Random Graphs and Statistical Inference: New Methods and Applications

Amin Coja-Oghlan (University of Frankfurt),
Pu Gao (University of Waterloo),
Mihyun Kang (TU Graz),
Joon Lee (University of Frankfurt),
Mike Molloy (University of Toronto at Scarborough)

August 8–August 13, 2021

1 Overview of the Field

Since the days of Erdős and Rényi the study of random graphs has developed into a substantial mathematical discipline with close ties to computer science, statistical physics and network science [4]. In computer science random graphs have been used, among other things, as gadgets in order to better understand the complexity of computational tasks. An outstanding example of this is work of Sly on the hardness of counting and sampling problems [8], which connections between complexity theory and statistical physics, particularly properties of the Boltzmann distribution. These properties also play a key role in the statistical mechanics of disordered systems, a burgeoning area that revolves around the idea of replica symmetry breaking suggested by Mézard and Parisi [6]. Additionally, random graphs provide models of complex real-world networks [5].

2 Recent Developments and Open Problems

In recent years applications of random graphs methods in statistical inference problems such as compressed sensing have had a major impact on statistics, machine learning and information theory. Once again random graphs come in as indispensable gadgets in randomised constructions, apart from providing intriguing benchmarks for the study of algorithms. This workshop brought together researchers from the various communities to further mutual understanding and the exchange of ideas.

Statistical physics ideas predict that inference problems generically undergo a “impossible” to “hard” to “easy” evolution as the signal-to-noise ratio increases [9]. Here “impossible” indicates that above a certain noise level the underlying inference task is information-theoretically ill-posed and therefore insoluble. By contrast, “hard” means that the task is information-theoretically feasible but a prohibitive amount of computational resources is required. Finally, “easy” problems can be solved by means of efficient algorithms.

An outstanding scientific problem is to turn these predictions into mathematical theorems. A further task is to investigate the scope of this evolution in the context of realistic network models. Here an exciting question is what degree of universality these predictions can be expected to possess; from a classical random graphs perspective, the question of invoking isoperimetry to fathom the scope of the Erdős-Rényi phase transition provides an excellent example [3]. Any stab at tackling these challenges clearly requires a broad interdisciplinary exchange of ideas, for which the workshop provided a platform.

3 Presentation Highlights

The presentations given by the workshop participants can broadly (albeit not sharply) be categorised into three areas: random graphs and networks, statistics and information theory, and algorithms and complexity. Beginning with random graphs and networks, there were several talks about random geometric graphs, a setting that still poses substantial mathematical challenges (Heydenreich, Frieze, Sly) [7]. A second type of problem, going back to the seminal work of Erdős and Rényi, on which there were several presentations on percolation (Holmgren, Erde). Additionally, there was an outstanding presentation on the modularity of random graphs (Skerman); modularity is a key concept from network science that gauges the decomposition of a network into communities. Furthermore, Cooley presented mathematical work on the Warning Propagation heuristic from statistical physics. Finally, Sen's presentation dealt with challenging large deviations problems on random graphs. This work also pertains to notions of convergence for graph sequences.

Moving on to statistics and information theory, we had talks on the information-theoretic nature of inference problems on random graphs. Loh presented work on the stochastic block model, a prominent benchmark in this area [1]. Furthermore, Aldridge presented random graph designs for the group testing problem, which gained prominence during the pandemic [2]. Further presentations addressed principal component analysis as well as the perceptron model (Zadik, Li, Sellke). On a more methodological note, Bandeira presented work on matrix concentration and free probability theory, a relatively new "non-commutative" generalisation of classical probability theory. Another presentation concerning new probabilistic methods was given by Ramdas. Barbier presented mathematical work on overlap concentration, a fundamental property predicted by physics heuristics. Complementing his contribution, Ravelomanana discussed a concrete inference problem where overlap concentration fails to hold.

Concerning algorithms and complexity, Wein gave an exciting presentation on low-degree hardness for the independent set problem on random graphs, a fundamental question in this area. A different random constraint satisfaction problem was discussed by Zhou. Max Hahn-Klimroth presented new efficient algorithms for the group testing problem. Finally, there was an exciting (but heuristic) contribution by Muntoni on algorithms for mitigation epidemic spreads by means of contact tracing data.

4 Scientific Progress Made

We can say with confidence that the workshop delivered on its main objective of bringing together researchers from different communities concerned with random graphs in a broad sense. The purpose of this "clash of cultures" has been to foster new applications of random graph models as well as the development and adaptation of new mathematical techniques, both algorithmical and analytical. The theory of random graphs has come a long way since its inception, branching out in several distinct directions over the years. Meetings like the present are vital to maintain communication between the different branches so as to facilitate the adaptation of new techniques, themes and models.

5 Outcome of the Meeting

Inspired by the success of the Banff meeting, we have since contributed to several other events that brought together researchers from the different disciplines that were present at the Banff meeting (albeit not necessarily the same individuals). Specifically, we held a follow-up meeting at Oberwolfach ("Random Graphs: Combinatorics, Complex Networks and Disordered Systems", 26 Mar - 31 Mar 2023, Organizers: Amin Coja-Oghlan, Tobias Friedrich, Mihyun Kang, Konstantinos Panagiotou) and Dortmund ("Random Graphs", 4 Sep - 8 Sep 2023, Organizers: Amin Coja-Oghlan, Lena Krieg, Max Hahn-Klimroth, Jan Nagel, Olga Scheftelowitsch). The feedback that we received from the participants of the Banff meeting as well as the aforementioned follow-ups was enthusiastic. In particular, participants emphasised that they met researchers with whom they would otherwise not likely have interacted.

References

[1] E. Abbe: Community detection and stochastic block models: recent developments. arXiv:1703.10146 (2017).

[2] M. Aldridge, O. Johnson, J. Scarlett: Group testing: an information theory perspective. arXiv:1902.06002.

[3] S. Diskin, J. Erde, M. Kang, M. Krivelevich: Isoperimetric inequalities and supercritical percolation on high-dimensional graphs. arxiv 2304.00016 (2023).

[4] P. Erdős, A. Rényi: On the evolution of random graphs. Magayar Tud. Akad. Mat. Kutato Int. Kozl. **5** (1960) 17–61.

[5] R. v. d. Hofstad: Random graphs and complex networks. Cambridge (2017).

[6] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press 2009.

[7] M. Penrose: Random geometric graphs. Oxford (2003).

[8] A. Sly: Computational transition at the uniqueness threshold. Proc. 51st FOCS (2010) 287–296.

[9] L. Zdeborová, F. Krzakala: Statistical physics of inference: thresholds and algorithms. Advances in Physics **65** (2016) 453–552.