区分確定的モンテカルロ法の拡散極限と早期収束診断

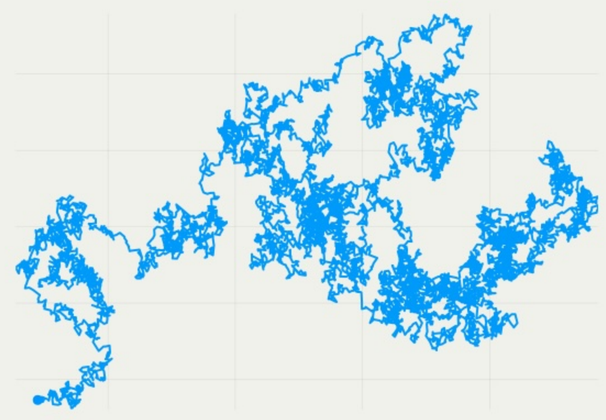
司馬博文　総合研究大学院大学　5年一貫博士課程4年

PDMP とは何か？

運動論的なダイナミクスで速い収束を狙う

	Drift	Diffusion	Jump	$t\in[0,\infty)$
Diffusion	$dX_t = a(X_t)\,dt$	$+\, b(X_t)\,dB_t$		
PDMP	$dX_t = a(X_t)\,dt$		$+\displaystyle\int_{\mathbb{R}^d} c(X_{t-},u)\eta(dtdu)$	

Piecewise Deterministic　　Markov Process



Langevin 拡散

- SDE で記述される過程
  → 数値離散化で誤差が入る
- 平衡状態のダイナミクス

vs.

