Mathematisches Forschungsinstitut Oberwolfach

Report No. 57/2017

DOI: 10.4171/OWR/2017/57

Network Models: Structure and Function

Organised by
Louigi Addario-Berry, Montreal
Shankar Bhamidi, Chapel Hill
Remco van der Hofstad, Eindhoven
Frank den Hollander, Leiden

10 December – 16 December 2017

Abstract. The focus of the meeting was on the mathematical analysis of complex networks, both on how networks emerge through microscopic interaction rules as well as on dynamic processes and optimization problems on networks, including random walks, interacting particle systems and search algorithms. Topics that were addressed included: percolation on graphs and critical regimes for the emergence of a giant component; graph limits and graphons; epidemics, propagation and competition; trees and forests; dynamic random graphs; local versus global algorithms; statistical learning on networks.

Mathematics Subject Classification (2010): 60C05 Combinatorial probability, 60D05 Geometric probability and stochastic geometry, 60K35 Interacting random processes; statistical mechanics type models; percolation theory, 60K37 Processes in random environments, 82B20 Lattice systems (Ising, dimer, Potts, etc.) and systems on graphs, 82B26 Phase transitions (general), 82B27 Critical phenomena, 82B41 Random walks, random surfaces, lattice animals, etc., 82C22 Interacting particle systems.

Introduction by the Organisers

Scientific scope of the meeting

The explosion of empirical network data – ranging from information and transportation networks (Internet, train and road traffic) via biochemical networks (gene expression and regulation) to social networks (Facebook and Twitter) – has had a profound impact on science and society. A variety of mathematical models have been proposed in past years to understand the growth and the evolution

of real-world networks, as well as their robustness against disturbances and attacks. These models, in turn, have been used to analyse dynamical processes on networks, such as algorithms that drive search engines like Google, epidemics for infectious diseases, spread of information in social networks, and distribution of energy and goods. Of particular interest is the interplay between the "structure" of networks, such as degree distribution, connectivity and community structure, and their "function", such as the evolution of processes living on them.

In mathematics, real-world networks are modelled as random graphs, and their performance is modelled by stochastic processes on these graphs. Mathematics has a lot to offer. Practitioners in the applied network sciences are looking for more guidance on current methodologies. Insight comes from a combination of different disciplines, including probability theory, statistics, combinatorics, statistical physics, information theory and algorithmic computer science. Fascinating limiting structures such as continuum random trees, graphons, and graphexes, have been uncovered; these provide unifying perspectives on the study of properties of very large graphs. An open problem session contributed to the open en positive atmosphere in the meeting.

Topics

The lectures focussed on the following research topics:

1. Percolation on graphs.
2. Phase transitions and critical phenomena in networks.
3. Epidemics and propagation on graphs.
4. Exploration of graphs and efficiency of algorithms.
5. Dynamics of graphs and on graphs.
6. Graph limits.
7. Statistical learning.

1. The presentations by Nicolas Broutin, Laura Eslava, Lorenzo Federico, Tim Hulshof and Lutz Warnke all revolved around near-critical percolation on graphs, including the hypercube, the Hamming graph, random regular graphs, inhomogeneous random graphs and dynamical Achlioptas processes on graphs. A detailed description of the largest clusters was obtained, with the aim to unveil universality. Augusto Teixeira spoke about percolation of words. He gave an overview of the history of the problem and settled conjectures about which words are easiest to percolate. David Aldous presented a universality result regarding the percolation threshold of arbitrary large finite graphs. The result states that if a sequence of graphs is such that a giant component emerges at a random time of order 1, then the "incipient" time at which the giant component starts to emerge is deterministic to first order. He then presented a fascinating conjecture which, informally speaking, states that the same phenomenon should occur for epidemic processes with recovery on arbitrary graphs.

2. Markus Heydenreich presented recent results for scale-free percolation on networks. This model combines aspects of the inhomogeneous random graph and long-range percolation on the Euclidean lattice. Scale freeness introduces new

critical exponents. The focus of his talk was on the dichotomy of transience vs recurrence of random walk on such spatial scale-free networks. Clara Stegehuis spoke about graphlets in scale-free random graphs, showing that certain graphlets prevail over others and identifying how their densities scale as well as their typical structure, which can be described in terms of a variational formula characterized by the degree structure of the vertices involved.

3. Mia Deijfen and Julia Komjathy spoke about epidemics on random graphs generated according to the configuration model with heavy-tailed degrees. Deijfen focussed on competing epidemics, showing that depending on the rate at which different types of infections spread along the graph, either one of the types dominates or both types coexist. Komjathy discussed the shortest-weight properties of first passage percolation on such graphs and the role of the explosive nature of its branching-process approximation. Sarah Penington told the audience about Branching Brownian Motion with competitions, and the link with a non-local version of the FKPP-equation. The main goal was to locate the front between high-mass and low-mass regions.

