大学共同利用機関法人 情報・システム研究機構
統計数理研究所
SOKENDAI
The Institute of Statistical Mathematics

メトロポリスを超えた枠組みで 我々はどこまで行けるか？

総合研究大学院大学 司馬博文

サンプリング問題

$U, \nabla U$ の情報のみを用い, 確率分布 $\pi(x) \propto e^{-U(x)}$ からのサンプルを構成せよ

例: $U^{(d)}(x) := \frac{1}{2}\sum_{i=1}^{d} x_i^2 \quad (x \in \mathbb{R}^d)$ 標準 Gauss

Piecewise Deterministic Monte Carlo

cf. Fearnhead+ (2018)

① 補助変数 V の分布 $\mu(v) \propto e^{-K(v)}$ を導入して問題を拡張 $\widetilde{\pi}(x,v) := \pi(x)\mu(v)$

② 力学系

$$\begin{cases} \dot{x}_t = f(x_t, v_t) \\ \dot{v}_t = g(x_t, v_t) \end{cases}$$

に従って運動する曲線

$t \mapsto (x_t, v_t)$

③ ランダムな時刻

$$\lambda(x,v) = (v|\nabla U(x))_+$$

の強度で到着するランダム時刻

$T_1, T_2, \cdots$

④ ランダムな速度ジャンプ

$$V_{T_i} \sim Q(x_{T_i-}, v_{T_i-})$$

で新しい速度をサンプリング

(⑤ リフレッシュ)

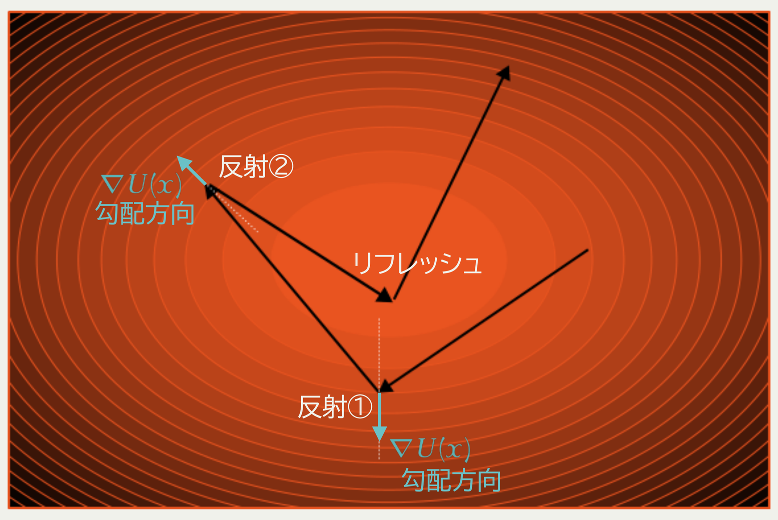
$$V_{T_j} \sim \mu(v)\,dv$$

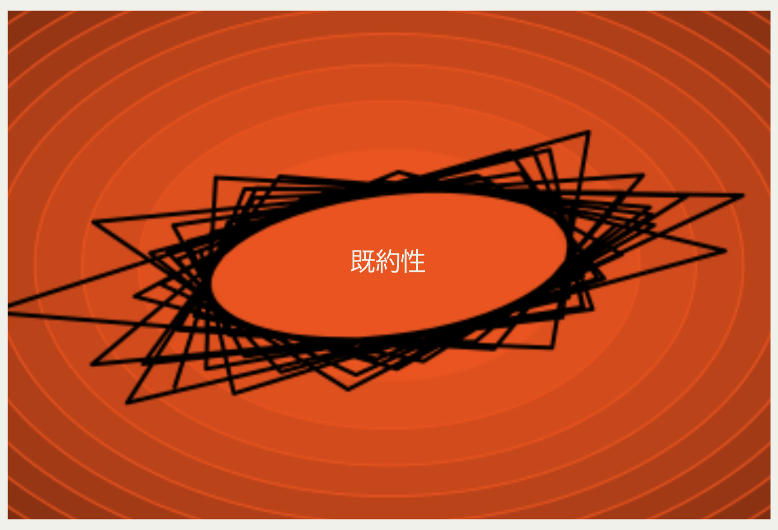
で平衡分布から取り直す

BPS

cf. Bouchard-Côté+ (2018)

