Imperial College London

A Piecewise Deterministic Markov Process
via (r, θ)-swaps in hyperspherical coordinates

Alexander Terenin (Imperial College London) · Daniel Thorngren (Université de Montréal)

Abstract

Recently, a class of stochastic processes known as piecewise deterministic Markov processes has been used to define continuous-time Markov chain Monte Carlo algorithms with a number of attractive properties. Not many processes in this class that are capable of targeting arbitrary invariant distributions are currently known. We derive a process whose transition function is nonlinear through solving its Fokker-Planck equation in hyperspherical coordinates. We explore implications to both the theory of piecewise deterministic Markov processes, and to Bayesian statisticians as well as physicists seeking to use these processes for simulation-based computation.

Bouncy Particle Sampler

- Deterministic Dynamics

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{v} \qquad \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \boldsymbol{0}$$

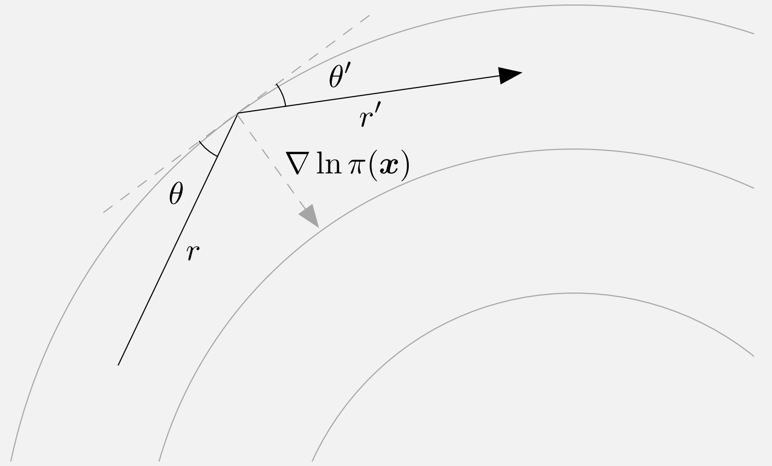
- Switching Rate

$$\lambda(\boldsymbol{x}, \boldsymbol{v}) = \max(0, -\boldsymbol{v} \cdot \nabla \ln \pi(\boldsymbol{x}))$$

- Transition Function

$$F_{\boldsymbol{x}}(\boldsymbol{v}) = \boldsymbol{v} - 2\frac{\boldsymbol{v} \cdot \nabla \ln \pi(\boldsymbol{x})}{||\nabla \ln \pi(\boldsymbol{x})||^2} \nabla \ln \pi(\boldsymbol{x})$$

Invariant measure: $\pi(\boldsymbol{x}, \boldsymbol{v}) = \pi(\boldsymbol{x})\pi(\boldsymbol{v})$ with $\pi(\boldsymbol{v})$ Gaussian



A non-BPS PDMP for Markov Chain Monte Carlo

- Deterministic Dynamics

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{v} \qquad \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \boldsymbol{0}$$

- Switching Rate

$$\lambda(\boldsymbol{x}, \boldsymbol{v}) = \max(0, -\boldsymbol{v} \cdot \nabla \ln \pi(\boldsymbol{x}))$$

- Transition Function

$$F_{\boldsymbol{x}}(\boldsymbol{v}) \quad \text{s.t.} \quad F_{\boldsymbol{x}}^{-1}(\boldsymbol{v}) \text{ exists} \quad \text{and} \quad \pi(\boldsymbol{x}, \boldsymbol{v}) \text{ is stationary}$$

Research question: does a process $(\boldsymbol{x}, \boldsymbol{v})$ with nonlinear $F_{\boldsymbol{x}}$ exist?