4. David Gamarnik and Ilias Zadik discussed the performance of certain graph algorithms, in particular, the occurrence of a gap between what is theoretically possible and what can be achieved in practice. The gap is linked to a certain geometric non-overlap property of random sets, and describes how well the global maximum can be approximated by local algorithms. Amin Coja-Oghlan described the theory of constraint satisfaction problems on graphs. He showed how to derive sharp satisfiability thresholds, and along the way simplified earlier proofs. Amandine Vber presented a fluid limit theorem for resource allocation algorithms for communication networks; the model she presented has the feature that different clients may evolve on different temporal scales and have very different orders of magnitude.

5. Hakan Guldas presented results for the mixing time of a random walk on a dynamic random graph, initially generated according to the configuration models and subsequently submitted to a random rewiring of the edges. In a regime of fast enough rewiring, the mixing time was computed to leading order. Alexander Fribergh looked at the "ant in the labyrinth": random walks on random high-dimensional percolation clusters. His focus was on universality in the scaling behaviour. Luca Avena spoke about random rooted forests on graphs and showed that they can be used to coarse-grain large networks and improve wavelet algorithms. Justin Salez presented his recent advances in understanding cutoff, mixing time and stationary distributions for random walks on random directed graphs.

6. Christian Borgs and Jennifer Chayes spoke about graph limits, in particular about graphons and graphexes. Both are tools to deal with graphs having large degree, either in the dense regime or in sparser regime. Various applications were discussed. Christina Goldschmidt focussed on scaling properties of the configuration model where the degree distribution has a heavy tail. Gregory Miermont gave

an overview of random triangulations. Svante Janson spoke about exchangeable random graphs and highlighted some of their key properties.

7. Elchanan Mossel described recent joint work with Anuran Makur and Yury Polyanskiy, regarding reconstruction thresholds for information flow on directed acyclic graphs (DAGs). They established that for random 3-regular DAGs, there is a natural phase transition in the width required to remember an input bit, as the noisiness of the system varies. He additionally described an idealized model of deep nets, and presented a result which shows that for certain models of data representation, if the data is sampled in a generative hierarchical way, then it is in fact necessary to use something like deep nets/deep learning in order to build a good classifier. Andrea Montanari presented a collection of results on semidefinite programming on random graphs, including the recovery program for group synchronization on Euclidean lattices and the quality of SDP relaxations for max cut-type problems on regular graphs.

Program of the workshop

Monday 11th December
08:55-09:00 Opening
09:00-10:00 Mia Deijfen
10:15-11:15 David Gamarnik
11:30-12:00 Lorenzo Federico
12:00-12:30 Alexander Fribergh
12:30-13:30 Lunch
13.30-16.30 Work and discussion
16.30-17:00 Julia Komjathy
17.00-17.30 David Aldous
17.30-18:30 Open problem session
18:30-20:00 Dinner

Tuesday 12th December
09:00-10:00 Christina Goldschmidt
10:15-11:15 Svante Janson
11:30-12:30 Gregory Miermont
12:30-13:30 Lunch
13.30-16.30 Hike
18:30-20:00 Dinner

Wednesday 13th December
09:00-10:00 Jennifer Chayes
10:15-11:15 Christian Borgs
11:30-12:00 Hakan Guldas
12:00-12:30 Tim Hulshof
12:30-13:30 Lunch
13.30-16.30 Work and discussion
16.30-17:00 Markus Heydenreich
17.15-17.45 Lutz Warnke

18:00-18:30 Sarah Penington
18:30-20:00 Dinner

Thursday 14th December
09:00-10:00 Amin Coja-Oghlan
10:15-11:15 Nicolas Broutin
11:30-12:00 Laura Eslava
12:00-12:30 Luca Avena
12:30-13:30 Lunch
13:30-16:30 Work and discussion
16:30-17:00 Clara Stegehuis
17:15-17:45 Ilias Zadik
18:00-18:30 Augusto Teixeira
18:30-20:00 Dinner

Friday 15th December
09:00-10:00 Elchanan Mossel
10:15-11:15 Andrea Montanari
11:30-12:00 Amandine Veber
12:00-12:30 Justin Salez
12:30-13:30 Lunch

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows".

Workshop: Network Models: Structure and Function

Table of Contents

Abstracts

Epidemics on general networks: a conjecture

David Aldous

The result in [1], stated there in terms of bond percolation, can be reformulated in terms of the final size of the SI epidemic on a general finite weighted n-vertex graph with initially $o(n)$ infectives. It shows (informally) that in terms of a "virulence" parameter θ, there is always a phase transition between subcriticality (epidemic size $o(n)$ w.h.p.) and supercriticality (epidemic size $\Omega(n)$ w.h.p.). The point is that (essentially) no assumptions about the weighted graph are needed.

