsymbol{X}_{k+1} = \begin{cases} \boldsymbol{X}^*_{k+1} & \text{if } \ \zeta_k < \alpha_h(\boldsymbol{X}_k, \boldsymbol{X}^*_{k+1}) \\ \boldsymbol{X}_k & \text{otherwise} \end{cases}$$

- discretization is stochastically stable

Structural Assumptions.

- Uniformly coercive potential energy

$$U(\boldsymbol{x}) \geq K|\boldsymbol{x}|, \ \ \forall \ \boldsymbol{x} \in \mathbb{R}^n.$$

- Local Lipschitz property

$$|\nabla U(\boldsymbol{x}) - \nabla U(\boldsymbol{y})| \leq K(U(\boldsymbol{x}) + U(\boldsymbol{y}))|\boldsymbol{x} - \boldsymbol{y}|, \ \ \ \forall \ \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

- one-sided Lipschitz drift

$$\langle -\nabla U(\boldsymbol{x}) + \nabla U(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \leq K|\boldsymbol{x} - \boldsymbol{y}|^2, \ \ \ \forall \ \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

- geometric ergodicity

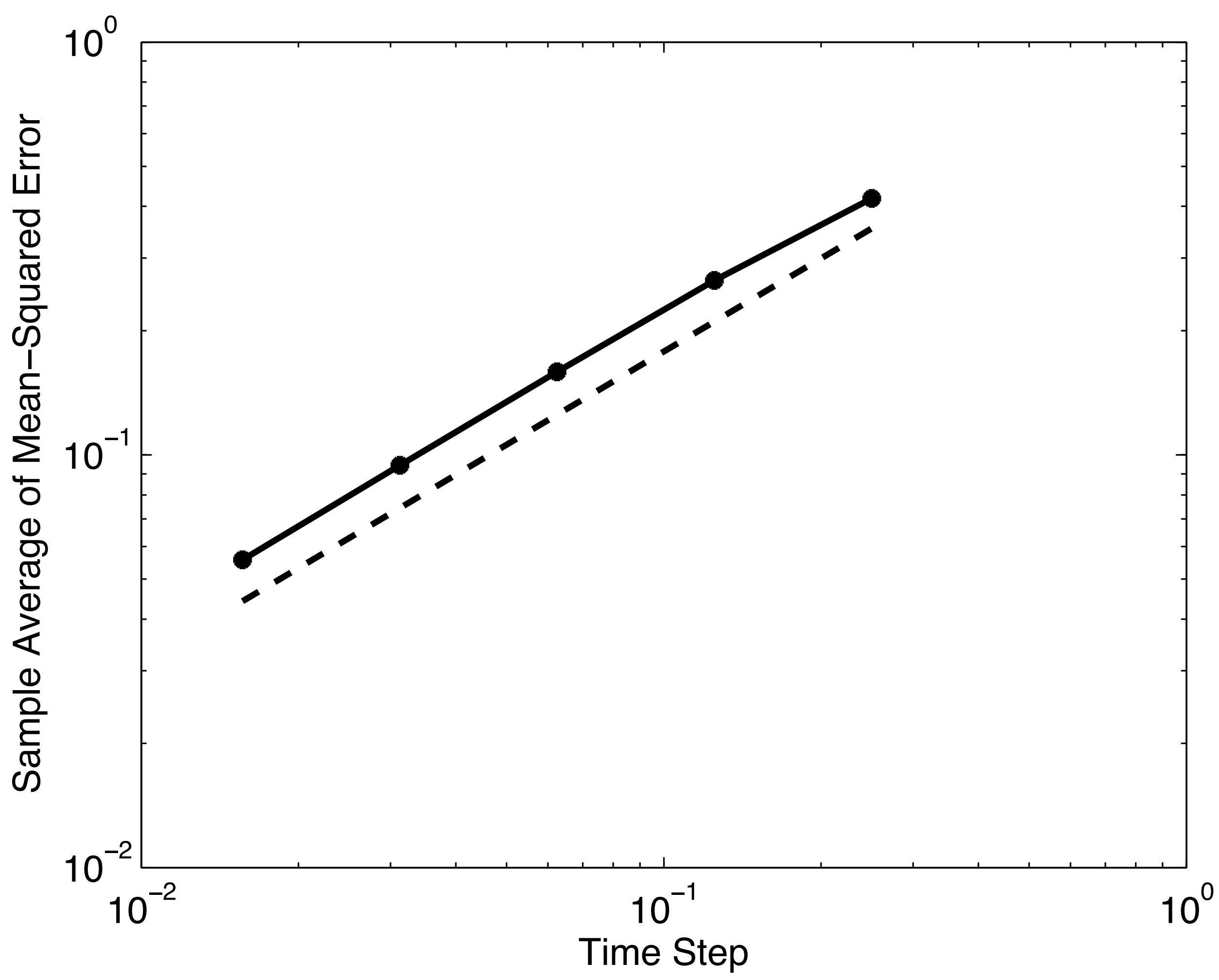
$$L\{U(\boldsymbol{x})^{\ell}\} \leq - \ \delta_{\ell} U(\boldsymbol{x})^{\ell} + M_{\ell}, \ \ \forall \ \boldsymbol{x} \in \mathbb{R}^n.$$