Bouncy Particle Sampler





④ 等高線に関する反射

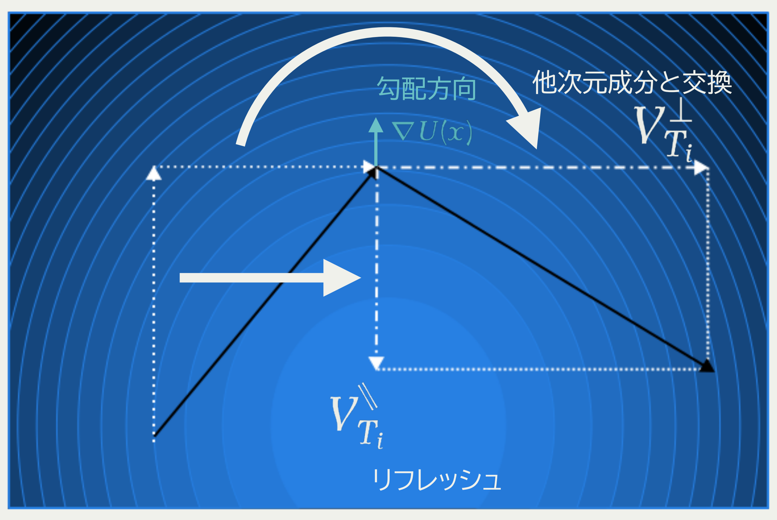
$$v_{T_i} \leftarrow v_{T_i-} - 2\frac{(\nabla U(x_{T_i})|v_{T_i-})}{\|\nabla U(x_{T_i})\|^2}\nabla U(x_{T_i})$$

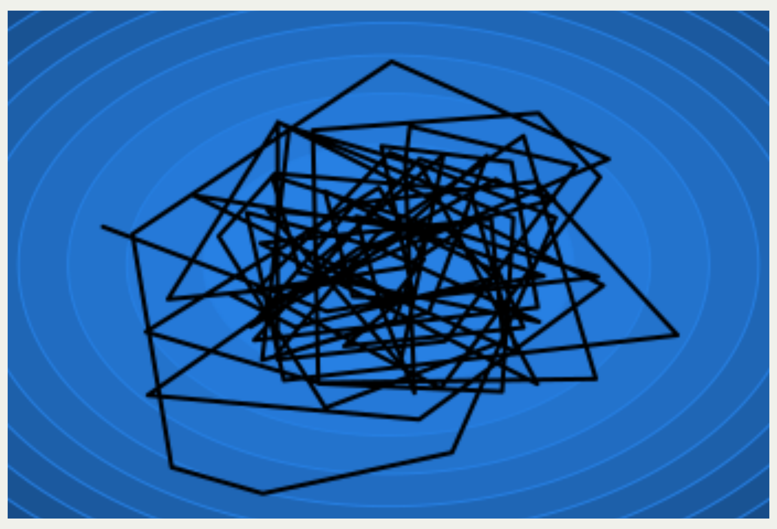
⑤ 定期的な速度リフレッシュ(必須) $\rho > 0$

FECMC

cf. Michel+ (2020)

Forward Event-Chain Monte Carlo





④ 反射とリフレッシュの組合せ: Parallel Refresh + Orthogonal Switch

別々に取り
$$\begin{cases} V^{\parallel}_{T_i} \sim Q^{\parallel}(x_{T_i-}) \\ V^{\perp}_{T_i} \leftarrow A v^{\perp}_{T_i-} \end{cases}$$
⇨ 合わせる
$$V_{T_i} \leftarrow V^{\parallel}_{T_i} + V^{\perp}_{T_i}$$

(⑤ 追加のリフレッシュ動作が不要) $\rho = 0$

Scaling Analysis: $d \to \infty$ 収束極限を比較することでアルゴリズムの性能を比較

ポテンシャルの過程 $Y_t^{(d)} := \frac{U^{(d)}(X^d_{dt}) - d}{\sqrt{d}}$ の収束極限はいずれも OU 過程 $dY_t = -\frac{\sigma^2}{4}Y_t\,dt + \sigma\,dB_t$ になる

定理 (Bierkens+2022)

$$\sigma_{\mathrm{BPS}}(\rho)^2 = 8\int_0^\infty e^{-\rho s}K(s,0)\,ds$$ で与えられる.

ただし K は次の生成作用素が定める Gauss-Markov 過程のカーネル

$$Gf(x) = f'(x) + x_+\left(f(-x) - f(x)\right)$$

定理

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

系 任意の $\rho > 0$ について

$$\sigma^2_{\mathrm{BPS}}(\rho) < \sigma^2_{\mathrm{FECMC}}$$

計算複雑性

単位長の軌跡をシミュレートするために必要な $\nabla U(x)$ の call 回数は $O(d)$

比較

Random Walk Metropolis-Hastings	$O(d)$
Metropolis-adjusted Langevin	$O(d^{1/2})$
Hamiltonian Monte Carlo	$O(d^{1/4})$

モンテカルロ分散

球面対称関数 $f : \mathbb{R}^d \to \mathbb{R}$ に関して

$$\mathrm{Var}\left[\frac{1}{N}\sum_{n=1}^{N} f(X^{\mathrm{BPS}}_{\delta n})\right] \geq \mathrm{Var}\left[\frac{1}{N}\sum_{n=1}^{N} f(X^{\mathrm{FECMC}}_{\delta n})\right]$$

が十分大きい次元 $d \gg 1$ で成立

＊ 一般の f でも成り立つ

極限での「速度」

$\widetilde{Y}_t := Y_{\alpha^2 t}$ と定めると

$$d\widetilde{Y}_t = -\frac{(\alpha\sigma)^2}{4}\widetilde{Y}_t\,dt + \sigma\alpha\,d\widetilde{B}_t$$

を満たす. この意味で

FECMC の方が極限で BPS より「速い」