We conjecture that analogous "phase transitions occur in completely general networks" results hold for other epidemic-type models where phase transitions are known in the usual specific network models. Here is an explicit conjecture.

An SIS model [contact process]: Given a network (finite connected graph with edge-weights w_e) and a rate function μ_v on vertices v. Introduce a parameter $0 < \theta < \infty$ and a (small) parameter $\varepsilon > 0$.

- Each v is in state S (susceptible) or I (infected); transition rates at v as follows.
- $I \to S$ at rate μ_v.
- $S \to I$ at rate $\varepsilon + \theta \sum\{w_{vy} : \ y \text{ infected } \}$.

Conceptually, you get infected by your contacts with "virulence" parameter θ, or from outside with low probability. Mathematically this is a finite state Markov chain and so has a stationary distribution; we study $X_{\theta,\varepsilon}$ = number of infected vertices, at stationarity.

Now consider a sequence of such networks/rate functions, indexed by n = number of vertices. The basic assumption we will make is: there exist $0 < \theta_* < \theta^* < \infty$ such that, for every sequence $\varepsilon_n \downarrow 0$ sufficiently slowly,

$$X^{(n)}_{\theta_*,\varepsilon_n} = o(n) \text{ in probability}; \quad X^{(n)}_{\theta^*,\varepsilon_n} = \Omega(n) \text{ in probability}.$$

Conjecture 1. *Under the above assumption (and perhaps further but weak assumptions), there exist $\theta_n \in [\theta_*, \theta^*]$ such that, for all $\varepsilon_n \downarrow 0$ sufficiently slowly,*

$$X^{(n)}_{\theta_n-\delta,\varepsilon_n} = o(n) \text{ in probability}; \quad X^{(n)}_{\theta_n+\delta,\varepsilon_n} = \Omega(n) \text{ in probability } \forall \delta > 0.$$

References

[1] D. Aldous, *The Incipient Giant Component in Bond Percolation on General Finite Weighted Graphs*, Electron. Commun. Probab **21** (2016), paper no. 68.

Data sets on networks: some probabilistic tools

Luca Avena

(joint work with Alexandre Gaudillière, Fabienne Castell, Clothilde Mélot)

Inspired by combinatorial problems from statistical physics, in the recent paper [1] we characterized properties and efficient sampling algorithms of certain random spanning forests on arbitrary weighted finite graphs. These objects are related to fundamental algebraic and probabilistic structures of a given weighted graph (or of the corresponding adjacency matrix). These studies branched into questions of a different nature and led to a number of applications within the analysis of real-world networks. This talk is on three main such applications recently obtained: 1) a downsampling procedure for "well-distributed" nodes in a network [1, 4]; 2) coarse-graining or reduction schemes for graphs and associated processes [2, 4]; 3) pyramidal wavelets-like algorithms for processing arbitrary signals on graphs [2, 4]. These applications are based on random forests together with other probabilistic tools, including the so-called Wilson's algorithm, coupling techniques and Markov chain intertwining dualities.

References

[1] L. Avena and A. Gaudillière, *Two applications of random spanning forests*, Journal of Theoretical Probability, DOI: 10.1007/s10959-017-0771-3 (2017).

[2] L. Avena, F. Castell, A. Gaudillière and C. Mélot, *Approximate and exact solutions of intertwining equations through random spanning forests*, PREPRINT, arXiv.org:1702.05992 (2017).

[3] L. Avena, F. Castell, A. Gaudillière and C. Mélot, *Intertwining wavelets or Multiresolution analysis on graphs through random forests*, PREPRINT, arXiv.org:1707.04616 (2017).

[4] L. Avena, F. Castell, A. Gaudillière and C. Mélot, *Random Forests and Networks Analysis*, arXiv:1711.01635 (2017).

Graphons and Graphexes as Limits of Sparse Graphs: Parts I and II

Christian Borgs, Jennifer Chayes

(joint work with Henry Cohn, Shirshendu Ganguly, Nina Holden, Christina Lee, Laci Lovasz, Devavrat Shah, Adam Smith, Victor Veitch, Yufei Zhao)

Graphons and graphexes are limits of graphs which allow us to model and estimate properties of large-scale networks. In this pair of talks, we review the theory of dense graph limits, and give two alternative theories for limits of sparse graphs, one leading to unbounded graphons over probability spaces, and the other leading to bounded graphons (and graphexes) over σ-finite measure spaces. Talk I, given by Jennifer, will review the general theory, highlight the unbounded graphons, and show how they can be used to consistently estimate properties of large sparse networks. This talk will also give an application of these sparse graphons to collaborative filtering on sparse bipartite networks. Talk II, given by Christian, will recast limits of dense graphs in terms of exchangeability and the Aldous Hoover Theorem, and generalize this to obtain sparse graphons and graphexes as limits of

subgraph sampled from sparse graph sequences. This will provide a dual view of sparse graph limits as processes and random measures, an approach which allows a generalization of many of the well-known results and techniques for dense graph sequences.