Randomized Hamiltonian Monte Carlo

- 分子運動論的なダイナミクス
  → ODE としての離散化はいくらか簡単
- 非平衡状態のダイナミクス
  → より速い緩和

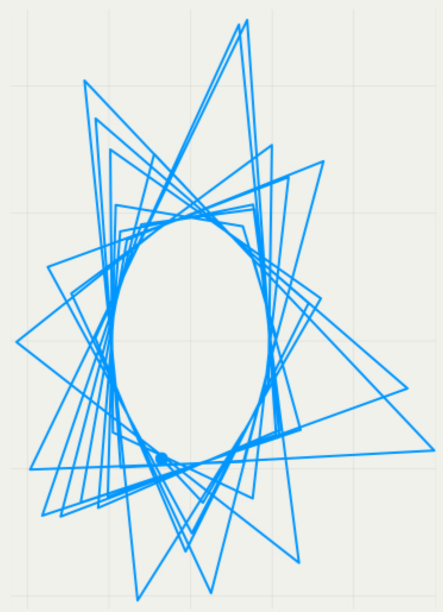
どの PDMP がより速い？

*今回は区分線型な手法に限る

Bouchard-Côté, Vollmer & Doucet (2018) The Bouncy Particle Sampler

Michel, Durmus & Sénécal (2020) Forward Event-Chain Monte Carlo

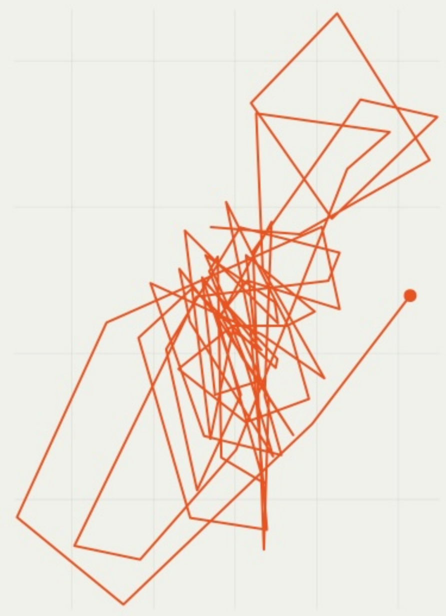
BPS with ρ=0.0



BPS with ρ=1.0

BPS with ρ=10.0

FECMC with $\rho=0$

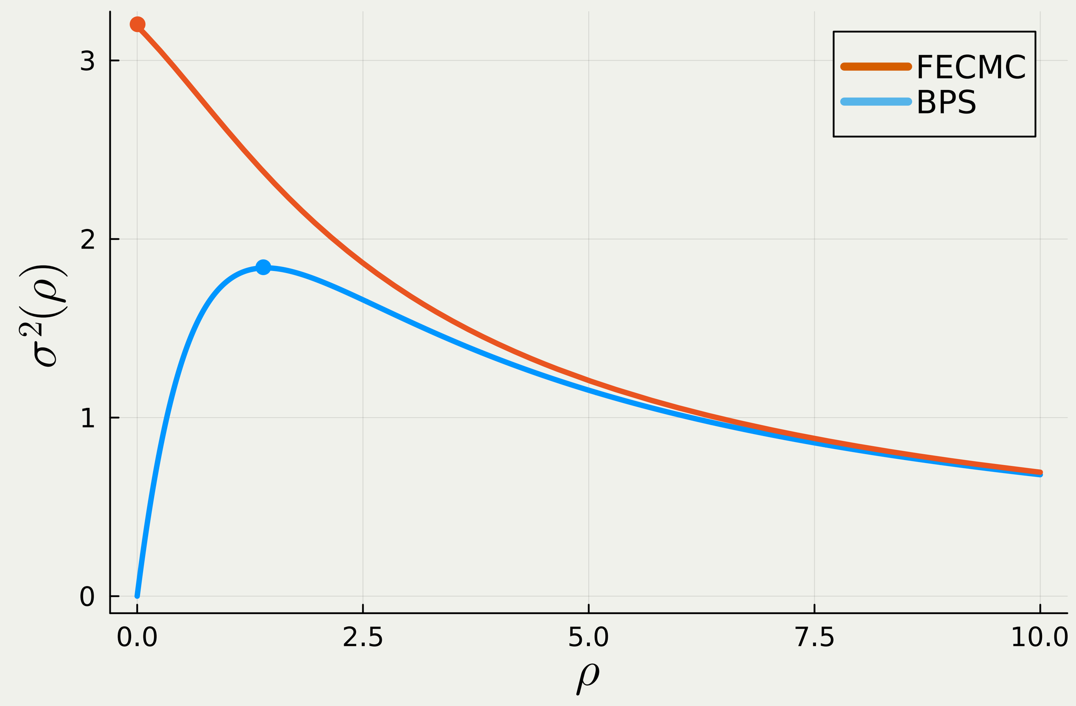


ハイパーパラメータ ρ をどう選ぶ？　本当に速くなってる？

標準 Gauss 分布

$$\pi(x)\propto e^{-U^{(d)}(x)}\qquad U^{(d)}(x)=\frac{\|x\|_2^2}{2}$$

について高次元極限 $d\to\infty$ を考える



$\sigma^2\approx$ ポテンシャル $U^{(d)}(X_t)$ の極限速度

スケーリング解析: $d\to\infty$ 極限でのダイナミクスを比較

Theorem 3.7 & 3.9　BPS, FECMC のポテンシャルの過程は Ornstein-Uhlenbeck 拡散に収束する

$$Y_t^{(d)} := \frac{2U^{(d)}(X^d_{dt})-d}{\sqrt{d}} \implies \mathrm{d}Y_t = -\frac{\sigma^2(\rho)}{4}Y_t\,\mathrm{d}t+\sigma(\rho)\,\mathrm{d}B_t$$

Theorem 4.1

$$\sigma^2_{\mathrm{FECMC}}(\rho)=\sqrt{\frac{32}{\pi}}\left(1-\frac{(\rho^2-\rho\sqrt{\frac{\pi}{2}}+\Omega(\rho))^2}{\rho^4\Omega(\rho)(2-\Omega(\rho))}\right)$$

$$\sigma^2_{\mathrm{BPS}}(\rho)=\frac{8}{\rho^4}\left(\rho^3-\rho^2\sqrt{\frac{8}{\pi}}+\rho-\sqrt{\frac{8}{\pi}}\frac{((1+\rho^2)\Omega(\rho)-\rho^2)^2}{\Omega(2\rho)}\right)$$