- growth conditions on gradient, Hessian, etc.

$$\left\|D^3 U(\boldsymbol{x})\right\| \vee \left\|D^2 U(\boldsymbol{x})\right\| \vee |\nabla U(\boldsymbol{x})| \leq K(1 + U(\boldsymbol{x})), \ \ \ \forall \ \boldsymbol{x} \in \mathbb{R}^n.$$

Theorem. *For all* $T > 0$, *there exists* $h_c > 0$ *and* $C(T) > 0$ *such that for all* $h < h_c$ *and* $t \in [0, T]$

$$\left(\mathbb{E}_\mu \mathbb{E}^{\boldsymbol{x}} \left\{ \left| \boldsymbol{X}_{\lfloor t/h \rfloor} - \boldsymbol{Y}(t) \right|^2 \right\}\right)^{1/2} \leq C(T) h^{3/4}.$$



Metropolis-Adjusted Langevin Truncated Algorithm (MALTA)

- modify proposal move

$$\boldsymbol{Z}^*_{k+1} = \boldsymbol{Z}_k - h\frac{\nabla U(\boldsymbol{Z}_k)}{1 \vee h|\nabla U(\boldsymbol{Z}_k)|} + \sqrt{2\beta^{-1}}(\boldsymbol{W}(t_{k+1}) - \boldsymbol{W}(t_k))$$

