Workshop: statistical methods for Hawkes processes

March 10^{th} 2020

Salle Paul Lévy (113 corridor $16-26$), 4 place Jussieu, 75005 Paris

9 a.m Welcoming

- 9.30-10.45 a.m **Eva Löcherbach**

A short primer on Hawkes processes (linear and non-linear) and their mean-field limits

In this talk I will give a short introduction to the theory both of linear and non-linear Hawkes processes. The talk will cover topics like the construction of such processes by thinning of a dominating Poisson random measure, possible simulation algorithms, stability results and the link with piecewise deterministic Markov processes (PDMP's) for exponential or Erlang-type memory kernels. I will try to give the ideas of some of the proofs in this first part.
In a second part I will then quickly explain how mean-field limits of large systems of interacting Hawkes processes are obtained and how we can prove the emergence of oscillatory behavior in the large system limit.

Coffee break

- 11.15-12 a.m **Simon Clinet**

Quasi likelihood analysis for marked point processes and application to Hawkes processes

We develop a quasi-likelihood analysis procedure for a general class of multivariate marked point processes. As a by-product of the general method, we establish under stability and ergodicity conditions the local asymptotic normality of the quasi-log likelihood, along with the convergence of moments of

quasi-likelihood and quasi-Bayesian estimators. To illustrate the general approach, we then turn our attention to a class of multivariate marked Hawkes processes with generalized exponential kernels, comprising among others the so-called Erlang kernels, and driven by a mark process whose distribution yields a Markovian representation of the global marked point process. We provide explicit conditions on the kernel functions and the mark dynamics under which a certain transformation of the original process is V -geometrically ergodic. We finally prove that the latter result, which is of interest in its own right, constitutes the key ingredient to show that the generalized exponential Hawkes process falls under the scope of application of the quasi likelihood analysis.

- 12.15a.m- 1p.m **Marcello Rambaldi**

  Disentangling and quantifying market participant volatility contributions

  Thanks to the access to labeled orders on the Cac40 index future provided by Euronext, we are able to quantify market participants contributions to the volatility in the diffusive limit. To achieve this result we leverage the branching properties of Hawkes point processes. We find that fast intermediaries (e.g., market maker type agents) have a smaller footprint on the volatility than slower, directional agents. The branching structure of Hawkes processes allows us to examine also the degree of endogeneity of each agent behavior. We find that high-frequency traders are more endogenously driven than other types of agents.

Lunch break ROOM $15 - 25$ 202

- 2.15 - 3 p.m **Judith Rousseau**

  Estimating the interraction functions and the graph of interactions in multivariate Hawkes processes using Bayesian nonparametric methods

  Hawkes processes form a class of point processes describing self and inter exciting/inhibiting processes. There is now a renewed interest of such processes in applied domains and in machine learning, but there exists only limited theory about inference in such models. To be more precise, the intensity function of a univariate Hawkes process has the following form:

$$\lambda(t) = \int_0^{t-} h(t-s)dN_s + \nu$$

where N is the Hawkes process and $\nu > 0$. Multivariate Hawkes processes have a similar intensity function which involves the interractions functions between the different components of the process. In this work we propose a generic

Bayesian non parametric procedure in such models and we study its theoretical properties, both in terms of the estimation of the parameters which are the impulsions and the interractions functions, and in terms of the graph of interactions. As a consequence of these results we also obtain theoretical guaranties for Bayesian tests on the existence of an interraction (or self excitation) function and we extend these results to the case of some non linear Hawkes processes.

Coffee break

- 3.30- 4.15 p.m **Martin Bompaire**

The first part of the presentation will be focused on convex optimization. We will recall the recent progresses made with stochastic first order methods using variance reduction techniques and then show how to adapt these techniques to optimize the log-likelihood of Hawkes processes parametrized with exponential kernels. Indeed, this loss does not meet the gradient-Lipschitz assumption required by the latest first order methods. We thus work under another smoothness assumption, and obtain a linear convergence rate for a shifted version of Stochastic Dual Coordinate Ascent that improves the current state-of-the-art. Besides, such objectives include many linear constraints that are easily violated by classic first order algorithms, but in the Fenchel-dual problem these constraints are easier to deal with. Hence, our algorithm's robustness is comparable to second order methods that are very expensive in high dimensions. Second, we will introduce a statistical learning library for Python called tick, that relies on C++ implementation and state-of-the-art optimization algorithms to provide very fast computations in a single node multi-core setting. Open-sourced and published on Github, this library implements a wide variety of algorithms for Hawkes processes and is easily extensible thanks to its very modular design.

- 4.30-5.15 p.m **Félix Cheysson**

Estimation of Hawkes processes from binned observations using Whittle likelihood

Hawkes processes are a family of stochastic processes for which the occurrence of any event increases the probability of further events occurring. When count data are only observed in discrete time, we propose a spectral approach for the estimation of Hawkes processes, by means of its Bartlett spectrum and Whittle parameter estimation method. To get asymptotic properties for the Whittle's estimator, we prove alpha-mixing properties for the series of counts, using the Galton-Watson properties of the cluster representation of Hawkes processes. Simulated datasets and an application to the incidence of measles in France illustrate the performances of the estimation, notably of the Hawkes excitation function, even when the time between observations is large.