$\Omega(\rho):=\rho e^{\frac{\rho^2}{2}}\int_\rho^\infty e^{-\frac{t^2}{2}}\,\mathrm{d}t$: exponentially scaled complementary error function

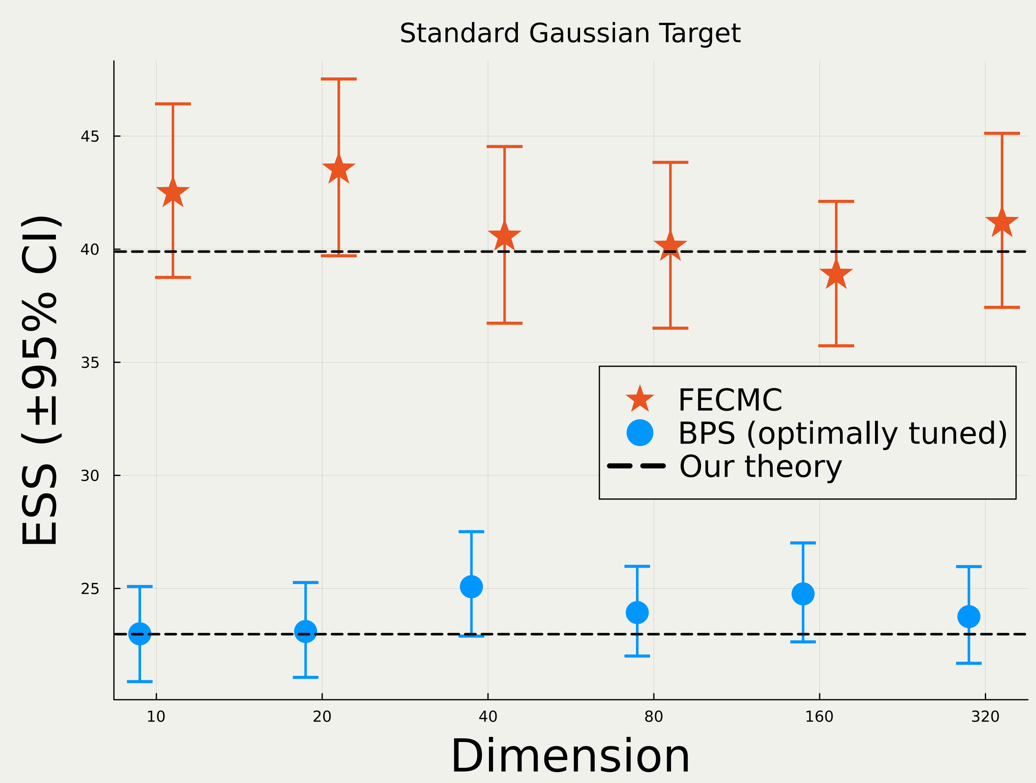
Corollary 4.3

$$\sigma^2_{\mathrm{FECMC}}(0)=\sqrt{\frac{32}{\pi}}$$

ESS: A benchmark　$U^{(d)}$ の有効サンプルサイズ (ESS) は

$$\mathrm{ESS}(U)\approx T\frac{\sigma^2(\rho)}{8}$$

$\mathrm{ESS}/T\approx$ 平均自乗誤差 (MSE: Mean Squared Error)



計算複雑性のスケーリング

	$\nabla U(x)$ evals / ESS	$\nabla U(x)$	総計
Potential			
BPS / FECMC	$O(d)$	$O(d^{0\sim1})$	$O(d^{1\sim2})$
Hamiltonian Monte Carlo		未解決	
Marginal coordinate			
Random Walk Metropolis	$O(d)$	$O(d)$	$O(d^2)$
Metropolis-adjusted Langevin	$O(d^{1/3})$	$O(d)$	$O(d^{4/3})$
Hamiltonian Monte Carlo	$O(d^{1/4})$	$O(d)$	$O(d^{5/4})$