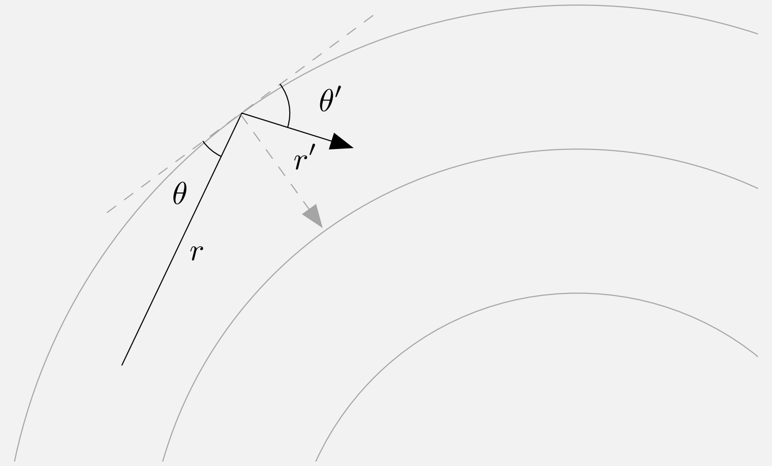
Hyperspherical process

Fokker-Planck equation

$$\lambda(\boldsymbol{x}, \boldsymbol{v})\, \pi(\boldsymbol{v}) - \lambda(\boldsymbol{x}, F_{\boldsymbol{x}}^{-1}(\boldsymbol{v}))\, \pi(F_{\boldsymbol{x}}^{-1}(\boldsymbol{v}) \mid \boldsymbol{x}) \left| \frac{\partial F_{\boldsymbol{x}}^{-1}(\boldsymbol{v})}{\partial \boldsymbol{v}} \right| = \\ = -v \cdot (\nabla \ln \pi(\boldsymbol{x}))\, \pi(\boldsymbol{v})$$

Goal: find a solution with nonlinear $F_{\boldsymbol{x}}$

1. Switch to hyperspherical coordinates: $(\boldsymbol{x}, \boldsymbol{v}) \to (\boldsymbol{x}, r, \theta, \boldsymbol{\phi})$
2. Swap magnitude for direction: $F_{\boldsymbol{x}}(r, \theta, \boldsymbol{\phi}) = (r'(\theta), \theta'(r), \boldsymbol{\phi})$



Solving the Fokker-Planck equation

Swap in hyperspherical coordinates: $F_{\boldsymbol{x}}(r, \theta, \boldsymbol{\phi}) = (r'(\theta), \theta'(r), \boldsymbol{\phi})$

$\implies$ Fokker-Planck equation reduces to ODE (k: constant)

(with boundary conditions to ensure $F_{\boldsymbol{x}}^{-1}$ exists) (p: dimension)

$$\frac{\mathrm{d}\theta'(r)}{\mathrm{d}r} = \frac{k r^p \exp\left(\frac{r^2}{-2}\right)}{\cos(\theta'(r)) \sin^{p-2}(\theta'(r))}$$

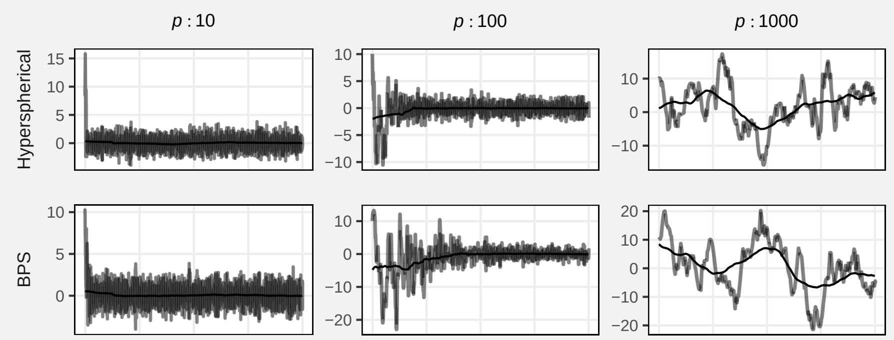
This ODE admits an analytic solution

(for numerical stability, approximate with limit as $p \to \infty$) (Φ: Gaussian CDF)

$$\theta'(r) \approx \frac{\pi}{2} - \frac{1}{2}\sqrt{\frac{-8}{p-2} \ln \Phi\left(\sqrt{2}(r - \sqrt{p})\right)}$$

Numerical simulation

Target: isotropic Gaussian



Better mixing properties in low dimension
Slows down in sufficiently high dimension

Works for Bayesian logistic regression in low dimension
Compatible with stochastic gradients and control variates
Similar behavior to bouncy particle sampler in practical settings

References

A. Terenin and D. Thorngren. A Piecewise Deterministic Markov Process via (r, θ) swaps in hyperspherical coordinates. In *Proceedings of the 38th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, 2018