Report on the 2008 HIM Junior Hausdorff Semester Program: Numerical methods in molecular simulation.

Frédéric Legoll[(2,3)], Tony Lelièvre[(1,3)], Mathias Rousset[(4)], Gabriel Stoltz[(1,3)].
Institutional affiliations:

- (1) CERMICS, Ecole Nationale des Ponts et Chaussées, 6 et 8 av. Blaise Pascal, 77455 Marne la Vallée, FRANCE.
- (2) LAMI, Ecole Nationale des Ponts et Chaussées, 6 et 8 av. Blaise Pascal, 77455 Marne la Vallée, FRANCE.
- (3) INRIA Rocquencourt, MicMac project-team, B.P. 105, 78153 Le Chesnay, FRANCE.
- (4) INRIA Lille, SIMPAF project-team, 59650 Villeneuve d'Ascq, FRANCE.

Our stay at the HIM was a great opportunity for us to work together on new projects. The local administrative staff helped us a lot for all practical organization aspects. The workshop we organized[1] gathered an audience of mathematicians from very different horizons, as well as chemists, physicists and biologists working on molecular simulations. Such an interdisciplinary meeting is still a very rare event. It was really appreciated by both the mathematical and applied communities. As another key fact, let us mention that this stay has been crucial to start writing the monograph [6], which is the first book in the applied mathematics community gathering the presentation and the study of all the standard numerical techniques used in molecular dynamics to sample the canonical measure, and compute the free energy. In addition to this monograph, our stay to HIM is directly related to 8 other publications in international peer-reviewed journals, see the references below.

During our stay at HIM, we studied various numerical questions raised by molecular dynamics simulations. Such simulations represent today a huge part of the CPU time used by scientists but this field has not yet been covered a lot by numerical analysts and experts in scientific computing. Below are some examples of achievements where the point of view of applied mathematicians has been useful to understand and improve numerical methods in this field.

1 Sampling methods in molecular dynamics on the computation of free energy

One of the aim of molecular dynamics simulations is to compute averages of some functionals with respect to a given probability measure (statistical ensemble), in order to get macroscopic information from microscopic descriptions. Examples include material properties like constitutive relations, heat capacity, or free energy.

Free energy calculations in molecular dynamics have become an outstanding and increasingly broad computational field in physics, chemistry and molecular biology within the past few years, by making possible the analysis of complex molecular systems. The monograph [6] proposes a new, general and rigorous presentation, intended both for practitioners interested in a mathematical treatment, and for applied mathematicians interested in molecular dynamics. From a mathematical viewpoint, computing a free energy amounts

[1] See `http://cermics.enpc.fr/~stoltz/Bonn/workshop.html`

to computing the marginal of a high-dimensional measure along some low-dimensional degrees of freedom. The difficulty comes from the fact that the high-dimensional measure is typically multimodal, so that standard sampling techniques do not work. Many techniques have been proposed in the literature, and the aim of this monograph is to review these algorithms, and to present new contributions concerning the mathematical analysis of these methods. In particular, we analyze constrained stochastic dynamics to sample measures supported by submanifolds, or adaptive importance sampling methods.

These adaptive importance sampling techniques appear to be very useful in practice, and we are still working on them today. In [7], a generalization of the standard approach to sample constraints in average is proposed. In [3], we adapt these kinds of techniques for the sampling of the posterior distribution (Bayesian inference). Such methods were not known before in the computational statistics community. From a more theoretical viewpoint, the work [5] which was achieved during our stay at HIM is an example of a theoretical result (two-scale criteria for logarithmic Sobolev inequalities) which comes from a question motivated by applications (namely the convergence analysis of adaptive sampling techniques). On this particular subject, we benefited a lot from discussions with Felix Otto.

2 Quantum chemistry and quantum physics

The diffusion Monte Carlo method is a powerful strategy to compute the ground state energy of a Schrödinger Hamiltonian $H = -\frac{1}{2}\Delta + V$ with high accuracy. One of the difficulty for this computation is the so-called fixed node approximation, which introduces a systematic bias in the computation of fermionic ground states. One step towards solving this problem is to understand how to move the nodes in order to reduce the bias. The aim of the paper [8] is to present a probabilistic representation of derivatives with respect to the position of the nodes.

The time at HIM was also inspirational for other works in quantum physics, in particular numerical studies of the thermal properties of carbon nanotubes [9] and a rigorous proof of the Gell-Mann and Low formula, a fundamental ingredient for perturbative approaches in Quantum Field Theory [1, 2].

3 Coarse-graining and effective dynamics

One difficulty in molecular dynamics is the presence of a large spectrum of time scales. Suppose there exists a separation between fast and slow time scales, and that the slow and fast degrees of freedom are known, how to use this information to obtain an effective dynamics on the slow degrees of freedom ? Indeed, it is generally the case that only the slow degrees of freedom are of interest (since they are typically those related to macroscopic, or conformational changes). This question is related to the reduction of dimensionality of dynamical systems.

In [4], we propose an analysis of a coarse-graining technique which uses conditional expectations to close the effective dynamics. Such coarse-graining method is actually used by many authors. The main contribution is an error analysis of such a technique, using relative entropy estimates. Again, this work was mainly achieved during our stay at the HIM.

References

[1] C. Brouder, G. Panati, and G. Stoltz. Adiabatic approximation, gell-mann and low theorem and degeneracies: A pedagogical example. *Phys. Rev. A*, 78:042102, 2008.

[2] C. Brouder, G. Panati, and G. Stoltz. Gell-Mann and Low formula for degenerate unperturbed states. *Ann. I. H. Poincare-Phy*, 10(7):1285–1309, 2010.

[3] N. Chopin, T. Lelièvre, and G. Stoltz. Free energy methods for Bayesian inference: efficient exploration of univariate Gaussian mixture posteriors. *Stat. Comput.*, 2011. To appear.

[4] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *Nonlinearity*, 23:2131–2163, 2010.

[5] T. Lelièvre. A general two-scale criteria for logarithmic Sobolev inequalities. *J. Funct. Anal.*, 256(7):2211–2221, 2009.

[6] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations: A mathematical perspective*. Imperial College Press, 2010.

[7] J.B. Maillet and G. Stoltz. Sampling constraints in average: The example of Hugoniot curves. *Appl. Math. Res. Express*, 2008:abn:004, 2009.

[8] M. Rousset. On a probabilistic interpretation of shape derivatives of Dirichlet ground-states with application to fermion nodes. *ESAIM-Math. Model. Num.*, 2011. To appear.

[9] G. Stoltz, M. Lazzeri, and F. Mauri. Thermal transport in isotopically disordered carbon nanotubes: a comparison between Green's functions and Boltzmann approaches. *J. Phys. Cond. Matter*, 21:245302, 2009.