効率的な収束診断

一般のターゲットでは σ^2 は不明
→ シミュレーション中に on the fly で推定

時間微分の過程に注目する

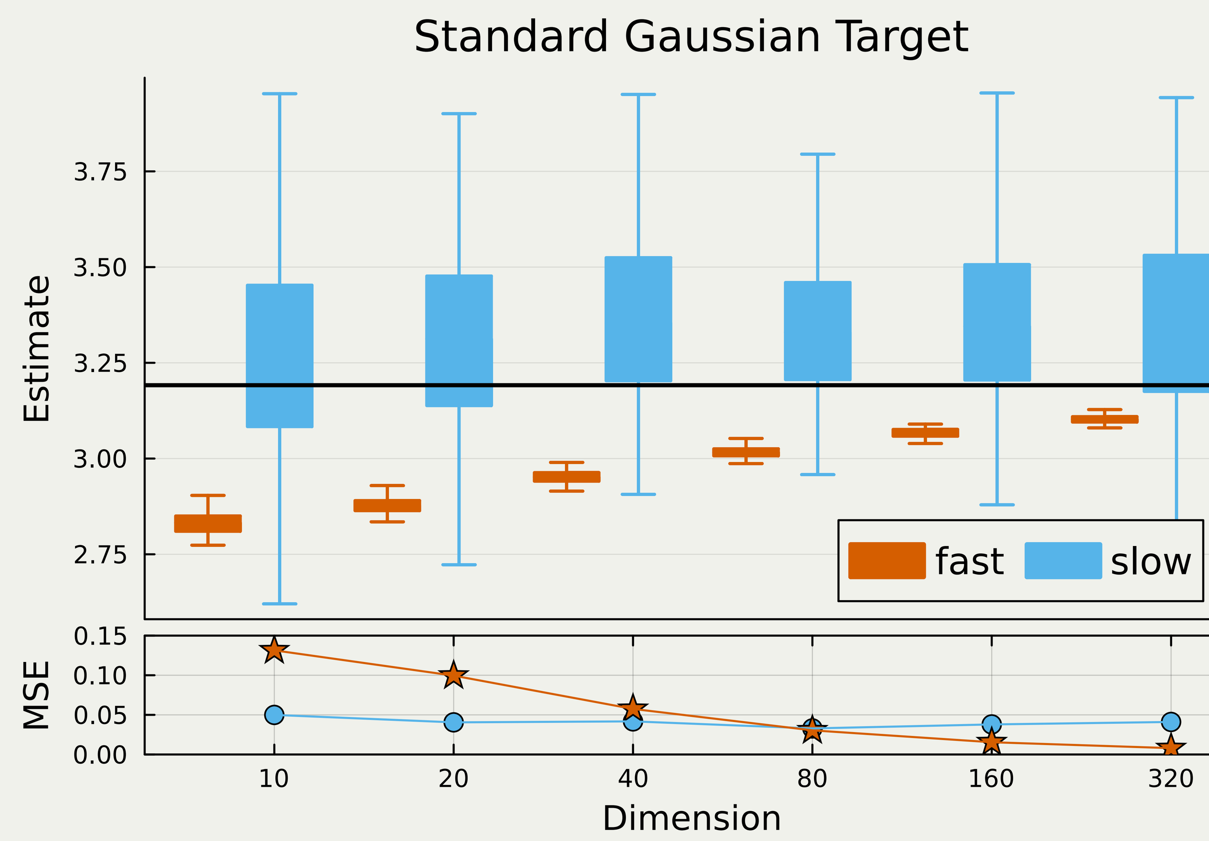
$$\frac{\mathrm{d}}{\mathrm{d}t}h(X_t)=(\nabla h(X_t)|V_t)=:R_t$$

→ MSE が $O(d^{-1})$ 低減可能！

Proposition 4.5

$$\frac{1}{\sqrt{T}}\int_0^{dT}h(X_t)\,\mathrm{d}t\overset{d,T\to\infty}{\Rightarrow}N\left(0,\frac{8}{\sigma^2}\right)$$

$$\frac{1}{\sqrt{T}}\int_0^{T}R_t\,\mathrm{d}t\overset{d,T\to\infty}{\Rightarrow}N\left(0,\frac{\sigma^2}{4}\right)$$



Comparison of two batch means estimators