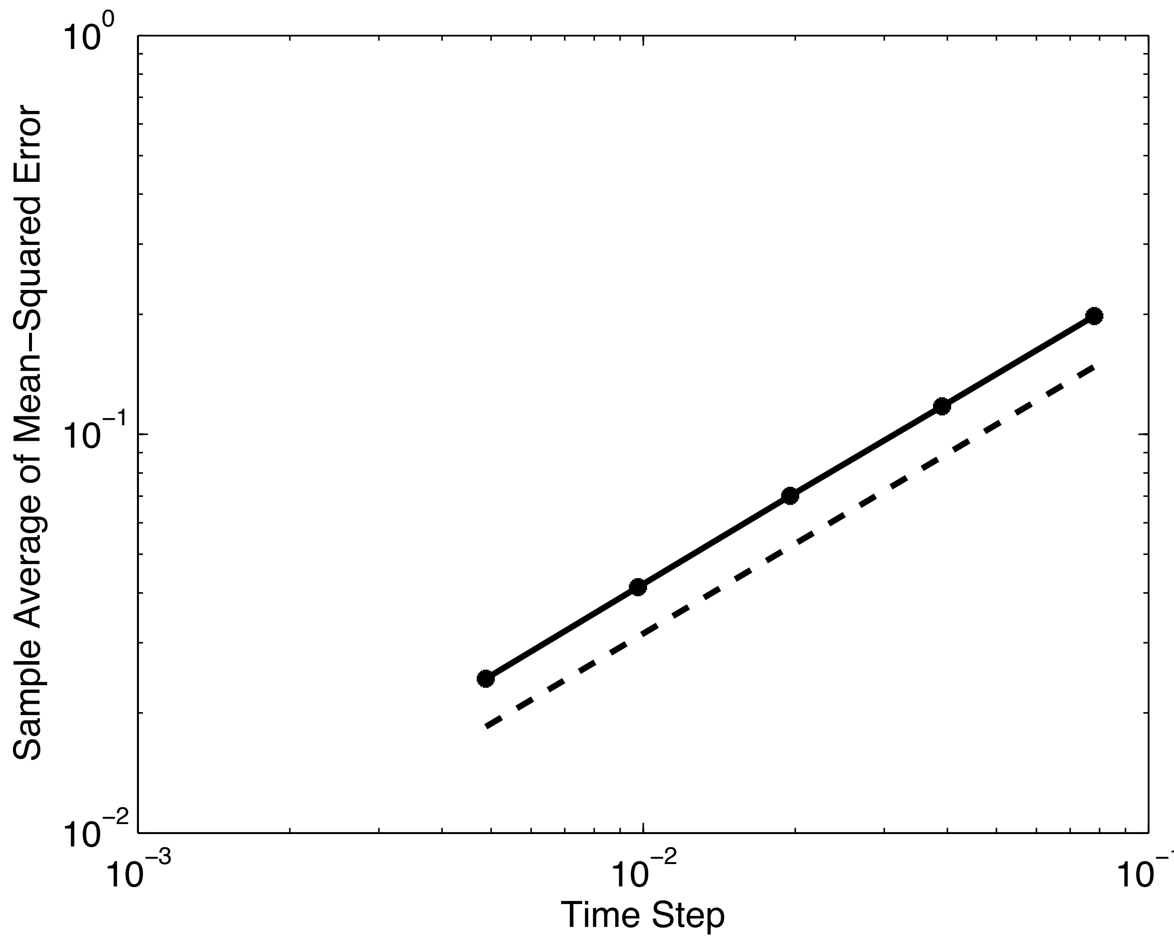
  - when $|\nabla U(\boldsymbol{Z}_k)| < 1/h$ proposal uses ULA drift
  - otherwise drift preserves direction of ULA drift, but normalizes its amplitude

- can show MALTA is *geometrically ergodic*

- how is MALTA related to original diffusion?

Theorem. *For every* $T > 0$ *and* $E_0 > 0$, *there exists an* $h_c(E_0) > 0$ *and* $C(T, E_0) > 0$, *such that for all* $h < h_c$ $\boldsymbol{x} : U(\boldsymbol{x}) \le E_0$ *and* $t \in [0, T]$

$$\left(\mathbb{E}^{\boldsymbol{x}} \left\{ \left| \boldsymbol{Z}_{\lfloor t/h \rfloor} - \boldsymbol{Y}(t) \right|^2 \right\}\right)^{1/2} \le C(T, E_0) h^{3/4}.$$



References on MALA & MALTA

- Roberts and Tweedie [1996]. Exponential Convergence of Langevin distributions and their discrete approximations. Bernoulli, **2**, 341-363.

- Roberts and Tweedie [1996]. Geometric Convergence and Central Limit Theorems for Multidimensional Hastings and Metropolis Algorithms. Biometrika, **1**, 95-110.

- Mengersen and Tweedie [1996]. Rates of Convergence of the Hastings and Metropolis Algorithm. The Annals of Statistics. **24**, 101-121.

- Jarner and Hansen [2000]. Geometric Ergodicity of Metropolis Algorithms. Stochastic Processes and Applications. **85**, 341-361.

- Atchade [2005]. An Adaptive Version for the Metropolis Adjusted Langevin Algorithm with Truncated Drift. Method and Comput in Applied Probability. **8**, 235-254.

- Bou-Rabee and Vanden-Eijnden [2009]. Pathwise Accuracy and Ergodicity of Metroplized Integrators for SDEs. Submitted.

Does pathwise accuracy depend on reversibility or nondegeneracy of noise?