References

[1] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions. Preprint, arXiv:1401.2906, 2014.

[2] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence II: LD convergence, quotients, and right convergence. *Ann. Probab., to appear*, 2018. Preprint available at arXiv:1408.0744.

[3] C. Borgs, J. T. Chayes, H. Cohn, and N. Holden. Sparse exchangeable graphs and their limits via graphon processes. *Journal of Machine Learning Research, to appear*, 2018. Preprint available at arXiv:1601.07134.

[4] C. Borgs, J. T. Chayes, H. Cohn, and V. Veitch. Sampling perspectives on sparse exchangeable graphs. Preprint, arXiv:1708.03237, 2017.

[5] C. Borgs, J. T. Chayes, H. Cohn, and S. Ganguly. Consistent nonparametric estimation for heavy-tailed sparse graphs. Preprint, arXiv:1508.06675, 2015.

[6] C. Borgs, J. T. Chayes, and A. Smith. Private graphon estimation for sparse graphs. In *Advances in Neural Information Processing Systems 28*, pages 1369–1377, 2015. Extended version with proofs available at arXiv:1506.06162.

[7] C. Borgs, J. Chayes, C. E. Lee, and D. Shah. Thy friend is my friend: Iterative collaborative filtering for sparse matrix estimation. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 4718–4729. 2017. Extended version with proofs available at arXiv:1712.00710.

[8] C. Borgs, J. T. Chayes, H. Cohn, and L.M. Lovász. Identifyability for graphexes and the weak kernel metric. Preprint, in preparation, 2018.

Scaling limits of inhomogeneous random graphs

Nicolas Broutin

(joint work with Thomas Duquesne and Minmin Wang)

1. Poisson inhomogeneous random graphs

We consider the following rank-1 inhomogeneous random graph model. For an integer $n \geq 1$, let $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ where $w_1 \geq w_2 \geq \cdots \geq w_n \geq 0$, be a weight sequence and define $\sigma_1 = \sigma_1(\mathbf{w}) = \sum_i w_i$. Given $\mathbf{w}$, one constructs a random graph $G(\mathbf{w}) = (V, E)$ on $V = \{1, 2, \ldots, n\}$ where each of the $\binom{n}{2}$ edges is present in E independently of the others, and the edge $\{i, j\}$ is included with probability $1-\exp(w_i w_j/\sigma_1)$. The model has a long history. It was first considered by Aldous [4] as an inhomogeneous version of the classical Erdős–Rényi random graph, and then further studied by Aldous and Limic [5]. It also appears in the works of Norros and Reittu [7]. In general, one could actually think of a graph process $G_t(\mathbf{w})$ where edges are included with probability $1 - \exp(t w_i w_j/\sigma_1)$. As t increases in $[0, \infty)$, one observes a sudden change in structure at some t_c: for

$t > t_c$, the graph contains with high probability a 'giant' connected component containing a linear proportion of the vertices. We are interested in the behaviour at the critical point t_c, and more specifically in the scaling limit of the graph seen as a measured metric space where the metric is the graph distance, and the measure is the counting measure on the vertices. The topologies of interest include the Gromov–Prokhorov (GP) or the Gromov–Hausdorff–Prokhorov (GHP) topologies, which somewhat correspond to convergence of the finite dimensional distribution and uniform convergence for the graphs, respectively. In the following, we assume that $\mathbf{w}$ is scaled so that $t_c = 1$, and drop all reference to t altogether.

2. The critical regime and scaling limits

Weight of connected components. Aldous [4] and then Aldous and Limic [5] considered the model in relation to the multiplicative coalescent, and were mostly interested in the $\mathbf{w}$-mass (sum of the weights) of the connected components. Let (W_t) be a standard Brownian motion, $\mathbf{c} = (c_1, c_2, \dots) \in \ell^3$ and $\alpha \in \mathbb{R}$, $\beta, \kappa \geq 0$. Then define

$$Y_t = -\alpha t - \frac{1}{2}\kappa\beta t^2 + \sqrt{\beta} W_t + \sum_i c_i(\mathbf{1}_{\xi_i \leq t c_i t\kappa} - c_i \kappa t), \tag{1}$$

where (ξ_i) are i.i.d. exponential r.v. with mean one. Then the Aldous–Limic limits are given in terms of the durations of the excursions of $Y_t - \inf\{Y_s : s \leq t\}$ away from zero.

