Bayesian static parameter estimation using Multilevel Monte Carlo

Kody Law

joint with A. Jasra (NUS), K. Kamatani (Osaka), & Y. Zhou

MANCHESTER 1824
The University of Manchester

UQ18, Garden Grove, CA, USA

April 18, 2018

Orientation

Aim: Approximate expectations with respect to a target probability distribution η_∞, which needs to be approximated by some η_L, *and the un-normalized target can only be evaluated up to a non-negative unbiased estimator.*

Solution: Apply an approximate coupling strategy so that the multilevel Monte Carlo (MLMC) method can be used within a pseudo-marginal algorithm [B02, AR08, ADH10].

- MLMC methods *reduce cost to* error$= \mathcal{O}(\varepsilon)$ [G08];
- Recently this methodology has been applied to *inverse UQ*, mostly in cases where η_L can be evaluated up to a normalizing constant [HSS13, DKST15, HTL16, BJLTZ17].
- Here it is assumed that we cannot,
- however we can construct a non-negative and unbiased estimator of the un-normalized target γ_L [JKLZ18].

Orientation

Aim: Approximate expectations with respect to a target probability distribution η_∞, which needs to be approximated by some η_L, *and the un-normalized target can only be evaluated up to a non-negative unbiased estimator*.

Solution: Apply an approximate coupling strategy so that the multilevel Monte Carlo (MLMC) method can be used within a pseudo-marginal algorithm [B02, AR08, ADH10].

- MLMC methods *reduce cost to* error$= \mathcal{O}(\varepsilon)$ [G08];
- Recently this methodology has been applied to *inverse UQ*, mostly in cases where η_L can be evaluated up to a normalizing constant [HSS13, DKST15, HTL16, BJLTZ17].
- Here it is assumed that we cannot,
- however we can construct a non-negative and unbiased estimator of the un-normalized target γ_L [JKLZ18].

Example: expectation for SDE [G08]

Estimation of expectation of solution of intractable stochastic differential equation (SDE).

$$dX = f(X)dt + \sigma(X)dW, \quad X_0 = x_0 .$$

Aim: estimate $\mathbb{E}(g(X_T))$.
We need to

(1) Approximate, e.g. by Euler-Maruyama method with resolution h:

$$X_{n+1} = X_n + hf(X_n) + \sqrt{h}\sigma(X_n)\xi_n, \quad \xi_n \sim N(0,1).$$

(2) Sample $\{X_{N_T}^{(i)}\}_{i=1}^{N}$, $N_T = T/h$.

Assumptions

(A1) $\forall\, y \in \mathsf{T}, \exists\, C > 0$ such that $\forall\, x \in \mathsf{S}, \theta \in \Theta$,

$$C \leq g_\theta(x, y) \leq C^{-1}.$$

And $\forall\, y \in \mathsf{T}$, $g_\theta(x, y)$ is globally Lipschitz on $\mathsf{S} \times \Theta$.

(A2) $\forall\, 0 \leq k \leq n$, $q \in \{1, 2\}$, $\exists\, \beta > 0$ such that $\forall$ $\varphi \in \mathcal{B}_b(\Theta \times \mathsf{S}^{k+1}) \cap \mathrm{Lip}(\Theta \times \mathsf{S}^{k+1})$ $\exists\, C > 0$

$$\left(\int_{\Theta \times \mathsf{S}^{2k+2}} |\varphi(\theta) - \varphi(\theta)|^q \Pi(\mathrm{d}\theta)\nu_\theta(\mathrm{d}z_0) \prod_{p=1}^{k} Q_{\theta,h,h'}(z_{p-1}, \mathrm{d}z_p) \right)^{3-q} \leq C(h')^\beta.$$

(A3) Suppose that $\forall\, n > 0$, $\exists\, \xi \in (0, 1)$ and $\nu \in \mathcal{P}(\mathsf{W})$ such that for each $w \in \mathsf{W}$, $\varphi \in \mathcal{B}_b(\mathsf{W}) \cap \mathrm{Lip}(\mathsf{W})$, h, h':

$$\int_{\mathsf{W}} \varphi(w') K(w, dw') \geq \xi \int_{\mathsf{W}} \varphi(v) \nu(dv).$$

K is η-reversible, that is,
$\int_{w \in B} \eta(dw) K(w, A) = \int_{w \in A} \eta(dw) K(w, B)$ for any $A, B \in \mathcal{W}$.