- Inertial Langevin dynamics $H \in C^\infty(\mathbb{R}^{2n}, \mathbb{R})$

$$d\mathbf{Y} = \mathbb{J}\nabla H(\mathbf{Y})dt - \gamma \boldsymbol{C}\nabla H(\mathbf{Y})dt + \sqrt{2\gamma\beta^{-1}}\boldsymbol{C}d\mathbf{W}$$

$$\mathbb{J} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{I} \end{bmatrix}.$$

- write: $\mathbf{Y}(t) = (\boldsymbol{Q}(t), \boldsymbol{P}(t))$
- solution process preserves

$$\mu(d\boldsymbol{q}, d\boldsymbol{p}) = Z^{-1}e^{-\beta H(\boldsymbol{q},\boldsymbol{p})}d\boldsymbol{q}d\boldsymbol{p}$$

- assume, for simplicity, separable Hamiltonian

$$H(\boldsymbol{q},\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T\boldsymbol{M}^{-1}\boldsymbol{p} + U(\boldsymbol{q}),$$

Geometric Langevin Algorithm (GLA)

- Split Langevin into Hamilton's equations:

$$\begin{cases} d\boldsymbol{Q} &= \boldsymbol{M}^{-1}\boldsymbol{P}dt \\ d\boldsymbol{P} &= -\nabla U(\boldsymbol{Q})dt \end{cases}$$

- and, Ornstein-Uhlenbeck equations:

$$\begin{cases} d\boldsymbol{Q} &= 0 \\ d\boldsymbol{P} &= -\gamma \boldsymbol{M}^{-1}\boldsymbol{P}dt + \sqrt{2\beta^{-1}\gamma}d\boldsymbol{W} \end{cases}$$

- Apply variational integrator to approximate solution of Hamilton's equations and use the exact flow for Ornstein-Uhlenbeck.

Variational Integrators

- Given $(\boldsymbol{q}_0, \boldsymbol{p}_0)$ and h

$$\begin{cases} \boldsymbol{p}_0 & = & -D_1 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h), \\ \boldsymbol{p}_1 & = & D_2 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h). \end{cases}$$

- Implicitly defines

$$\theta_h : \quad (\boldsymbol{q}_0, \boldsymbol{p}_0) \mapsto (\boldsymbol{q}_1, \boldsymbol{p}_1),$$

- Discrete Lagrangian is *self-adjoint* if

$$L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h) = L_d(\boldsymbol{q}_1, \boldsymbol{q}_0, h)$$

- e.g., Stormer-Verlet discrete Lagrangian

$$L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h) = \frac{1}{2h}(\boldsymbol{q}_1 - \boldsymbol{q}_0)^T \boldsymbol{M}(\boldsymbol{q}_1 - \boldsymbol{q}_0) - \frac{h}{2}(U(\boldsymbol{q}_0) + U(\boldsymbol{q}_1)).$$

Ornstein-Uhlenbeck Equations

- Use exact flow

$$\psi_{t_k+h,t_k}:$$

$$(\boldsymbol{q},\boldsymbol{p}) \mapsto \left(\boldsymbol{q}, e^{-\gamma \boldsymbol{M}^{-1}h}\boldsymbol{p} + \sqrt{2\beta^{-1}\gamma}\int_{t_k}^{t_k+h} e^{-\gamma \boldsymbol{M}^{-1}(t_k+h-s)} d\boldsymbol{W}(s)\right).$$

- Has the following transition density

$$o_h(\boldsymbol{p}_0,\boldsymbol{p}_1) = \frac{1}{(2\pi)^{n/2}|\det(\boldsymbol{\Sigma}_h)|}\exp\left(-\frac{1}{2}\left(\boldsymbol{p}_1 - e^{-\gamma \boldsymbol{M}^{-1}h}\boldsymbol{p}_0\right)^T \boldsymbol{\Sigma}_h^{-1}\left(\boldsymbol{p}_1 - e^{-\gamma \boldsymbol{M}^{-1}h}\boldsymbol{p}_0\right)\right)$$

$$\boldsymbol{\Sigma}_h = \beta^{-1}\left(\boldsymbol{Id} - \exp(-2\gamma \boldsymbol{M}^{-1}h)\right)\boldsymbol{M}.$$