Mass of connected component. Bhamidi, van der Hofstad and van Leeuwaarden [11, 14] proved that the scaling limit for the number of nodes in the largest connected components are actually the same as the $\mathbf{w}$-weight considered by Aldous and Limic when $c_i = 0$ for all i, or when c_i are regularly varying and $\beta = 0$.

Metric of the connected components. The first results concerning the metric were proved by Addario-Berry, B. and Goldschmidt [1, 2] for the special case where all the weights are identical (which corresponds to the classical Erdős–Rényi random graph model). They proved in particular that the scaling limit consists of a sequence of compact measured metric space that one can construct from (Y_t), in the special case when $c_i = 0$ for all $i \geq 1$, and an additional independent Poisson point process. The limit is thus purely 'Brownian'. The connected components are described as random real trees encoded by the excursions of the reflected process, in which one identifies a finite number of pairs of points given by the Poisson point process to create the graph. The law of the spanning trees may also be described as a change of measure of the Aldous' continuum random tree [3].

Bhamidi, Sen and X. Wang [12] then considered the regime where the inhomogeneity is weak enough that it does modify the nature of the limit (that is $\mathbf{c} = \mathbf{0}$), and the weights only affect the result via deterministic multiplicative constants. The proof relies on a description of the spanning tree in terms of birthday trees (a weighted version of uniform trees) [6]. The strongly inhomogeneous regime has been considered by Bhamidi, van der Hofstad and Sen [13], and also relies on birthday trees and their scaling limits, inhomogeneous continuum random trees [6].

While the previous results already give a pretty good picture of the limits of $G(\mathbf{w_n})$ for sequences of weights of increasing lengths, many aspects are still poorly understood. The first one is the set of possible limits: One also expects that all the limits obtained by Aldous and Limic in terms of (Y_t) defined in (1) should lift to limit metric spaces, but the previous results only concern specific cases. Furthermore, the proofs for the metric limits are regime specific (either 'weakly' or 'strongly' inhomogeneous), while the Aldous-Limic results treat the amount of inhomogeneity transparently. Even for the limits that have been constructed, many properties remain intriguing such as compactness of the connected components, the fractal dimensions, and the extend of the basin of attraction.

3. Results and proof ideas

From every process (Y_t) as in (1) we construct a sequence of measured metric spaces. The metric is given by the process of local times

$$H_t = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}(Y_s \leq \inf_{s \leq u \leq t} Y_u + \epsilon) ds \,.$$

The very fact that the limit above exist is non-trivial. The excursion of the process (H_t) away from zero each encode a tree, the tree is then modified by creating finitely many "short-cuts" whose location are given by the marks of a Poisson point process with rate κ under the excursions of Y_t above its infimum process (which are in one-to-one correpondence with the excursions of H_t). The compactness and fractal dimensions of the metric space are determined in terms of a Laplace exponent associated to (Y_t).

We also obtain limit theorems: for any limit object above, there exists a sequence of $(\mathbf{w}_n)$ such that the corresponding graphs $G(\mathbf{w}_n)$, suitably rescaled converge to it; and there are also sequences for which the convergence takes place for GP but not for the stronger GHP topology, and this even if the limit is compact.

The central idea consists in a novel representation of the discrete inhomogeneous random graphs as embedded inside a Galton–Watson forest. Intuitively, we build an infinite graph on $\mathbb{N}$, and then prune it in such a way to (a) keep only one copy of each of the weights, and (b) make sure that the metric inside the remaining part is just the retriction of the metric on $\mathbb{N}$. The limit graphs are then constructed on a subforest of a Levy forest. This makes it possible to leverage the results on the scaling limits of Galton–Watson forests by Le Gall & Le Jan [10] and Duquesne & Le Gall [8] and avoid the need to deal frontally with the law of the graph. In particular, this allows to deduce transparently a condition for the compactness of the connected components, as well as their fractal dimension (Hausdorff and packing) from the corresponding results for Levy trees (and the Laplace exponent mentioned above is that of the Levy process in which (Y_t) is embedded).

References

[1] L. Addario-Berry, N. Broutin,, and C. Goldschmidt, The continuum limit of critical random graphs. *Probab. Theory Related Fields 152*, 3-4 (2012), 367–406.

[2] L. Addario-Berry, N. Broutin,, and C. Goldschmidt, Critical random graphs: limiting constructions and distributional properties. *Electronic Journal of Probability 15* (2010), 741–774.

[3] D. Aldous, The continuum random tree. I. *The Annals of Probability 19* (1991), 1–28.

[4] D. Aldous, Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab. 25*, 2 (1997), 812–854.

[5] D. Aldous and V. Limic, The entrance boundary of the multiplicative coalescent. *Electron. J. Probab. 3* (1998), No. 3, 59 pp.

[6] D. Aldous and J. Pitman, Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probability Theory and Related Fields 118* (2000), 455–482.

