Extreme statistics for random geometric models

Supervision

- B. Błaszczyszyn, ENS-PSL, département d'informatique, Bartek.Blaszczyszyn@ens.fr
- N. Chenavier, ULCO, LMPA J. Liouville, nicolas.chenavier@univ-littoral.fr

1 Context

Stochastic Geometry (SG) is a branch of probability theory that explores random geometric objects in (not necessarily) Euclidean spaces, including point processes, arrangements of lines, hyperplanes, closed sets, and models constructed from these components [9]. Notably, random geometric graphs, tessellations, and germ-grain models are extensively studied within SG (see Figure 1 for a realization of a Boolean and random tessellations). The Poisson point process [11] is a commonly employed model for point processes. It is characterized by complete independence between random points. This model can be used as an approximation of point processes with weak dependencies, supported mathematically by the Poisson limit theorem.

Conceptually in line with statistical physics, SG focuses on evaluating ensemble averages on large collections of particles with random interactions, deriving macroscopic laws from local interactions. SG has direct connections with percolation theory, random graphs, ergodic theory, integral geometry and convex analysis.

If the initial development of SG is credited to Kolmogorov for proposing models of analyzing crystal growth in materials, it has diverse historical roots in various models, including for instance astronomy and ecology. Modern applications of SG are found in life sciences, image analysis, and wireless communications [2], where Euclidean geometry is crucial. Random graph theory is also widely utilized in computer and general communication sciences.

Within this interconnected approach between theoretical developments and practical applications, SG has been actively pursued in various countries. Notable endeavors are evident in Europe, spanning the Czech Republic, Croatia, Germany, France [1], and the UK. Furthermore, significant contributions have been made by researchers in Australia, India, and the USA. The global community convenes at the biennial conference "SGSIA", with the upcoming edition taking place this year in Germany [2].

[1] The Thematic Network "Stochastic Geometry" `https://gdr-geostoch.math.cnrs.fr/` unites the French community of scientists working on theoretical or applied SG. It fosters collaboration by supporting missions, invitations, and events, including the annual conference

[2] `https://sgsia24.math.kit.edu/`

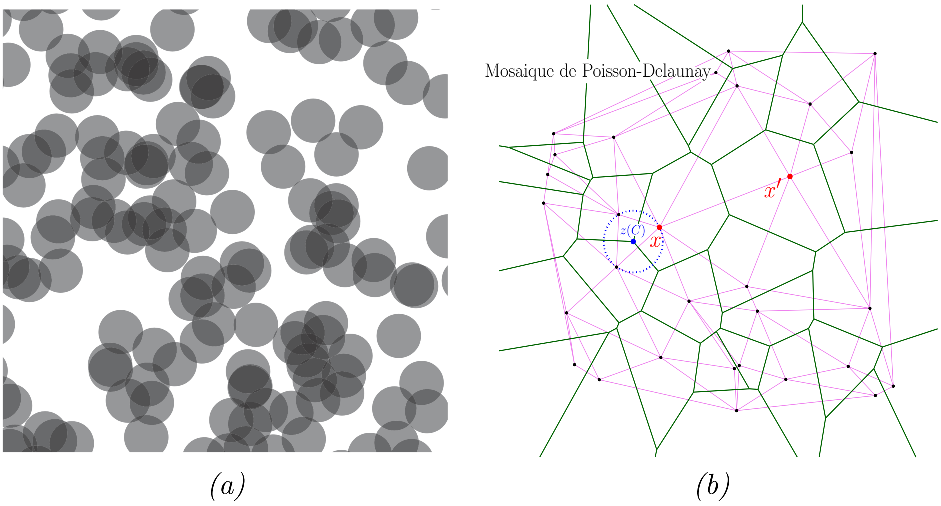

Figure 1: *(a)* A Boolean model. *(b)* The Voronoi (green) and the Delaunay (pink) tessellations.

Spatial statistics are functions that capture the local characteristics of individual points or local elements interacting within more complex random models. Examples of these functions, sometimes called “score functions”, include for instance the degrees of vertices in random geometric graphs formed from point processes (equivalently the number of balls intersecting a given one, see Figure 1 a) or characteristics of cells in a random tessellation (see Figure 1 b). These functions serve as a bridge between theoretical models and the analysis of real-world phenomena.

To grasp the average behavior of these statistics, we begin by utilizing the Law of Large Numbers (LLN). In SG this process involves assessing how the statistic behaves for the so-called typical point or element of the model. Palm theory [9, Chapter 4.4] defines this concept by conditioning the model to have the typical point at the origin. This concept is linked to the LLN through ergodicity.

Following this, the Central Limit Theory (CLT) is employed to characterize the fluctuations of these statistics around the mean. A substantial body of literature explores CLT for geometric statistics in models driven by Poisson or binomial point processes. In this context, various methods are available, including the martingale approach and the Malliavin-Stein method, which leverages the concept of independence, see e.g. [10]. For more general models, approaches based on moments/cumulants are often utilized [5].

Extreme statistics, a specialized field within statistical analysis, is dedicated to comprehending and modeling the behavior of exceptional events or values within a dataset. In our context, it explores the tails of distributions related to spatial statistics. The insights gained from extreme statistics are crucial in various fields of applications where understanding and managing extreme occurrences play a vital role.

2 Research project

Extreme statistics lie at the core of our proposal. Three approaches are identified, and all are related to spatial dependence.