Main result

Theorem (JKLZ18)

Assume (A1-3). Then $\forall\ n > 0$, $\exists\ \beta > 0$ *such that* $\forall$ $\varphi \in \mathcal{B}_b(\Theta \times \mathsf{S}^{n+1}) \cap \mathit{Lip}(\Theta \times \mathsf{S}^{n+1})$ $\exists\ C > 0$ *such that*

$$\mathbb{E}[\bar{E}_l^{N_l}(\varphi)^2] \leq \frac{Ch_l^{\beta}}{N_l}.$$

(A4) $\exists\, \gamma, \alpha, C > 0$ such that the cost to simulate $E_l^{N_l}$ is controlled by $\mathrm{C}(E_l^{N_l}) \leq C N_l h_l^{-\gamma}$, and the bias is controlled by

$$|\mathbb{E}_{\Pi_{h_L}}(\varphi(\theta, X_{0:n})) - \mathbb{E}_{\Pi_0}(\varphi(\theta, X_{0:n}))| \leq C h_L^{\alpha}\,.$$

Corollary

Assume (A1-4). $\forall\, n > 0$ *and* $\varphi \in \mathcal{B}_b(\Theta \times S^{n+1}) \cap Lip(\Theta \times S^{n+1})$ $\exists$ $C > 0$ *such that* $\forall\, \epsilon > 0$ *one can choose* $(L, \{N_l\}_{l=1}^{L})$ *so*

$$\mathbb{E}\left[\,|\sum_{l=0}^{L} E_l^{N_l}(\varphi) - \mathbb{E}_{\Pi_0}(\varphi(\theta, X_{0:n}))|^2\right] \leq C\epsilon^2\,,$$

with a total cost (per time step)

$$\mathrm{COST} \leq C \begin{cases} \epsilon^{-2}, & \text{if} \quad \beta > \gamma, \\ \epsilon^{-2}|\log(\epsilon)|^2, & \text{if} \quad \beta = \gamma, \\ \epsilon^{-\left(2 + \frac{\gamma - \beta}{\alpha}\right)}, & \text{if} \quad \beta < \gamma. \end{cases}$$

Langevin SDE

$$\mathrm{d}X_t = \frac{1}{2}\nabla \log \pi(X_t)\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad X_0 = x_0$$

$$Y_k | X_k \sim \mathcal{N}(0, \tau^2 \exp X_k)\,,$$

$$\theta \sim \mathcal{G}(1, 1)\,,$$

$$\sigma \sim \mathcal{G}(1, 0.5)\,.$$

- $\pi(x)$ denote the probability density function of a Student's t-distribution with θ degrees of freedom.
- $x_0 = 0$.
- 1,000 observations simulated with $\theta = 10$, $\sigma = 1$, and $\tau^2 = 1$.

