

Master 2 internship proposal

Physique et Mécanique des Milieux Hétérogènes

Contact: Y.-E. Keta (yann-edwin.keta@espci.fr), P. Mangold (paul.mangold@polytechnique.edu)

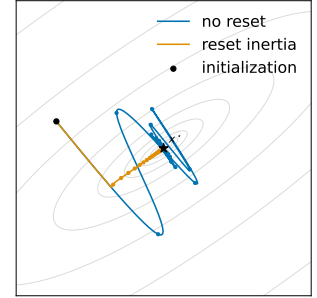
Internship location: PMMH (Jussieu, Paris 5e) and CMAP (Polytechnique, Saclay)

PhD opportunity: Yes

Stochastic resetting of velocities in Heavy Ball minimisation methods

This internship aims to bridge the gap between methods from statistical physics and optimisation for machine learning. Consider a function $F: \mathbb{R}^d \rightarrow \mathbb{R}^d$, and the inertial gradient flow, defined by the following set of continuous or discrete equations.

$$\begin{cases} \tau \dot{\mathbf{v}}(t) &= -\mathbf{v}(t) - \frac{\mathbf{F}'(\mathbf{x}(t))}{\zeta} \\ \dot{\mathbf{x}}(t) &= \mathbf{v}(t) \end{cases} \Rightarrow \begin{cases} \mathbf{v}_{t+1} &= (1 - \frac{\gamma}{\tau})\mathbf{v}_t - \frac{\gamma}{\tau} \frac{\mathbf{F}'(\mathbf{x}_t)}{\zeta} \\ \mathbf{x}_{t+1} &= \mathbf{x}_t + \gamma \mathbf{v}_t \end{cases} \quad (1)$$



This scheme, or variants of it, is often called the Heavy Ball method due to its similarity with inertial equations of motion. This method is known to converge on convex functions, with accelerated rates on quadratic functions F . However, the inertial component of this method makes iterates overshoot the minimum (see Figure).

On quadratic functions, with deterministic gradients, this overshooting phenomenon can be corrected by periodically resetting the inertial component $\mathbf{v}_t$ to zero [1], while preserving the accelerated convergence. Later, Ref. [2] showed empirically that resetting can also benefit when gradients are stochastic, although without formal proof of their claims.

In the context of statistical physics, resetting schemes have been shown to accelerate search algorithms [3] and create nonequilibrium currents [4] due to their effectiveness at biasing dynamics against unrewarding directions of motion. In the context of the optimisation of functions with multiple minima, this method may thus help bias the dynamics towards deeper regions of F and uncover deeper minima.

This internship aims at computing analytically the rate of convergence of the Heavy Ball method for quadratic and convex functions in the presence of stochastic resetting and possibly stochastic gradients. This analysis will show whether an optimal resetting rate may be chosen with minimal to no prior knowledge of the minimized function F . Lastly the dynamics of the Heavy Ball method with stochastic resetting will be characterised for functions F with rugged landscapes inspired by the potential energy of dense and disordered systems [5].

- [1] B. O'Donoghue and E. Candès, “Adaptive Restart for Accelerated Gradient Schemes”, *Foundations of Computational Mathematics* **15**, 715 (2015).
- [2] B. Wang, T. Nguyen, T. Sun, A. L. Bertozzi, R. G. Baraniuk, and S. J. Osher, “Scheduled Restart Momentum for Accelerated Stochastic Gradient Descent”, *SIAM Journal on Imaging Sciences* **15**, 738 (2022).
- [3] M. R. Evans and S. N. Majumdar, “Diffusion with Stochastic Resetting”, *Physical Review Letters* **106**, 160601 (2011).
- [4] P. K. Ghosh, S. Nayak, J. Liu, Y. Li, and F. Marchesoni, “Autonomous ratcheting by stochastic resetting”, *The Journal of Chemical Physics* **159**, 031101 (2023).
- [5] P. G. Debenedetti and F. H. Stillinger, “Supercooled liquids and the glass transition”, *Nature* **410**, 259 (2001).

Expected skills: We expect prospective students to be comfortable with linear algebra and be proficient with at least one multipurpose programming language (*e.g.* Python). Candidates with knowledge of stochastic calculus, statistical physics, and/or optimisation are most welcome.