Prof. Suliana Manley - Laboratory of Experimental Biophysics
Institute of Physics of Biological Systems

It is our pleasure to invite you to the following seminar on:

Monday February 24th, 2020 – 14:00 pm– BSP 407, Cubotron

Probabilistic Pipeline and Unsupervised Learning of biomolecule random walks

Jean-Baptiste Masson (PhD)
Decision and Bayesian Computation
Computational biology and Neuroscience department
Institut Pasteur, Paris, France

Twenty years after its inception, the field of Single Molecule (SM) biology undergoes a transition towards a data-generating science [1-3]. At the nanometer scale, the dynamics of individual biomolecules is inherently controlled by random processes, due to thermal noise and stochastic molecular interactions. By accessing the distribution of molecular properties, rather than simply their average value, the great advantage of SM measurements is thus to identify static and dynamic heterogeneities, and rare behaviours.

In recent years, these experimental limits have been progressively alleviated with the advent of new, game-changing methods. Thanks to photoactivatable probes (protein-based or synthetic dyes), millions of individual trajectories can now be recorded in live cells in a few minutes. PALM/STORM images can be reliably acquired over many hours (or even days), yielding up to hundreds of millions of individual localizations.

As SM experiments enter the age of « big data », the development of a proper and unifying statistical framework becomes more necessary than ever. « big data » approaches certainly open up new research venues for our understanding of biological processes, as they enable the inference of molecular dynamics. Yet they also come with a price. Often, adding more data brings both more information and more variability and noise. Specific tools are required to handle the complex structure of results associated to large datasets and to account for the sources of experimental and systemic variability.

Here, we show a global probabilistic pipeline: TRamWAy [4-7] that automatically analyse single molecule experiments from images to random walk analysis. TRamWAy relies on deep neural network to deconvolve single molecule images, Belief propagation coupled to graph summing to perform probabilistic assignments between images, and both supervised and unsupervised Bayesian analysis to extract information from random walks.
We demonstrate the approach on two datasets: Glycine receptors in synapses and GAG dynamics during the formation of the Virion in HIV-1 [4]. We demonstrate two ways of applying the probabilistic pipeline TRamWAy. In the first we use model-based learning with automated results extraction and statistics. In the second we show that unsupervised learning with structured inference allows full analysis without assigning a model to the biomolecules dynamics.

[1] M. El-Beheiry et al, InferenceMAP, **Nat. Meth. 12**, 594–595 (2015)
[2] M. El-Beheiry et al,VISP, **Nat. Meth., 10**, pages 689–690 (2013)
[3] InferenceMAP: https://goo.gl/HiwoxC
[4] Flodorer et al , **Sci. Rep.** (in Press)
[5] Remorino A *et al*, **Cell Rep** 2017 Nov;21(7):1922-1935
[6] Knight S. C *et al,* **Science** vol 350, Issue 6262, p823-826 (2015).
[7] TRamWAy: https://goo.gl/McgJXR