[7] I. Norros and H. Reittu, On a conditionally Poissonian graph process. *Advances in Applied Probability 38* (2006), 59–75.

[8] T. Duquesne, and J.-F. Le Gall, Random trees, Lévy processes and spatial branching processes. *Astérisque*, 281 (2002), vi+147.

[9] T. Duquesne, and J.-F. Le Gall, Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields 131*, 4 (2005), 553–603.

[10] J.-F. Le Gall and Y. Le Jan, Branching processes in Lévy processes: the exploration process. *Ann. Probab. 26*, 1 (1998), 213–252.

[11] S. Bhamidi, R. v. d. Hofstad and J. van Leeuwaarden, Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electronic Journal of Probability 15* (2010), 1682–1702.

[12] S. Bhamidi, S. Sen, and X. Wang, Continuum limit of critical inhomogeneous random graphs. Probability Theory and Related Fields (to appear).

[13] S. Bhamidi, R. v. d. Hofstad and S. Sanchayan, The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. Probability Theory and Related Fields (to appear), 2015.

[14] S. Bhamidi, R. v. d. Hofstad and J. van Leeuwaarden, Novel scaling limits for critical inhomogeneous random graphs. *The Annals of Probability 40* (2012), 2299–2361.

Random linear equations

Amin Coja-Oghlan

(joint work with Peter Ayre, Pu Gao, Noela Muller)

Suppose that $\mathbf{A}$ is a random $m \times n$ matrix over the field $\mathbb{F}_q$ for a prime power q with precisely $k \geq 3$ non-zero entries per row. Choose $\mathbf{y} \in \mathbb{F}_q^m$ uniformly at random. The main result of this talk provides the precise threshold value m/n up to which the random linear system $\mathbf{A}x = \mathbf{y}$ possesses a solution.

Theorem 1. *Let $k \geq 3$, let $q > 1$ be a prime power and let P be a permutation-invariant distribution on $\mathbb{F}_q^{*\,k}$. Set*

$$\rho_{k,d} = \sup\left\{x \in [0,1] : x = 1 - \exp(-dx^{k-1})\right\} \quad \text{for } d > 0, \tag{1}$$

and define

$$d_k = \inf\left\{d > 0 : \rho_{k,d} - d\rho_{k,d}^{k-1} + (1-1/k)d\rho_{k,d}^{k} < 0\right\}. \tag{2}$$

Suppose that $m = m(n)$ is a sequence such that $km/n \sim d$. Then

$$\lim_{n\to\infty} \mathbb{P}\left[\exists x \in \mathbb{F}_q^n : \mathbf{A}x = \mathbf{y}\right] = \begin{cases} 1 & \text{if } d < d_k, \\ 0 & \text{if } d > d_k, \end{cases}$$

and thus for $d < d_k$

$$\lim_{n\to\infty} \mathbb{P}\left[\operatorname{rank}(\mathbf{A}) = m\right] = 1. \tag{3}$$

As a corollary we obtain a formula for the random of the random matrix $\mathbf{A}$.

Corollary 1. *Let $k \geq 3$, let $q > 1$ be a prime power and let P be a permutation-invariant distribution on $\mathbb{F}_q^{*\,k}$. Then for any $d > d_k$, as $n \to \infty$ we have*

$$\frac{\operatorname{rank}(\mathbf{A})}{m} \longrightarrow 1 + \frac{k}{d}\rho_{k,d} - k\rho_{k,d}^{k-1} + (k-1)\rho_{k,d}^{k} \qquad \textit{in probabiliy.}$$

These results extend and generalise prior work on special cases. Particularly the case $q = 2$, known as the random k-XORSAT problem, received considerable attention [1, 5, 6, 8, 9]. More precisely, in the case $q = 2$ and $k = 3$ the satisfiability threshold was determined by Dubois and Mandler [6] via the second moment method. This argument was subsequently extended to $k > 3$ [5, 9], although this extension required rather technical calculations. Extensions to the cases $q = 3, 4$ were obtained via a technically even more intricate second moment argument as well [7].

By contrast, the present proof is based on a different technique inspired by ideas from mathematical physics, particularly the so-called Aizenman-Sims-Starr scheme [2] and its adaptation to statistical inference problems [4]. The full details of the proof can be found in [3].

References

[1] D. Achlioptas, M. Molloy: The solution space geometry of random linear equations. Random Structures and Algorithms **46** (2015) 197–231.

[2] M. Aizenman, R. Sims, S. Starr: An extended variational principle for the SK spin-glass model. Phys. Rev. B **68** (2003) 214403.

[3] P. Ayre, A. Coja-Oghlan, P. Gao, N. Müller: The satisfiability threshold for random linear equations. arXiv:1710.07497 (2017).

