Interplay between
Monte Carlo methods and automatic differentiation

Edouard Pauwels (Toulouse School of Economics)
Samuel Vaiter (CNRS et Université Côte d'Azur, Laboratoire J.A. Dieudonné, Nice)

Internship proposal (Spring 2025, 5 to 6 months) with PhD fundings (3 years)

Context. Algorithmic differentiation [8], known as "backpropagation" in AI and learning architectures [6], is a critical tool for modern parametric optimization. It enables the joint propagation of iterates and their derivatives concerning external parameters in iterative algorithms. The asymptotic properties of these derivative processes were explored in early works [5, 1], especially in relation to implicit differentiation [7].

With the rise of machine learning, there has been renewed interest in differentiating gradient-based algorithms, particularly using implicit differentiation [12] and forward propagation asymptotics [10]. This area has expanded into applications for nonsmooth optimization [2] and quantitative convergence estimates [11]. In particular, recent studies of gradient-type algorithms in empirical risk minimization have tackled the asymptotic behavior of derivatives under stochastic conditions [9].

Building on this foundation, we now turn to the theoretical challenges involved in differentiating Monte Carlo-based techniques, with a focus on creating robust mathematical foundations for these methods.

Goals of the internship. The primary goal of the internship is to develop a rigorous framework for differentiating Monte Carlo-based techniques, with potential applications in probabilistic programming and differentiable rendering. Although existing results provide conditions for interchanging (generalized) derivatives and integrals under specific smoothness constraints [3], directly differentiating Monte Carlo sampling remains complex.

To address this, we propose a two-phase approach:

1. Establishing a nonsmooth Reynold's theorem: This phase will involve formulating an adapted version of Reynold's theorem to support nonsmooth settings, enabling differentiation and integration interchange in cases common to Monte Carlo methods with nonsmooth distributions.

2. Developing differentiable Monte Carlo strategies through stratification: By leveraging structured stratification in probabilistic models, we aim to construct sampling strategies compatible with differentiation. This step will explore stratification as a means to maintain differentiability, leading to methods that can efficiently handle nonsmooth elements within Monte Carlo approaches.

This research will provide deeper theoretical insights into differentiable probabilistic algorithms, with applications spanning differentiable rendering, probabilistic programming, and other areas where differentiable sampling is essential.

Candidate profile. We seek a second year Master student in applied mathematics or computer science. A prior background in optimisation is welcome, along with a taste for numerical simulations.
Location. The supervision of the internship is shared between Nice and Toulouse. The internship may either take principally take place at Toulouse School of Economics or at Université Côte d'Azur depending on the particular interests of the intern.
Application procedure. Send your CV and a copy of your last year transcripts to both Edouard Pauwels (`edouard.pauwels@tse-fr.eu`) and Samuel Vaiter (`samuel.vaiter@cnrs.fr`).

References

[1] Beck, T. (1994). Automatic differentiation of iterative processes. Journal of Computational and Applied Mathematics.

[2] Bolte, J., Pauwels, E., Vaiter, S. (2022). Automatic differentiation of nonsmooth iterative algorithms. NeurIPS.

[3] Bolte, J., Le, T., Pauwels, E. (2023). Subgradient sampling for nonsmooth nonconvex minimization. SIAM J. Optim.

[4] Bottou, L., Curtis, F. E., Nocedal, J. (2018). Optimization methods for large-scale machine learning. Siam Review.

[5] Gilbert, J. C. (1992). Automatic differentiation and iterative processes. Optimization Methods and Software.

[6] Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep learning.

[7] Griewank, A., Faure, C. Piggyback differentiation and optimization (2003). In Large- scale PDE-constrained optimization. Springer.

[8] Griewank, A., Walther, A. (2008). Evaluating derivatives: principles and techniques of algorithmic differentiation. SIAM.

[9] Iutzeler, F., Pauwels, E., Vaiter, S. (2024) Derivatives of Stochastic Gradient Descent. NeurIPS.

[10] Lorraine, J., Vicol, P., Duvenaud, D. (2020). Optimizing millions of hyperparameters by implicit differentiation. AISTATS.

[11] Mehmood, S., Ochs, P. (2020). Automatic differentiation of some first-order methods in parametric optimization. AISTATS.

[12] Pedregosa, F. (2016). Hyperparameter optimization with approximate gradient. ICML.