Importance Sampling in high dimension

J. Beh[†,1,2], F. Simatos[§,2], J. Morio[§,1,2]

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

PhD expected duration: Jan. 2023 – Dec. 2025

[1] ONERA/DTIS, Université de Toulouse, F-31055 Toulouse
{jason.beh,jerome.morio}@onera.fr

[2] Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, 31000 Toulouse
florian.simatos@isae-supaero.fr

Abstract

Rare event analysis often involves the estimation of the rare event probability $p = \mathbb{P}_f(X \in A)$, where $f = N(0, I)$ is the d-dimensional standard Gaussian distribution, which is a fairly general setting owing to isoprobabilistic transformations [8]. Whereas Adaptive Splitting [3] concerns the modification of the trajectories of the samples towards the region of interest A, Importance Sampling (IS) considers an auxiliary distribution g which allocates more probability mass in A than f. Given n_g samples $(Y_i)_{i=1\dots n}$ generated according to g to whom $\mathbb{1}(\cdot \in A) f$ is absolutely continuous, the IS estimator is written as

$$\hat{p}_g = \frac{1}{n_g} \sum_{i=1}^{n_g} \frac{f(Y_i)}{g(Y_i)} \mathbb{1}(Y_i \in A)$$

In low dimension, IS estimators are often employed due to the desired variance reduction property compared to Monte Carlo estimator. However, in high dimension, IS estimators suffer from convergence issues and become extremely sensitive to the choice of auxiliary distribution. This motivates a theoretical study on the convergence of IS estimators in the high-dimensional setting, $d \to +\infty$.

As $d \to +\infty$, two settings can arise: either the probability to be estimated is bounded away from zero: $\inf_d p > 0$, or the probability tends to zero with the dimension: $p \to 0$. The first setting $\inf_d p > 0$, considered by [1, 4], occurs when p involves a stochastic process which is approximated by a finite sum of random variables by principal component analysis. Then, the probability to estimate becomes p_d, which tends to $p > 0$ when $d \to \infty$. In this setting, We will discuss our work on the convergence of the Cross-Entropy scheme [2] as well as its projection and improved variants [5, 10, 9].

The second setting, $p \to 0$ as $d \to \infty$, considered by [7, 6], occurs in specific settings such as in Highly Reliable Markovian Systems or in static network reliability estimation. This setting is more complex since the properties of IS estimators are reliant on the rate of convergence of p to 0. To tackle this setting, we first establish necessary and sufficient conditions for general IS estimators to be consistent, and conditions to verify a Central Limit Theorem towards a normal distribution. We then translate these conditions into the necessary rate of growth of the sample size n_g for various auxiliary distributions in a classical large deviation setting, $A = \{x \in \mathbb{R}^d : \sum_{j=1}^{d} x(j) \geq d^\gamma\}$ with $\gamma > 1/2$. It will be observed that the 'optimal' Gaussian density for IS largely depends on the error metric considered.

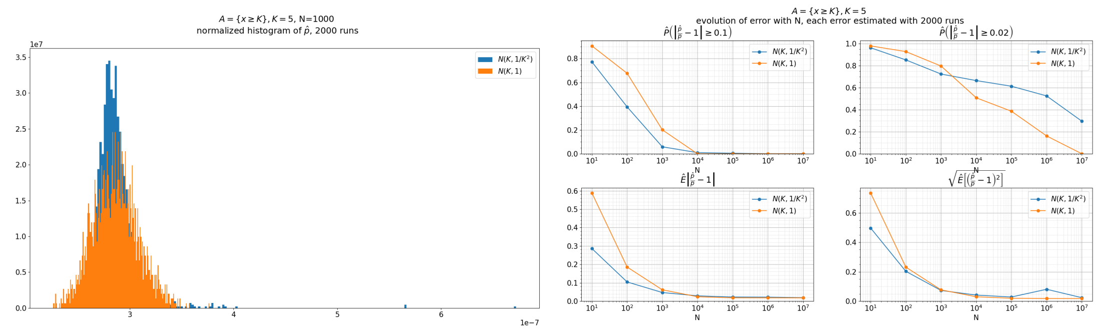

Figure 1: The histogram ($n_g = 1000$) and the evolution of usual error metrics with n_g of $\hat{p}_g$ for two choices of auxiliary density: which is better?

Short biography (PhD student)

I hold an M. Sc. in engineering from ISAE-SUPAERO in Toulouse. I completed my final-year internship at ISAE-SUPAERO which led to my current PhD thesis co-funded by EUR-MINT and ONERA, under the supervision of F. Simatos and J. Morio. The main goal of the thesis is to study the cause of the well-known curse-of-dimensionality for adaptive importance sampling, and to provide means to circumvent it based on this knowledge.

References

[1] S. Au and J. Beck. Important sampling in high dimensions. *Structural Safety* 25.2 (2003), pp. 139–163.

[2] J. Beh, Y. Shadmi, and F. Simatos. Insight from the Kullback–Leibler divergence into adaptive importance sampling schemes for rare event analysis in high dimension. *arXiv preprint* (2023). arXiv: `2309.16828 [math.ST]`.

[3] F. Cérou, A. Guyader, and M. Rousset. Adaptive multilevel splitting: Historical perspective and recent results. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29.4 (2019), p. 043108.

[4] S. Chatterjee and P. Diaconis. The sample size required in importance sampling. *The Annals of Applied Probability* 28.2 (2018).

[5] M. El Masri, J. Morio, and F. Simatos. Improvement of the cross-entropy method in high dimension for failure probability estimation through a one-dimensional projection without gradient estimation. *Reliability Engineering & System Safety* 216 (2021), p. 107991.

[6] A. Guyader and H. Touchette. Efficient Large Deviation Estimation Based on Importance Sampling. *Journal of Statistical Physics* 181.2 (2020), pp. 551–586.

[7] P. L'Ecuyer et al. Asymptotic robustness of estimators in rare-event simulation. *ACM Trans. Model. Comput. Simul.* 20.1 (2010).

[8] R. Lebrun and A. Dutfoy. Do Rosenblatt and Nataf isoprobabilistic transformations really differ? *Probabilistic Engineering Mechanics* 24.4 (2009), pp. 577–584.

[9] I. Papaioannou, S. Geyer, and D. Straub. Improved cross entropy-based importance sampling with a flexible mixture model. *Reliability Engineering & System Safety* 191 (2019), p. 106564.

[10] F. Uribe et al. Cross-Entropy-Based Importance Sampling with Failure-Informed Dimension Reduction for Rare Event Simulation. *SIAM/ASA Journal on Uncertainty Quantification* 9.2 (2021), pp. 818–847.