Extreme characteristics of typical objects. Closely related to Palm theory, we delve into the tail distribution of the characteristics of the typical object. An illustrative example is found in the study conducted in [7], which explores the form of large cells within the typical cell of a Poisson-Voronoi tessellation. In such instances, this approach necessitates a comprehensive understanding of the Palm distribution for these models.

A challenging question would insist in going beyond Poisson assumption.

Extreme statistics via LLN and CLT. The typical point or element of the model is a mathematical concept that captures the spatial interaction of the intricate model. When extreme statistics are not amenable to direct probabilistic analysis in the Palm setting, they can still be evaluated using the limiting approach via the LLN and the CLT. These methods provide, at the very least, the spatial frequency of the extreme observations and their Gaussian fluctuations. This represents a rather straightforward application of well-established limiting theory for models in SG, extending well beyond the Poisson/independence assumptions.

One could consider addressing carefully chosen extreme problems, such as examining large values of degrees in geometric models, as in [6], or the size of cells in tessellations within models based on point processes exhibiting attraction or repulsion.

Extreme statistics via Poisson-Boolean approximations. When extreme observations in the large model appear genuinely rare in terms of their low spatial frequency, their analysis might be feasible through a limiting approach using Poisson limit theory. In spatial statistics, this limit assumes a Poisson germ-grain model, where regions with large values of the statistics in question are independently scattered throughout space, see e.g. [8]. This conceptualization not only captures the spatial frequency of these events, identified by the intensity of the germs in the Poisson-Boolean model, but also the geometry of these regions of large values, which are modelized by the grains. This information may not be directly accessible through the Palm study of the statistics of the typical point.

Aldous'book [1] introduces this approach, which leverages various techniques to prove Poisson convergence, including the Chen-Stein method (see e.g. [10]). Similar to the case of the CLT, the Poisson limit is expected to hold for a large class of non-Poisson models exhibiting reasonably weak dependence and/or true sparsity of extreme observations.

This approach could be applied to specific examples of extreme statistics, with a particular emphasis on addressing the geometry of regions exhibiting extreme statistical behavior.

Capturing spatial dependence in SG models is a crucial aspect of limiting analyses, be it the CLT or Poisson limit. This connection is inherently linked to ergodicity, and several sufficient conditions are explored in this domain. These range from classical mixing considerations to various conditional properties and Brillinger mixing. Also, in recent developments in the theory of point processes, the concept of "de-correlation" was introduced in [5]. This concept, based on the evaluation of moment measures, aims to capture the joint "asymptotic independence" of the structure. It has proven useful in establishing the CLT and is expected to be beneficial in proving the Poisson limits mentioned earlier.

When not only asymptotic but also local capturing of dependence is required, a notion involving the comparison of moments and void probabilities with those of a Poisson point process was suggested in [4]. This comprehensive approach aims to capture the impact of positive and negative dependence among points and their effects on characteristics like coverage and percolation in random geometric models. Both the local and asymptotic approaches are naturally applicable to determinantal and permanental processes [3, Chapter 5]. These processes, inspired by fermions and bosons in statistical physics, respectively, exhibit inherent repulsion or attraction in point processes.

Our goal is to explore these two approaches to establish limiting or Palm results for some extreme statistics. The latter technique, directly related to moment measures and void probabilities, may be particularly relevant for studying extreme degree values in random geometric graphs and characteristics of large cells in tessellations.

References

[1] D. Aldous. *Probability approximations via the Poisson clumping heuristic*, volume 77. Springer Science & Business Media, 2013.

[2] F. Baccelli and B. Błaszczyszyn. *Stochastic Geometry and Wireless Networks Volume 2: Applications.* Foundations and Trends in Networking : Vol. 1: Theory, Vol. 2: Applications No 1-2, 2009.

[3] F. Baccelli, B. Błaszczyszyn, and M. Karray. *Random Measures, Point Processes, and Stochastic Geometry.* 2020. In preparation.

[4] B. Błaszczyszyn and D. Yogeshwaran. Clustering comparison of point processes with applications to random geometric models. In V. Schmidt, editor, *Stochastic Geometry, Spatial Statistics and Random Fields: Models and Algorithms*, volume 2120 of *Lecture Notes in Mathematics*, chapter 2, pages 31–71. Springer, 2014.

[5] B. Błaszczyszyn, D. Yogeshwaran, and J. E. Yukich. Limit theory for geometric statistics of point processes having fast decay of correlations. *Ann. Probab.*, 47(2):835–895, 2019.

[6] G. Bonnet and N. Chenavier. The maximal degree in a Poisson-Delaunay graph. *Bernoulli*, 26(2):948–979, 2020.

[7] P. Calka, Y. Demichel, and N. Enriquez. Large planar Poisson-Voronoi cells containing a given convex body. *Ann. H. Lebesgue*, 4:711–757, 2021.

[8] N. Chenavier and M. Otto. Compound poisson approximation under β-mixing and stabilization. *Disponible sur https://arxiv.org/pdf/2310.15009.pdf*, 2023.

[9] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, third edition, 2013.

[10] L. Decreusefond, M. Schulte, and C. Thäle. Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. Probab.*, 44(3):2147–2197, 2016.

[11] G. Last and M. Penrose. *Lectures on the Poisson Process.* Institute of Mathematical Stat. Cambridge University Press, 2017.