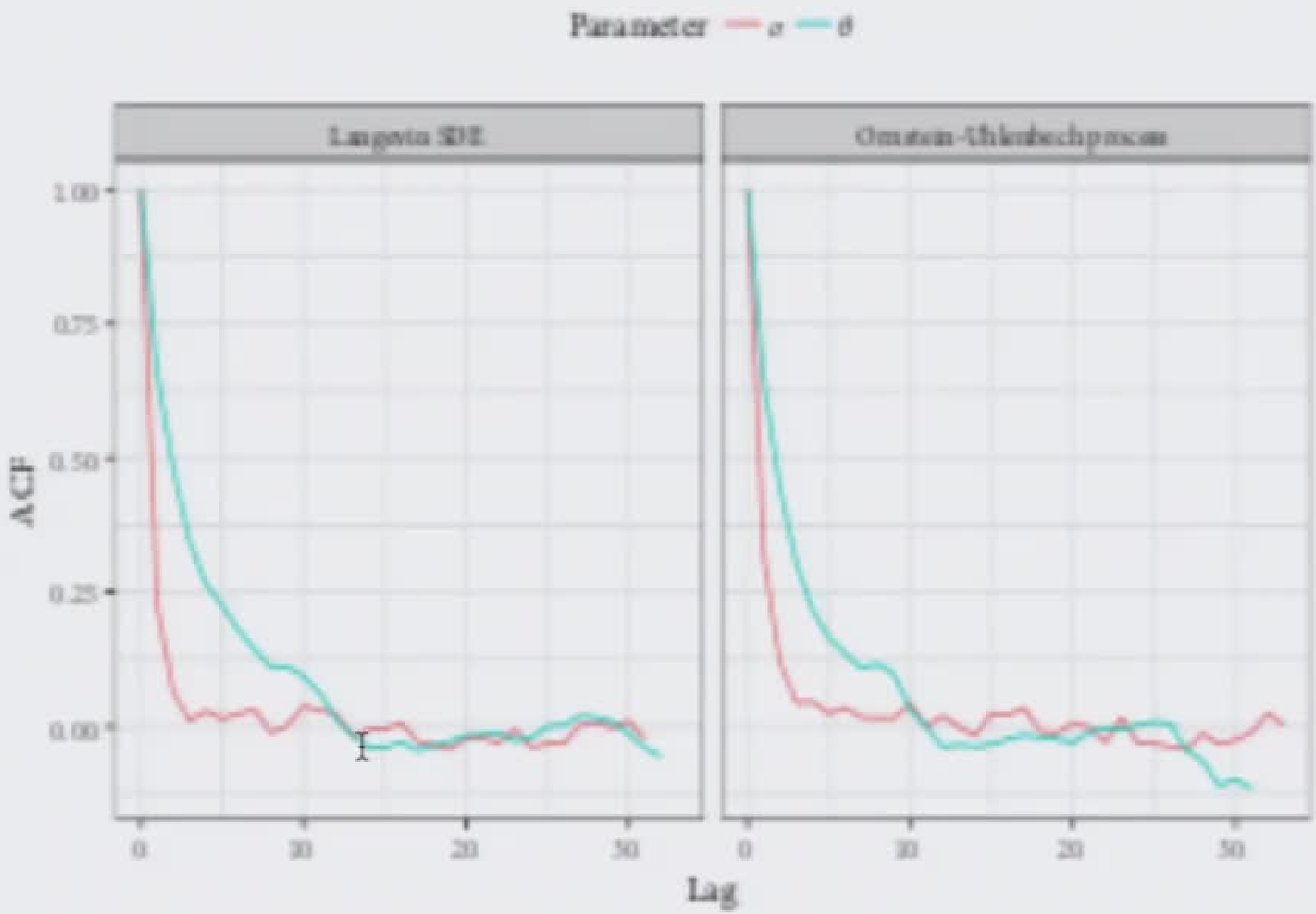

Figure: Autocorrelation of a typical PMCMC chain.

Langevin SDE

$$
\begin{aligned}
\mathrm{d}X_t &= \frac{1}{2}\nabla \log \pi(X_t)\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad X_0 = x_0 \\
Y_k | X_k &\sim \mathcal{N}(0, \tau^2 \exp X_k)\,, \\
\theta &\sim \mathcal{G}(1, 1)\,, \\
\sigma &\sim \mathcal{G}(1, 0.5)\,.
\end{aligned}
$$

- $\pi(x)$ denote the probability density function of a Student's t-distribution with θ degrees of freedom.
- $x_0 = 0$.
- 1,000 observations simulated with $\theta = 10$, $\sigma = 1$, and $\tau^2 = 1$.

