Automated adaptive multilevel splitting

J. Menesch, T. Taimre and R. Vaisman

School of Mathematics and Physics, The University of Queensland, Brisbane, Australia
Email: j.menesch@uq.edu.au

Abstract: Consider the hitting problem – the estimation of the probability that a stochastic differential equation (SDE), given by $\mathrm{d}\boldsymbol{X}_t^\varepsilon = a(\boldsymbol{X}_t^\varepsilon)\mathrm{d}t + B(\boldsymbol{X}_t^\varepsilon, \varepsilon)\mathrm{d}\boldsymbol{W}_t$ with $\boldsymbol{X}_0^\varepsilon = \boldsymbol{x}_0$, hits a set R before another set O:

$$p^\varepsilon := \mathbb{P}(\tau_R < \tau_O \mid \boldsymbol{X}_0^\varepsilon = \boldsymbol{x}_0) \text{ where } \tau_A := \inf\{t > 0 \mid \boldsymbol{X}_t^\varepsilon \in A\}. \quad (1)$$

When this is rare, say $p^\varepsilon < 10^{-6}$, estimation requires rare-event methods, and these methods typically involve the use of an importance function (IF) or sampling distribution. Splitting methods are well-suited for hitting problems, and specifically, adaptive multilevel splitting (AMS) is a good choice due to its practicality, e.g., unbiasedness and automated choice of levels. However, AMS requires an IF, whose choice determines the algorithm's efficiency (Cérou et al. 2019). The optimal IF is the committor function, which is generally not known nor easily guessed. Although determining p^ε in (1) can oftentimes be reformulated as the solution to an equivalent PDE, numerical solution to the PDE is intractable beyond a few dimensions. Other methods, such as those in transition path theory, also rely on the committor function and aim to approximate it directly.

We propose an iterative heuristic, called *AutoIF*, for solving the hitting problem, in which a sequence of modified, increasingly rare problems are solved via AMS, with the ensemble of points generated by AMS at each iteration being used to approximate a new IF for the next iteration. Specifically, the rarity is encoded by the SDEs noise parameter ε, with the first ε chosen such that the problem is not rare, and the last ε corresponding the target problem. In the first iteration, AMS is run with a user-defined IF being used for the first iteration. The weighted and labelled ensemble of points generated during AMS are stored and subsequently used with a function approximation algorithm to generate a new IF for the next iteration. In our tests, we used a cell-based method (following Lopes and Lelièvre (2019)) and logistic regression for this task. The process is repeated for each ε in the sequence, and after the last ε, the final approximated IF is used with AMS to find the hitting probability for the target problem. The heuristic is motivated by the observation that AMS efficiently generates an ensemble of points which can be utilized in solving the problem. This allows gradual approximation of IFs for rarer and rarer problems, thus eventually leading to an IF for the target problem which may provide a better approximation of the committor function than anything that can be approximated outright.

We tested the AutoIF method on 3 SDEs – an Ornstein-Uhlenbeck spiral, three-wells potential, and rugged Mueller potential – across 3 levels of rarity ($p^\varepsilon \approx 10^{-3}, 10^{-6}, 10^{-9}$) and 3 dimensionalities ($d = 2, 5, 10$). Results were compared against a typical user-defined IF – the Euclidean distance to the center of R. The primary performance metric was relative time-variance product (RTVP). We found that AutoIF generally outperforms AMS for smaller p^ε due to decreased relative error, but the function approximation method started to hinder the method in higher dimensions. In addition, AutoIF performed best for the spiral and three-wells problems but struggled with the rugged Mueller problems. Finally, the cell method almost always outperformed logistic regression in terms of RTVP, albeit with greater likelihood of encountering numerical issues.

Overall, we found that AutoIF is strongest in high-rarity problems that have complex committor functions. The proposed heuristic is a promising choice for rare-event problems which are out of reach for traditional algorithms. Further improvements could include a better choice of function approximation algorithm and a detailed analysis of the IF updating procedure.

REFERENCES

Cérou, F., Delyon, B., Guyader, A., Rousset, M., 2019. On the asymptotic normality of adaptive multilevel splitting. SIAM-ASA Journal on Uncertainty Quantification 7(1), 1–30.

Lopes, L.J.S., Lelièvre, T., 2019. Analysis of the adaptive multilevel splitting method on the isomerization of alanine dipeptide. Journal of Computational Chemistry, 40(11), 1198–1208.

Keywords: *Rare-event simulation, adaptive multilevel splitting, importance function, committor function*