- transition density does not depend on initial or terminal position

Strang-Type Splitting

$$\tilde{\boldsymbol{X}}_{k+1} := (\tilde{\boldsymbol{Q}}_{k+1}, \tilde{\boldsymbol{P}}_{k+1}) = \psi_{t_k+h, t_k+h/2} \circ \theta_h \circ \psi_{t_k+h/2, t_k}(\tilde{\boldsymbol{Q}}_k, \tilde{\boldsymbol{P}}_k)$$

- Stormer-Verlet based GLA:

$$\begin{cases} \tilde{\boldsymbol{Q}}_{k+1} = \tilde{\boldsymbol{Q}}_k + h\boldsymbol{M}^{-1} e^{-\gamma \boldsymbol{M}^{-1} h/2} \tilde{\boldsymbol{P}}_k - \frac{h^2}{2} \boldsymbol{M}^{-1} \nabla U(\tilde{\boldsymbol{Q}}_k) \\ \qquad + h\sqrt{2\beta^{-1}\gamma} \int_{t_k}^{t_k+h/2} \boldsymbol{M}^{-1} e^{-\gamma \boldsymbol{M}^{-1}(t_k+h/2-s)} d\boldsymbol{W}(s), \\ \tilde{\boldsymbol{P}}_{k+1} = e^{-\gamma \boldsymbol{M}^{-1} h} \tilde{\boldsymbol{P}}_k - \frac{h}{2} e^{-\gamma \boldsymbol{M}^{-1} h/2} \left(\nabla U(\tilde{\boldsymbol{Q}}_k) + \nabla U(\tilde{\boldsymbol{Q}}_{k+1}) \right) \\ \qquad + \sqrt{2\beta^{-1}\gamma} \int_{t_k}^{t_k+h} e^{-\gamma \boldsymbol{M}^{-1}(t_k+h-s)} d\boldsymbol{W}(s). \end{cases}$$

- For globally Lipschitz potential forces, it is first-order accurate and geometrically ergodic.

- Otherwise, it is plagued with the same stochastic instability as forward Euler.

GLA Probability Transition Density

- For any $A \in \mathcal{B}(\mathbb{R}^{2n})$ transition kernel is given by

$$Q_h((\boldsymbol{q}_0, \boldsymbol{p}_0), A) = \int_{\mathbb{R}^{2n} \times A} o_{h/2}(\boldsymbol{p}_0, \boldsymbol{p}_0^*) \cdot o_{h/2}(\boldsymbol{p}_1^*, \boldsymbol{p}_1) \cdot \delta_{\theta_h(\boldsymbol{q}_0, \boldsymbol{p}_0^*)}(d\boldsymbol{q}_1, d\boldsymbol{p}_1^*) d\boldsymbol{p}_0^* d\boldsymbol{p}_1$$

- Zero of Dirac-delta given by

$$\begin{cases} \boldsymbol{p}_0^* &= -D_1 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h), \\ \boldsymbol{p}_1^* &= D_2 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h). \end{cases}$$

- To make it explicit, perform change of variables.

$$q_h((\boldsymbol{q}_0, \boldsymbol{p}_0), (\boldsymbol{q}_1, \boldsymbol{p}_1)) =$$

$$|\det(D_{12} L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h))| \cdot o_{h/2}(\boldsymbol{p}_0, -D_1 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h)) \cdot o_{h/2}(D_2 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h), \boldsymbol{p}_1)$$

- Formula works if variational integrator is implicit.

MAGLA

- compute proposal move

$$(\boldsymbol{Q}^*_{k+1}, \boldsymbol{P}^*_{k+1}) = \psi_{t_k+h, t_k+h/2} \circ \theta_h \circ \psi_{t_k+h/2, t_k}(\boldsymbol{Q}_k, \boldsymbol{P}_k).$$

- accept with probability

$$\alpha_h((\boldsymbol{q}_0, \boldsymbol{p}_0), (\boldsymbol{q}_1, \boldsymbol{p}_1)) = 1 \wedge \frac{q_h((\boldsymbol{q}_1, \boldsymbol{p}_1), (\boldsymbol{q}_0, -\boldsymbol{p}_0))\pi(\boldsymbol{q}_1, \boldsymbol{p}_1)}{q_h((\boldsymbol{q}_0, \boldsymbol{p}_0), (\boldsymbol{q}_1, -\boldsymbol{p}_1))\pi(\boldsymbol{q}_0, \boldsymbol{p}_0)}$$

