

Post-doctoral position in Strasbourg

Simulation-based inference for marked point processes and application to histology.

Contact:

etienne.birmele@math.unistra.fr
geffray@math.unistra.fr

General Information

Position Type: Postdoctoral position in Statistics

Duration: 24 months, non-renewable, starting between September 2026 and January 2027

Salary: According to the University of Strasbourg pay scale, based on the candidate's experience

Location: Institut de Recherche Mathématique Avancée (IRMA), University of Strasbourg

Supervisors: Etienne Birmelé and Ségolen Geffray

Keywords: Computational statistics, marked point processes, simulation-based inference, topological data analysis, applications in histology.

Scientific Context

HisToGraM Project

This postdoctoral position is part of the interdisciplinary HisToGraM project, funded by Inserm for the period 2025–2029. It brings together probabilists and statisticians from Strasbourg and Calais with an Inserm team of bioinformaticians and physicians from Hôpital Beaujon in the parisian area.

The aim of this project is to develop statistical tools for differential histology, enabling the comparison of tissue structures of an organ under different conditions.

From a mathematical point of view, the data can be represented as colored points in a plane, where each point symbolizes a cell, and its color indicates its type or morphology. Data from the same condition can be viewed as realizations of a marked point process, for instance log-gaussian Cox processes [WGJM16], with parameters depending on the condition. The challenge is to discriminate between different types of point processes and parameterizations, and to determine the most relevant statistics for this purpose. We are particularly interested in statistics derived from topological data analysis [dMDES24], pattern counting [RS14], or geometric extremes [BC20].

These studies require both theoretical and computational developments, especially in simulation-based inference. Real data from pancreatic cancer will be used to develop the method in a relevant way and to test its results in terms of interpretability and tissue classification potential.

Postdoctoral Mission

The main mission of this postdoctoral position will be to develop the project's component related to simulation-based inference [DG84, CBL20], which allows for the consideration of realistic but theoretically intractable point processes. This methodology uses simulation to determine the empirical

distributions of well-chosen statistics, enabling model selection, parameter estimation, or hypothesis testing. It has been successfully applied in ecology [Woo10], systems biology [TS10], and tumor growth modeling [JRTH17].

The recruited post-doc will primarily aim to develop this approach to:

- empirically evaluate the discriminative power of the indicators considered in the project;
- perform model selection and inference among marked point processes;
- construct tests for comparing two tissue populations.

They will also be invited to participate to the HisToGraM meetings and hence in the other aspects of the project.

The work carried out will be published in international peer-reviewed journals or conferences, and the code will be integrated into the R package developed as part of the project.

Application

Applicants should have a PhD in Statistics. They may alternatively have a PhD in Mathematics or in a related field and show a strong motivation to transition into Statistics.

They should provide a detailed CV outlining their academic background, including a list of publications. They should also include a description of their research work (maximum 3–4 pages) and a cover letter explaining their professional goals, scientific interests, and the names and contact details of two referees who can provide letters of recommendation.

Applications may be written in french or in english, and should be submitted by **May 15th**.

References

- [BC20] Gilles Bonnet and Nicolas Chenavier. The maximal degree in a poisson–delaunay graph. *Bernoulli*, 2020.
- [CBL20] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020.
- [DG84] Peter J Diggle and Richard J Gratton. Monte carlo methods of inference for implicit statistical models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 46(2):193–212, 1984.
- [dMDES24] Sebastiano Cultrera di Montesano, Ondrej Draganov, Herbert Edelsbrunner, and Morteza Saghafian. Chromatic topological data analysis. *arXiv preprint arXiv:2406.04102*, 2024.
- [JRTH17] Nick Jagiella, Dennis Rickert, Fabian J Theis, and Jan Hasenauer. Parallelization and high-performance computing enables automated statistical inference of multi-scale models. *Cell systems*, 4(2):194–206, 2017.
- [RS14] Pedro Ribeiro and Fernando Silva. Discovering colored network motifs. In *Complex Networks V: Proceedings of the 5th Workshop on Complex Networks CompleNet 2014*, pages 107–118. Springer, 2014.
- [TS10] Tina Toni and Michael PH Stumpf. Simulation-based model selection for dynamical systems in systems and population biology. *Bioinformatics*, 26(1):104–110, 2010.
- [WGJM16] Rasmus Waagepetersen, Yongtao Guan, Abdollah Jalilian, and Jorge Mateu. Analysis of multispecies point patterns by using multivariate log-gaussian cox processes. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 65(1):77–96, 2016.
- [Woo10] Simon N Wood. Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, 466(7310):1102–1104, 2010.