Computational Complexity Comparison via Scaling Limits for Emerging MCMC Samplers

Hirofumi Shiba*†

January 20, 2026

1 A Brief History of PDMP: From Computational Physics to Bayesian Statistics

Piecewise deterministic Markov process (PDMP) serves as a mathematical umbrella for a class of 'rejection-free' continuous-time samplers that were originally proposed in computational physics, such as event-chain Monte Carlo Bernard et al. (2009) and the rejection-free scheme of Peters and de With (2012). In essence, most PDMP samplers work by simulating a piecewise linear trajectory whose invariant measure matches the target distribution. One could therefore have called them 'piecewise-linear Markov processes'. We adopt the terminology PDMP following the classical operations-research literature of Davis (1984), where this class of stochastic processes was introduced and studied from a general probabilistic viewpoint.

The event-chain algorithm, a predecessor of PDMP samplers, has become a cornerstone of computational physics. A prominent achievement is the two-dimensional melting transition simulation by Bernard and Krauth (2011). Through large-scale simulations up to the order of 10^6 particles, the authors reported the spontaneous emergence of local hexatic structures. Such empirical successes are cheifly attributed to the remarkable scalability of event-chain-type algorithms in short-range interacting systems, where they can outperform the state-of-the-art approaches such as Hamiltonian Monte Carlo (HMC).

Bouchard-Côté et al. (2018) translated these insights into the statistical vocabulary and popularized the framework by introducing the Bouncy Particle Sampler (BPS), one of the first general purpose PDMP samplers. One of their narratives is as follows. From a high-dimensional scaling viewpoint, HMC typically exhibits computational complexity of order $O(d^{5/4})$ in dimension d, whereas PDMP samplers can achieve $O(d)$ scaling in sparse Markov random fields. PDMP methods can also be favourable in large data regimes: with stochastic-gradient implementations and effective control variates, the cost can become essentially independent of the number of data points; see Bierkens et al. (2019).

2 Comparing PDMP Efficiency via High-Dimensional Scaling Limits

Despite this early momentum, the development of PDMP methods that are broadly competitive across Bayesian models appears to have slowed down. In fact, their most compelling properties seems to always come with the caveat 'under a suitable implementation', and the resulting methods can be less plug-and-play than HMC. A key bottleneck is the lack of a unified theoretical framework for comparing PDMP algorithms. For instance, one of the original contributer for event-chain algoriths developed Forward Event-Chain Monte Carlo (FECMC) Michel et al. (2020), leveraging physical intuition to design an algorithm that is often substantially more efficient than BPS in practice. However, the original work provided limited theoretical explanation for the observed gains, and the method remained less widely adopted than its empirical performance might suggest.

We propose to quantify the claim that FECMC is faster than BPS as follows. For simplicity, let the target π^d be the d-dimensional standard Gaussian distribution, and denote by $(X_t^d)_{t\geqslant 0}$ the continuous-time Markov process produced by either FECMC or BPS targeting π^d. A particularly hard-to-explore direction is given by the negative log-density, or potential

$$h^d(x) := -\log \pi^d(x),$$

which, for the standard Gaussian, reduces to $h^d(x) = \|x\|_2^2/2$. As d grows, the scalar projection $\tilde{Y}_t^d := h^d(X_t^d)$ becomes increasingly slow as the exploration must effectively coordinate across d independent directions. We therefore consider the accelerated (and normalised) process

$$Y_t^d := \frac{h^d(X_t^d) - \mathrm{E}_{\pi^d}[h^d(X)]}{\sqrt{\mathrm{Var}_{\pi^d}[h^d(X)]}}, \qquad t \geqslant 0,$$

