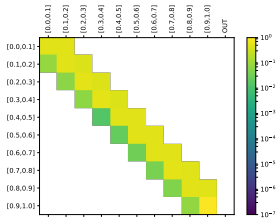
Inria Sophia Antipolis
Group: Algorithms-Biology-Structure
Frederic.Cazals@inria.fr
Web: `http://team.inria.fr/abs`



Matrix showing the diffusivity of a random walk across energy strata in a high dimensional space. From [1].

Master internship proposal

Approximation algorithms for densities of states and polytope volume calculations

Keywords. High-dimensional spaces, sampling, Markov chains, multi-phase Monte Carlo; statistical physics, density of states, volume calculations.

Context. A classical problem in computational mathematics is to compute the volume of a polytope, given as the intersection of n half-spaces in $\mathbb{R}^d$. A classical problem in computational physics is the calculation of the density of states (DoS) of a system (the *number* of states with a given energy), from which the partition function can be derived. Both problems are known to be hard [2] and hard to approximate: by concentration arguments, the volume or mass (in a probabilistic sense) is generally concentrated near a *typical set* of small dimension. For this reason, probabilistic algorithms are resorted to. For polytopes, multi-phase Monte Carlo methods provide an (ε, δ) approximation in time $O^\star(d^3)$ [2, 3]. These algorithms are typically based on random walks such as hit-and-run, or ball walk, whose mixing times have been analyzed. For DoS, multi-phase adaptive Monte Carlo methods of the Wang-Landau type [4] have been designed, and their asymptotic convergence proved [5]. However, no error bounds are known.

Goals. Consider a physical system with a potential energy. In short, the The Wang-Landau algorithm [4] is an adaptive Monte Carlo method computing the DoS in the form of a histogram corresponding to energy strata. We recently contributed an enhanced Wang-Landau algorithm [1], based upon a novel the proposal (the random walk used in the Metropolis-Hastings kernel) exploiting properties of high dimensional spaces to target the aforementioned typical set. This proposal *learns* global geometric properties of the function defining the potential energy, so as to remain *diffusive* near the aforementioned typical set. Our method has been implemented and provides (for the first time) converged DoS calculations on small bio-molecules, in particular for non harmonic models of single energy basins.

In this context, the goal of this internship will be twofold. On the design side, the goal will be to generalize our method to cope with energy function possibly having a large number of local minima. On the analysis side, the goal will be understand the convergence properties of the method.

Background. Theoretical computer science, approximation algorithms, applied mathematics, geometry.

Misc. The duration is of six months. Possibility to follow-up with a PhD thesis in the scope of the Institut d'Intelligence Artificielle (3IA) Côte d'Azur.

References

[1] A. Chevallier and F. Cazals. Wang-landau algorithm: an adapted random walk to boost convergence. *J. of Computational Physics (Under revision)*, 2019.

[2] S. Levy. *Flavors of Geometry*. Cambridge University Press, 1997.

[3] B. Cousins and S. Vempala. A practical volume algorithm. *Mathematical Programming Computation*, 8(2):133–160, 2016.

[4] F. Wang and D.P. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Physical review letters*, 86(10):2050, 2001.

[5] G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre, and G. Stoltz. Convergence of the Wang-Landau algorithm. *Mathematics of Computation*, 84(295):2297–2327, 2015.