NANO-D

Algorithms for Modeling and Simulating Nanosystems

Internship Position Opening

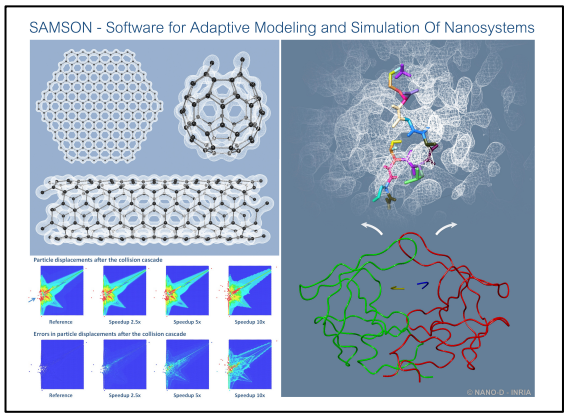
Convex Relaxation for Computational Structural Biology

About the NANO-D research group at INRIA

The NANO-D group, led by Stephane Redon at INRIA Grenoble, develops novel multiscale, adaptive modeling and simulation methods, which automatically focus computational resources on the most relevant parts of the nanosystems under study. All algorithms developed by the group are gathered into SAMSON, an open-architecture software platform designed by NANO-D (SAMSON: Software for Adaptive Modeling and Simulation Of Nanosystems).

During the twentieth century, the development of macroscopic engineering has been largely stimulated by progress in numerical design and prototyping: cars, planes, boats, and many other manufactured objects are nowadays designed and tested on computers. Digital prototypes have progressively replaced actual ones, and effective computer-aided engineering tools have helped cut costs and reduce production cycles of these macroscopic systems.

The twenty-first century is most likely to see a similar development at the atomic scale. Indeed, the recent years have seen tremendous progress in nanotechnology - in particular in the ability to control matter at the atomic scale. Similar to what has happened with macroscopic engineering, powerful and generic computational tools will be employed to engineer complex nanosystems, through modeling and simulation. The NANO-D group is funded through national, regional, and a European ERC Starting Grant (http://nano-d.inrialpes.fr).



Internship topic description

Developing efficient ways to predict structure of individual proteins and their assemblies is a major challenge in computational biology. However, these problems are highly combinatorial and it is NP-complete or NP-hard to find even a reasonable approximate solution of these problems. We strongly believe that the best way to meet these challenges is to use global optimization methods with guaranteed tight bounds to the solution of the original problem. Thus, we propose to adapt the state-of-the-art convex relaxation techniques such as semidefinite relaxation and sum-of-squares to the structure of particular problems that emerge in computational structural biology.

Numerous problems of computational biology and structural bioinformatics can be formulated as combinatorial optimization problems, i.e., optimization programs with discrete (integer or binary) variables. Many of these problems (e.g., Linear and Quadratic Programming with integer variables) are theoretically hard and computationally intractable in general. Some important challenges of structural biology (such as prediction of rotameric states, inverse folding, multi-component protein assembly, etc) can be formulated as a Combinatorial Quadratic Program (CQP) with a vector variable x having 0-1 components:

$$\begin{cases} \min_x & x^T Q_0 x + b_0^T x \\ s.t. & x^T Q_j x + b_j^T x + c_j \begin{cases} \leq 0 \\ or = 0 \end{cases} \qquad j = 1 \ldots M \end{cases} \tag{1}$$

Here, the vector variable x represents realization probabilities of different system states (such as rotameric states, or amino acid types at a given position in the protein fold), the interaction matrix Q0 gives the pair-wise interaction energy between the system states, and the vector b0 gives the interaction energy between each of the system states and the rigid background. The vector x is block-structured, and the subvectors xj may be subject to different constraints such as Σi xij = 1. The binary constraint on the entry xi can be turned into a quadratic one by imposing the condition xi2 = xi. In this formulation the problem becomes a Quadratically Constrained Quadratic Problem (QCQP).

The overall goal of the current project is to study and solve QCQP problems arising in computational structural biology.

Requirements

We are looking for creative, passionate and hard-working individuals with exceptional talent for computer science and mathematics. Excellent oral, written and interpersonal communication skills are essential (working language will be English – knowledge of French is a plus).

- Strong **computer science** background
- Strong knowledge of **general and applied math**
- Strong oral, written and interpersonal **communication skills** (working language: English – knowing French is a plus)
- Good knowledge of **Matlab/ C / C++**
- Ability to work **independently** and with a **team**

Supervisors:
Sergei Grudinin (sergei.grudinin@inria.fr)
Anatoli Juditsky (anatoli.iouditski@imag.fr)

Scolarship:
500,51€ per month, will be readjusted on Sep 1st 2015.

Duration
6-8 months

About Grenoble

Grenoble is the capital city of the French Alps. Combining the urban life-style of southern France with a unique mountain setting, it is ideally situated for outdoor activities. The Grenoble area is today an important centre of industry and science (second largest in France). Dedicated to an ambitious policy in the arts, the city is host to numerous cultural institutions. With 60,000 students (including 6,000 foreign students), Grenoble is the third largest student area in France.