*Institute of Statistical Mathematics (ISM), Tokyo, Japan
†Email: shiba.hirofumi@ism.ac.jp

and study its scaling limit as $d \to \infty$ under stationarity $X_0^d \sim \pi^d$. For both FECMC and BPS, the limiting dynamics take the same functional form:

$$dY_t = -\frac{\sigma^2}{2} Y_t \, dt + \sigma \, dB_t, \tag{1}$$

where B_t is standard Brownian motion. Crucially, the diffusion coefficient σ depends on the algorithm. Writing σ_F and σ_B for FECMC and BPS respectively, we obtain $\sigma_F > \sigma_B$. Since (1) is an Ornstein–Uhlenbeck process, the difference amounts to a time change $t \mapsto (\sigma_F^2/\sigma_B^2)t$, so that $\sigma_F > \sigma_B$ rigorously means that FECMC explores the geometry of $h^d(x) = -\log \pi^d(x)$ faster than BPS.

This comparison has an important implication for Monte Carlo estimation: For estimating $\mathrm{E}_{\pi^d}[h^d(X)]$, the asymptotic variance under FECMC is smaller by a factor of approximately $\sigma_F^2/\sigma_B^2 \approx 1.77$ compared with BPS. When CPU time is also taken into account, the performance gap is even more pronounced: in moderate-to-high dimensions (for $d \geqslant 10$), we observe that the effective sample size (ESS) for $\mathrm{E}[h^d(X)]$ can differ by around a $15\times$.

3 Fast Proxy Trick: Early Detection of Mixing Slowdown for PDMP Samplers

A closer inspection of the proof reveals that the key object behind $\widetilde{Y}_t^d$ is the time derivative thereof, denoted by

$$R_t^d := (X_t^d, V_t^d),$$

where V_t^d is the velocity (momentum) variable. In common practice, the velocity variable V_t^d is treated solely as an auxiliary component introduced purely for computational convenience, hence discarded after sampling. However, we claim that V_t^d carries nontrivial information, especially about the transient behaviour of the sampler when a separation of time scales is present.

Exploiting this slow-fast structure, we obtain as a corollary that one is able to estimate the asymptotic variance of Monte Carlo averages for h^d with an accuracy that does not degrade with the dimension d, by examining R^d instead of $\widetilde{Y}^d$. More broadly, our results suggest the following principle: even when the sampler slows down by a factor $O(d)$ along a difficult direction such as $h^d(X_t^d)$, this slowdown can be detected on an $O(1)$ time scale by monitoring its time derivative. Such coupling between slow and fast processes should abound in other momentum-based methods such as HMC, where auxiliary variables may mix faster than the position components. In PDMP algorithms, such time derivatives are typically readily available without heavy computational overheads because they appear internally in the event-rate computation. This 'fast proxy' trick points to a promising route toward efficient, principled convergence diagnostics for PDMP samplers.

References

E. P. Bernard and W. Krauth. Two-Step Melting in Two Dimensions: First-Order Liquid-Hexatic Transition. *Phys. Rev. Lett.*, 107:155704, Oct 2011.

E. P. Bernard, W. Krauth, and D. B. Wilson. Event-chain Monte Carlo algorithms for hard-sphere systems. *Phys. Rev. E*, 80:056704, Nov 2009.

J. Bierkens, P. Fearnhead, and G. Roberts. The Zig-Zag Process and Super-Efficient Sampling for Bayesian Analysis of Big Data. *The Annals of Statistics*, 47(3):1288–1320, 2019.

A. Bouchard-Côté, S. J. Vollmer, and A. Doucet. The Bouncy Particle Sampler: A Nonreversible Rejection-Free Markov Chain Monte Carlo Method. *Journal of the American Statistical Association*, 113(522):855–867, 2018.

M. H. A. Davis. Piecewise-Deterministic Markov Processes: A General Class of Non-Diffusion Stochastic Models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(3):353–376, 1984.

M. Michel, A. Durmus, and S. Sénécal. Forward Event-Chain Monte Carlo: Fast Sampling by Randomness Control in Irreversible Markov Chains. *Journal of Computational and Graphical Statistics*, 29(4):689–702, 2020.

E. A. J. F. Peters and G. de With. Rejection-free Monte Carlo Sampling for General Potentials. *Physical Review E*, 85(2), 2012.