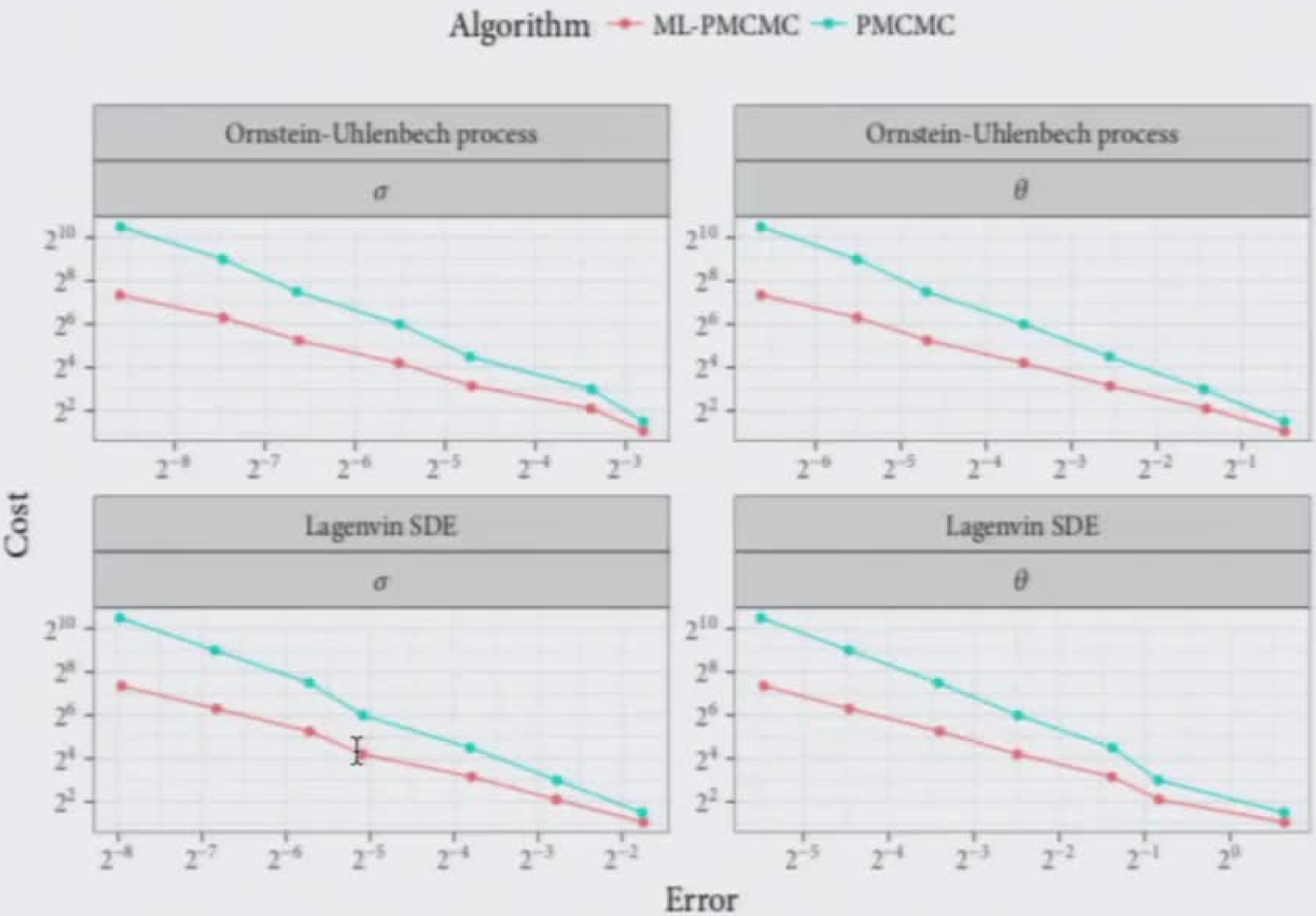

Figure: Cost vs. MSE for the 2 parameters for each of the 2 SDEs.