[4] A. Coja-Oghlan, F. Krzakala, W. Perkins and L. Zdeborova: Information-theoretic thresholds from the cavity method. Proc. 48th STOC (2017) 146–157.

[5] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, M. Rink: Tight thresholds for cuckoo hashing via XORSAT. arXiv:0912.0287v3 (2010).

[6] O. Dubois, J. Mandler: The 3-XORSAT threshold. Proc. 43rd FOCS (2002) 769–778.

[7] A. Goerdt, L. Falke: Satisfiability thresholds beyond k-XORSAT. Proc. 7th International Computer Science Symposium in Russia (2012) 148–159.

[8] M. Ibrahimi, Y. Kanoria, M. Kraning, A. Montanari: The set of solutions of random XOR-SAT formulae. Annals of Applied Probability **25** (2015) 2743–2808.

[9] B. Pittel, G. Sorkin: The satisfiability threshold for k-XORSAT. Combinatorics, Probability and Computing **25** (2016) 236–268.

Competing first passage percolation on the configuration model

Mia Deijfen

(joint work with Daniel Ahlberg, Remco van der Hofstad, Svante Janson)

Competing first passage percolation describes the growth of two competing infections on an underlying graph structure. We study the model on a graph generated by the configuration model, that is, each vertex $i \in \{1, \dots, n\}$ is equipped with D_i half-edges, where $\{D_i\}$ are independent and identically distributed, and the half-edges are then paired uniformly at random to create edges. Starting from two vertices chosen uniformly at random, the infections spread via the edges in the graph in that an uninfected vertex becomes type 1 (2) infected at rate λ_1 (λ_2) times the number of nearest neighbors of type 1 (2).

Our assumptions on the degree distribution ensure that the graph contains a giant component occupying all but a vanishing fraction of the vertices as $n \to \infty$, and hence that almost all vertices will with high probability be infected when the process terminates. The question that we will be interested in is the outcome of this competition. Specifically, will both types occupy a strictly positive fraction of the vertices in the limit as $n \to \infty$?

When $\mathbb{E}[D^2] < \infty$, we show that the answer is yes if $\lambda_1 = \lambda_2$. Specifically, the fraction of vertices that are ultimately infected by type 1 then converges to a continuous random variable $V \in (0, 1)$ as $n \to \infty$. If $\lambda_1 \neq \lambda_2$, on the other hand, then the type with the larger intensity occupies all but a vanishing fraction of the vertices. When the degrees obey a power law with exponent $\tau \in (2, 3)$ – so that the mean is finite but the variance infinite – then with high probability one of the infection types will occupy all but a finite number of vertices. Furthermore, which one of the infections wins is random and both infections have a positive probability of winning regardless of the values of λ_1 and λ_2.

In both cases, the initial growth of the infections can be approximated by continuous time branching processes and it turns out that the type that is in the lead after this phase will win the competition. When the degree distribution has infinite variance, the approximating branching processes will have infinite mean (due to size biasing). This means that they explodes in finite time and both types will then have a positive probability of exploding first. When the degrees have finite variance, the branching processes will have finite mean and the relation between the growth rates will then determine the outcome.

The methods controlling the growth after the branching process phase are specific for exponentially distributed passage times. A natural extension would be to study more general passage time distributions, possibly different for the two types. Another extension would be to study more general initial conditions, where the initial number of one or both types may grow with n or where the vertices are chosen based on degree. Is it for instance possible for a weaker type in the finite variance case to capture a positive fraction of the vertices if it can start from one or more high degree vertices, while the stronger type starts from a vertex with small degree?

References

[1] D. Ahlberg, M. Deijfen, S. Janson, *Competing first passage percolation on random graphs with finite variance degrees*, arxiv:1711.02902.

[2] M. Deijfen and R. van der Hofstad, *The winner takes it all*, The Annals of Applied Probability **26** (2016), 2419–2453.

The size of the giant component in the random *d*-process

Laura Eslava

(joint work with Lutz Warnke)

Graph processes $(G(i), i \geq 0)$ are usually defined as follows. Starting from the empty graph on n vertices, at each step i a random edge is added from a set of available edges. For the d-process, edges are chosen uniformly at random among all edges joining vertices of current degree at most $d-1$.

The fact that, during the process, vertices become *saturated* when reaching degree d makes the process depend heavily on its history. However, it shares several qualitative properties with the classical Erdos-Rényi graph process. For example, there exists a critical time t_c at which a giant component emerges, whp (that is, the largest component in $G(tn)$ goes from logarithmic to linear order).

In this talk we consider $d \geq 3$ fixed and describe the growth of the size of the giant component. In particular, we show that whp the largest component in $G((t_c + \varepsilon)n)$ has asymptotic size cn, where $c \sim c_d \varepsilon$ is a function of time ε as $\varepsilon \to 0+$.