- in other words

$$\boldsymbol{X}_{k+1} := (\boldsymbol{Q}_{k+1}, \boldsymbol{P}_{k+1}) =$$

$$\begin{cases} (\boldsymbol{Q}^*_{k+1}, \boldsymbol{P}^*_{k+1}) & \text{if } \ \zeta_k < \alpha_h((\boldsymbol{Q}_k, \boldsymbol{P}_k), (\boldsymbol{Q}^*_{k+1}, \boldsymbol{P}^*_{k+1})) \\ \varphi(\boldsymbol{Q}_k, \boldsymbol{P}_k) & \text{otherwise} \end{cases}$$

- stochastically stable, and straightforward to classify as an ergodic chain

Momentum Flip

$$\boldsymbol{X}_{k+1} := (\boldsymbol{Q}_{k+1}, \boldsymbol{P}_{k+1}) =$$

$$\begin{cases} (\boldsymbol{Q}^*_{k+1}, \boldsymbol{P}^*_{k+1}) & \text{if} \;\; \zeta_k < \alpha_h((\boldsymbol{Q}_k, \boldsymbol{P}_k), (\boldsymbol{Q}^*_{k+1}, \boldsymbol{P}^*_{k+1})) \\ \varphi(\boldsymbol{Q}_k, \boldsymbol{P}_k) & \text{otherwise} \end{cases}$$

- momentum flipped twice in the algorithm
  1) if proposal rejected momentum is flipped (local accuracy lost)
  2) acceptance probability involves momentum flip

- motivation: inertial Langevin dynamics is not reversible, but composed with a momentum flip is reversible

Acceptance probability of MAGLA

Lemma. *Consider GLA based on a self-adjoint discrete Lagrangian. Then acceptance probability of MAGLA satisfies*

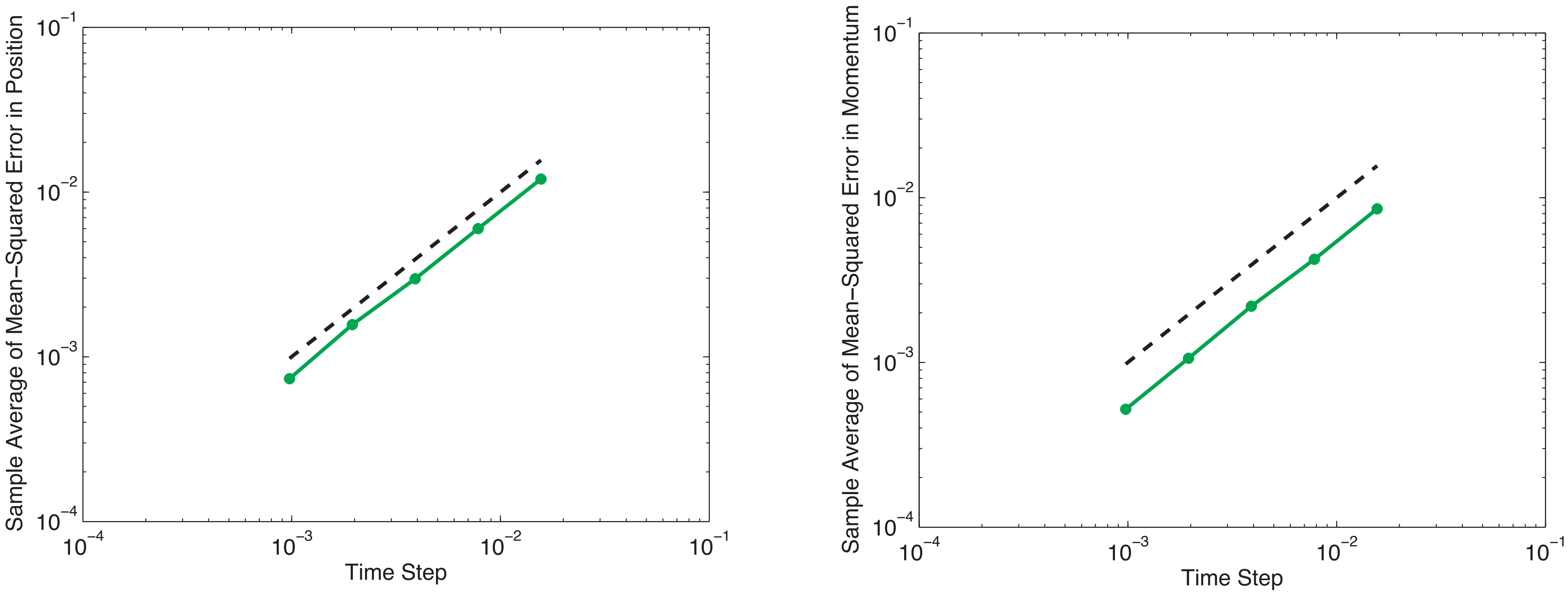
$$\alpha_h((\boldsymbol{q}_0, \boldsymbol{p}_0), (\boldsymbol{q}_1, \boldsymbol{p}_1)) = 1 \wedge \exp\left(-\beta \Delta E(\boldsymbol{q}_0, \boldsymbol{q}_1)\right),$$

