Adaptive importance sampling of stochastic processes with graph-based mean hitting times

Guillaume Chennetier[†,1,2], Hassane Chraibi[1], Anne Dutfoy[1], Josselin Garnier[§,2]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Apr. 2021 – May. 2024

[1] EDF Lab Paris-Saclay, Boulevard Gaspard Monge, 91120 Palaiseau, France
{guillaume.chennetier,hassane.chraibi,anne.dutfoy}@edf.fr

[2] CMAP, Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau Cedex, France
{guillaume.chennetier,josselin.garnier}@polytechnique.edu

Abstract

Importance sampling is one of the most famous and perhaps the most direct method of variance reduction in simulation-based inference when estimating quantities of the form $\mathbb{E}_{X\sim\mathbf{p}}\left[\varphi(X)\right]$.

$$\overline{\varphi} := \mathbb{E}_{X\sim\mathbf{p}}\left[\varphi(X)\right] = \mathbb{E}_{X\sim\mathbf{g}}\left[\varphi(X)\frac{\mathbf{p}(X)}{\mathbf{g}(X)}\right] \approx \frac{1}{n}\sum_{k=1}^{n}\varphi(X_k)\frac{\mathbf{p}(X_k)}{\mathbf{g}(X_k)} \quad \text{with } X_1,\ldots,X_n \sim \mathbf{g}. \tag{1}$$

Generating a sample $X_1,\ldots,X_n$ under an importance distribution $\mathbf{g}$ as close as possible to $\mathbf{g}_{\mathsf{opt}} \propto |\varphi| \times \mathbf{p}$ results in a low-variance estimator of $\overline{\varphi}$. A central use case is the rare event setting $\varphi(X) = \mathbb{1}_{X\in\mathbf{F}}$ (with $\overline{\varphi}$ very small) for which Monte Carlo methods are too expansive.

Simulating X according to an efficient importance distribution is notoriously challenging in high dimension. Various approaches have been proposed to address this problem (notably within the MASCOT-NUM community) such as projection into well-chosen subspaces [1, 2] or the use of generative models [3]. The dimension of the space is also an issue when X is not a vector but the trajectory of a stochastic process. This is a common case in reliability assessment where the operation of an industrial system is modeled by a stochastic process (see our previous work [4]).

We propose a new family of importance distributions tailored for dynamic rare event simulation with any non-diffusive Markov process, i.e any piecewise deterministic Markov process (PDMP). These processes evolve according to deterministic differential equations whose parameters are subject to random jumps. Their simulation cost prohibits the learning of a good importance distribution in a non-parametric framework, and the hybrid geometry of their state space prevents the application of most known dimension reduction methods.

The optimal importance distribution for PDMPs was characterized in [5] from the so-called "committor function" of the process. We present a methodology that combines the approximation of this committor function and a cross-entropy procedure to sequentially estimate $\overline{\varphi}$:

1. The PDMP is approximated by a time-homogeneous Markov random walk on a graph.

2. Mean hitting times of a target region of the graph are explicit for the simplified process, and form the basis of a family of approximations of the committor function.

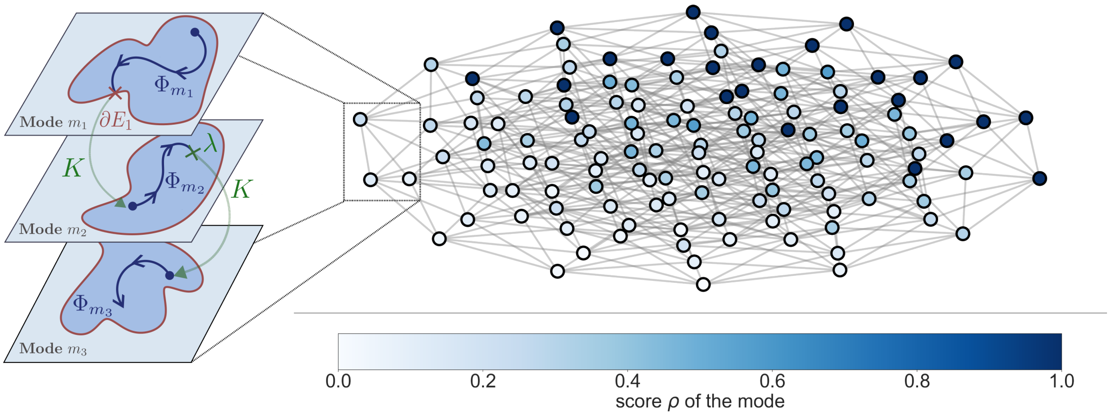

Figure 1: PDMP as a random walk on a graph.

3. This leads to a parametric family of importance distributions, for which we seek a good representative using a cross-entropy method with a new multilevel mechanism.

Our method achieves a variance reduction factor of approximately 10,000 compared to standard Monte Carlo estimation on a large-scale industrial test case.

Short biography (PhD student)

After a master's degree in statistics at Sorbonne University, Guillaume Chennetier did his internship then PhD thesis between École polytechnique and EDF R&D Saclay in rare event simulation for piecewise deterministic Markov processes in order to access the reliability of industrial systems involved in the operation of nuclear and hydraulic power plants.

References

[1] Maxime El Masri. *High dimensional importance sampling through projections on a low dimensional subspace*. PhD thesis, ISAE-Institut Supérieur de l'Aéronautique et de l'Espace, 2022.

[2] Felipe Uribe, Iason Papaioannou, Youssef M Marzouk, and Daniel Straub. Cross-entropy-based importance sampling with failure-informed dimension reduction for rare event simulation. *SIAM/ASA Journal on Uncertainty Quantification*, 9(2):818–847, 2021.

[3] Julien Demange-Chryst, François Bachoc, Jérôme Morio, and Timothé Krauth. Variational autoencoder with weighted samples for high-dimensional non-parametric adaptive importance sampling. *arXiv preprint arXiv:2310.09194*, 2023.

[4] Guillaume Chennetier, Hassane Chraibi, Anne Dutfoy, and Josselin Garnier. Adaptive importance sampling based on fault tree analysis for piecewise deterministic markov process. *arXiv preprint arXiv:2210.16185*, 2022.

[5] Hassane Chraibi, Anne Dutfoy, Thomas Galtier, and Josselin Garnier. On the optimal importance process for piecewise deterministic markov process. *ESAIM: Probability and Statistics*, 23:893–921, 2019.