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

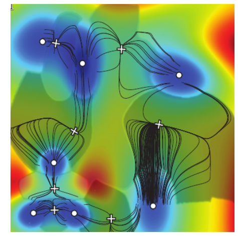

Figure 1: Saddle searches: 2D illustrations. Saddles marked with X. From [1].

MASTER INTERNSHIP PROPOSAL

EFFICIENTLY LOCATING SIGNIFICANT SADDLE POINTS ON HIGH DIMENSIONAL ENERGY SURFACES

Keywords. High-dimensional spaces, saddle points, multi-scale analysis, numerical algorithms.

Context. Energy landscapes (EL), or fitness landscapes, are central concepts in (bio-)physics and optimization. In short, such a landscape is defined by a real valued multivariate function. When the function is not convex–the vast majority of cases, it is crucial to identify local minima (or *comparable* height) as well as saddle points connecting them.

On the one hand, the enumeration of local minima is a central topic for which efficient randomized algorithms have been developed, including hybrid methods [2] combining the classical methods known as basin hoping and `T-RRT`. On the other hand, the enumeration of saddle points lags behind. Classical methods have been used in physics [3], with tests conducted on benchmarks [4]. Recently, these methods have been improved using so-called mode following [1, 5].

Goals. A key difficulty in dealing with high dimensional energy landscapes is that the number of critical points (local minima, saddle points of all indices) is generally exponential in the dimension. Fortunately, such landscapes exhibit a remarkable hierarchical structure and can be simplified using *topological persistence* [6, 7]. Therefore, locating local minima and the associated saddle points before simplifying those which are spurious *i.e.* non persistent is clearly sub-optimal.

The goal of this internship will be to develop a novel strategy, using insights on the geometry of energy landscapes, to avoid these spurious steps. The methods will be implemented in (advanced/generic) C++ in the Structural Bioinformatics Library, see `http://sbl.inria.fr`.

Conditions. Internship with *gratification*. Possibility to follow-up with a PhD thesis.

References

[1] Andreas Pedersen, Sigurdur F Hafstein, and Hannes Jónsson. Efficient sampling of saddle points with the minimum-mode following method. *SIAM Journal on Scientific Computing*, 33(2):633–652, 2011.

[2] A. Roth, T. Dreyfus, C.H. Robert, and F. Cazals. Hybridizing rapidly growing random trees and basin hopping yields an improved exploration of energy landscapes. *J. Comp. Chem.*, 37(8):739–752, 2016.

[3] D. Sheppard, R. Terrell, and G. Henkelman. Optimization methods for finding minimum energy paths. *The Journal of chemical physics*, 128(13):134106, 2008.

[4] Samuel T Chill, Jacob Stevenson, Victor Ruehle, Cheng Shang, Penghao Xiao, James D Farrell, David J Wales, and Graeme Henkelman. Benchmarks for characterization of minima, transition states, and pathways in atomic, molecular, and condensed matter systems. *Journal of chemical theory and computation*, 10(12):5476–5482, 2014.

[5] M. Plasencia Gutieérrez, C. Argaéez, and H. Jónsson. Improved minimum mode following method for finding first order saddle points. *Journal of chemical theory and computation*, 13(1):125–134, 2017.

[6] F. Cazals, T. Dreyfus, D. Mazauric, A. Roth, and C.H. Robert. Conformational ensembles and sampled energy landscapes: Analysis and comparison. *J. Comp. Chem.*, 36(16):1213–1231, 2015.

[7] J. Carr, D. Mazauric, F. Cazals, and D. J. Wales. Energy landscapes and persistent minima. *The Journal of Chemical Physics*, 144(5):4, 2016.