where

$$\begin{aligned}\Delta E(\boldsymbol{q}_0, \boldsymbol{q}_1) = &\frac{1}{2} D_2 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h)^T \boldsymbol{M}^{-1} D_2 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h) + U(\boldsymbol{q}_1) \\ &- \frac{1}{2} D_1 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h)^T \boldsymbol{M}^{-1} D_1 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1, h) - U(\boldsymbol{q}_0).\end{aligned}$$

Theorem. *For all* $T > 0$, *there exists* $h_c > 0$ *and* $C(T) > 0$ *such that for all* $h < h_c$ *and* $t \in [0, T]$

$$\left(\mathbb{E}_\mu \mathbb{E}^{\boldsymbol{x}} \left\{ \left| \boldsymbol{X}_{\lfloor t/h \rfloor} - \boldsymbol{Y}(t) \right|^2 \right\} \right)^{1/2} \leq C(T) h.$$



Proof Relies on Two Main Ingredients

(1) Local accuracy of Metropolized Integrator, e.g.,

$$\mathbb{E}^{\boldsymbol{x}}\left\{\left|\boldsymbol{X}_1-\boldsymbol{Y}(h)\right|^2\right\}=$$

$$\underbrace{\mathbb{E}^{\boldsymbol{x}}\{\left|\boldsymbol{X}_1^*-\boldsymbol{Y}(h)\right|^2\alpha_h(\boldsymbol{x},\boldsymbol{X}_1^*)\}}_{\text{Accepted Proposal Move}}$$

$$+\underbrace{\mathbb{E}^{\boldsymbol{x}}\{\left|\varphi(\boldsymbol{x})-\boldsymbol{Y}(h)\right|^2(1-\alpha_h(\boldsymbol{x},\boldsymbol{X}_1^*))\}}_{\text{Rejected Proposal Move}}$$

(2) Bounds on moments of Metropolized integrator uniform in time

Local accuracy of MALA

- first term bounded by

$$\underbrace{\mathbb{E}^{\boldsymbol{x}}\{|\boldsymbol{X}_1^* - \boldsymbol{Y}(h)|^2 \alpha_h(\boldsymbol{x}, \boldsymbol{X}_1^*)\}}_{\text{Accepted Proposal Move}} \le \mathbb{E}^{\boldsymbol{x}}\left\{|\boldsymbol{X}_1^* - \boldsymbol{Y}(h)|^2\right\} \le O(h^3)$$

- second term bounded by

$$\underbrace{\mathbb{E}^{\boldsymbol{x}}\{|\varphi(\boldsymbol{x}) - \boldsymbol{Y}(h)|^2 (1 - \alpha_h(\boldsymbol{x}, \boldsymbol{X}_1^*))\}}_{\text{Rejected Proposal Move}}$$

$$\le \underbrace{\mathbb{E}^{\boldsymbol{x}}\{|\varphi(\boldsymbol{x}) - \boldsymbol{Y}(h)|^4\}^{1/2}}_{O(h)} \underbrace{\mathbb{E}^{\boldsymbol{x}}\{(1 - \alpha_h(\boldsymbol{x}, \boldsymbol{X}_1^*))^2\}^{1/2}}_{O(h^{3/2})}$$