Outline

Outline

Model	Parameter	ML-PMCMC	PMCMC
OU	θ	-1.022	-1.463
	σ	-1.065	-1.522
Langevin	θ	-1.060	-1.508
	σ	-1.023	-1.481

Table: Estimated rates of convergence of MSE with respect to cost for various parameters, fitted to the curves.

Summary

- New approximate coupling strategy can be used to apply MLMC to PMCMC for static parameter estimation [JKLZ18.i].
- Same strategy can be employed for multi-index MCMC [JKLZ18.ii].
- In progress: SMC^2 [JLX18+]

References

- **[G08]**: Giles. "Multilevel Monte Carlo path simulation." Op. Res., 56, 607-617 (2008).
- **[H00]** Heinrich. "Multilevel Monte Carlo methods." LSSC proceedings (2001).
- **[CGST11]** Cliffe, Giles, Scheichl, & Teckentrup. "MLMC and applications to elliptic PDEs." Computing and Visualization in Science, 14(1), 3 (2011).
- **[D04]**: Del Moral. "Feynman-Kac Formulae." Springer: New York (2004).
- **[B02]** Beaumont. "Estimation of population growth..." Genetics 164(3) 1139-1160 (2003).
- **[AR08]** Andrieu & Roberts. "The pseudo-marginal approach." Annals of Stat. 37(2) 697-725 (2009).
- **[ADH10]** Andrieu, Doucet, and Holenstein. "Particle MCMC methods." JRSSB 72(3) 269-342 (2010).

References

- **[JKLZ18.i]**: Jasra, Kamatani, Law, Zhou. "MLMC for static Bayesian parameter estimation." SISC 40, A887-A902 (2018).
- **[JKLZ18.ii]**: Jasra, Kamatani, Law, Zhou. "A Multi-Index Markov Chain Monte Carlo Method." IJUQ 8(1), 61-73 (2018).
- **[JLX18+]**: Jasra, Law, Xu. "Multi-index SMC^2." In preparation.
- **[BJLTZ15]**: Beskos, Jasra, Law, Tempone, Zhou. "Multilevel Sequential Monte Carlo samplers." SPA 127:5, 1417–1440 (2017).
- **[HSS13]**: Hoang, Schwab, Stuart. Inverse Prob., 29, 085010 (2013).
- **[DKST13]**: Dodwell, Ketelsen, Scheichl, Teckentrup. "A hierarchical MLMCMC algorithm." SIAM/ASA JUQ 3(1) 1075-1108 (2015).

References

- **[CHLNT17]** Chernov, Hoel, Law, Nobile, Tempone. "Multilevel ensemble Kalman filtering for spatio-temporal processes." arXiv:1608.08558 (2017).
- **[JLS17]** Jasra, Law, Suciu. "Advanced Multilevel Monte Carlo Methods." arXiv:1704.07272 (2017).
- **[DJLZ16]**: Del Moral, Jasra, Law, Zhou. "Multilevel Sequential Monte Carlo samplers for normalizing constants." ToMACS 27(3), 20 (2017).
- **[JKLZ15]**: Jasra, Kamatani, Law, Zhou. "Multilevel particle filter." SINUM 55(6), 3068–3096 (2017).
- **[JLZ16]**: Jasra, Law, and Zhou. "Forward and Inverse UQ MLMC Algorithms for an Elliptic Nonlocal Equation." IJUQ 6(6), 501–514 (2016).
- **[HLT15]**: Hoel, Law, Tempone. "Multilevel ensemble Kalman filter." SINUM 54(3), 1813–1839 (2016).