Local accuracy of Verlet based MAGLA

- first term bounded by accuracy of GLA

$$\underbrace{\mathbb{E}^{\boldsymbol{x}}\{|\boldsymbol{X}_1^* - \boldsymbol{Y}(h)|^2 \alpha_h(\boldsymbol{x}, \boldsymbol{X}_1^*)\}}_{\text{Accepted Proposal Move}} \le \mathbb{E}^{\boldsymbol{x}} \left\{ |\boldsymbol{X}_1^* - \boldsymbol{Y}(h)|^2 \right\} \le O(h^3)$$

- second term bounded by

$$\underbrace{\mathbb{E}^{\boldsymbol{x}}\{|\varphi(\boldsymbol{x}) - \boldsymbol{Y}(h)|^2 (1 - \alpha_h(\boldsymbol{x}, \boldsymbol{X}_1^*))\}}_{\text{Rejected Proposal Move}}$$

$$\le \underbrace{\mathbb{E}^{\boldsymbol{x}}\{|\varphi(\boldsymbol{x}) - \boldsymbol{Y}(h)|^4\}^{1/2}}_{O(1)} \underbrace{\mathbb{E}^{\boldsymbol{x}}\{(1 - \alpha_h(\boldsymbol{x}, \boldsymbol{X}_1^*))^2\}^{1/2}}_{O(h^3)}$$

References on GLA

- Bou-Rabee and Owhadi [2009]. Geometric Langevin Algorithm. Submitted.

- Vanden-Eijnden and Ciccotti [2006]. Second-Order Integrators for Langevin Equations with Holonomic Constraints. Chem. Phys. Letters, **429**, 310-316.

- Pavliotis, Stuart and Zygalakis [2009]. Calculating Effective Diffusivities in the Limit of Vanishing Diffusion. preprint.

- Shardlow [2003]. Splitting in Dissipative Particle Dynamics. SIAM J. of Sci. Comp., **24**, 12671282.

- Serrano, De Fabritis, Espanol, and Coveney [2006]. A Stochastic Trotter Scheme for Dissipative Particle Dynamics. Mathematics and Computers in Simulation. **72**, 190-194.

- Skeel and Izaguirre [2002]. An Impulse Integrator for Langevin Dynamics. Mol. Phys., **100**, 3885-3891.

References on MAGLA

- Secama, Lelievre, Stoltz, Cances and Caffarel [2006]. An Efficient Sampling Algorithm for Variational Monte Carlo. J. Chem. Phys., **125**, 114105.

- Bussi and Parinello [2007]. Accurate Sampling using Langevin Dynamics. Physical Review E, **75**, 056707.

- Akhmatskaya, Bou-Rabee, and Reich [2009]. A Comparison of GHMC with and without Momentum Flip. J. of Comp. Physics, doi: 10.101/j.jcp.2008.12.014.

- Bou-Rabee and Vanden-Eijnden [2009]. Pathwise Accuracy and Ergodicity of Metroplized Integrators for SDEs. Submitted.

Outlook and Generalizations

Higher-Order Discretizations of Overdamped and Inertial Langevin

- Milstein and Tretyakov [2004]. Stochastic Numerics for Mathematical Physics. Springer.

Adaptive integrators

- Lamba, Mattingly, and Stuart [2007]. An Adaptive Euler-Maruyama Scheme for SDEs: Convergence and Stability. IMA J. of Num. Anal., **27**, 479-506.

Multistep integrators

- Lew, Marsden, West, and Ortiz [2003]. Asynchronous Variational Integrators. Arch. Rat. Mech. Anal., **167**, 85-145.

Inertial Langevin with Constraints

- Vanden-Eijnden and Ciccotti [2006]. Second-Order Integrators for Langevin Equations with Holonomic Constraints. Chem. Phys. Letters, **429**, 310-316.

Nose-Hoover-Langevin

- Leimkuhler and Reich [2009]. A Metropolis-Ajusted Nose-Hoover Thermostat. To appear in M2AN.

Energy-Stepping Integrators with Terraced Potential (Ortiz et al.)