The growth, linear in ε, is a new common qualitative feature shared with the Erdos-Rényi graph process and can be generalized to hypergraph processes with different max-allowed degree sequences. This is work in process jointly with Lutz Warnke.

References

[1] L. Eslava, L. Warnke *The size of the giant component in the random d-process*, in preparation.

[2] L. Warnke, N. Wormald, *The phase transition in random graph processes*, in preparation.

Critical Percolation on the Hamming graph

Lorenzo Federico
(joint work with Remco van der Hofstad, Frank den Hollander, Tim Hulshof)

Hamming graph percolation

The d-dimensional Hamming graph $H(d,n)$ is defined as the cartesian product of d complete graphs on n vertices each. Note that for specific choices of the parameters we obtain two very famous models: percolation on $H(1,n)$ is an Erdős-Rényi random graph (ERRG), while percolation on $H(d,2)$ is hypercube percolation. We study the asymptotic behaviour for d fixed and $n \to \infty$. The main goal of this project is to compare the critical phase of percolation to critical ERRG.

The critical point. The first result obtained is the determination of the critical point, done in [2]. We proved that the critical point for percolation on $H(d,n)$, has an expansion in negative powers of the degree m, similar to the one derived by van der Hofstad and Slade in [4] for percolation on the hypercube or on $\mathbb{Z}^d$. We obtain the following formula for the critical window

$$p_c^{(d)} = \frac{1}{m} + \frac{2d^2-1}{2(d-1)^2}\frac{1}{m^2} + O(m^{-3}) + O(m^{-1}V^{-1/3}),$$

which is asymptotically exact for $d \leq 6$. This is the first known exact determination of the critical window for percolation on finite graphs. The lower bound on $p_c^{(d)}$ is obtained with a branching process domination of the cluster exploration that explicitly exploits the geometric structure of $H(d,n)$, in particular the fact that all lines are complete graphs and thus in the percolation model they are equivalent to ERRG, while the upper bound is obtained via lace expansion.

The critical scaling limit. The other important result is the determination of the scaling limit for component sizes and surplus edges at criticality. Aldous in [1] derived it for the ERRG and since then it has been extended to a wide variety of non-geometric models, yet this is the first extension to a random graph model with non-trivial underlying geometry. The difficulty lays in the fact that the exploration algorithms typically used to derive these limits fail in presence of geometry, due to the lack of independence in the positions of vertices of a cluster. We solve this problem using a coupling between percolation and Branching Random Walks (which is defined for general regular graphs), exploiting the very fast mixing time of a RW on $H(d,n)$. We thus derive the following scaling limit for the sizes and surplus edges of the largest critical components, as $d = 2,3,4$ and $n \to \infty$

$$\Big(\big(n^{-2d/3} |\mathcal{C}^\lambda_{(j)}|, n^{1-2d/3} \mathbf{Sp}(\mathcal{C}^\lambda_{(j)}) \big) \Big)_{j\geq 1} \xrightarrow{d} \Big(\big(\mathbf{C}^\lambda_j, \frac{1}{2(d-1)^2} \mathbf{C}^\lambda_j \big) \Big)_{j\geq 1}.$$

The scaling limit for component sizes is identical to the one of the ERRG (even with the same limiting variables), while the number of surplus edges, instead of converging in distribution to a tight random variable, is diverging and is concentrated conditioned on component size. This is due to the fact that the lines of the

Hamming graph are complete graphs, and thus after percolation form some denser subgraphs, which contain a lot more short cycles than the average over the entire graph. This is coherent with the general idea that geometry increases clustering in random graphs.

We also have strong evidence that if we only consider the surplus created by long cycles this has a similar scaling to the surplus of the ERRG.

References

[1] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, **25**(2):812–854, (1997).
[2] L. Federico, R. v. d. Hofstad, F. d. Hollander, and T. Hulshof. Expansion of percolation critical points for Hamming graphs. arXiv:1701.02099, Preprint (2017).
[3] L. Federico, R. v. d. Hofstad, F. d. Hollander, and T. Hulshof. The scaling limit for critical percolation on the Hamming graph. In preparation
[4] R. v. d. Hofstad and G. Slade. Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: the first three terms. *Combin. Probab. Comput.*, **15**(5):695–713, (2006).

The ant in high-dimensional labyrinths

Alexander Fribergh
(joint work with Gérard Ben Arous, Manuel Cabezas)

One of the most famous open problem in random walks in random environments is to understand the behaviour of a simple random walk on a critical percolation cluster, a model known as the ant in the labyrinth. We will present new results on the scaling limit for the simple random walk on the critical branching random walk in high dimension which converges, after scaling, to the brownian motion on the integrated super-brownian motion. In the light of lace expansion, we believe that the limiting behaviour of this model should be universal for simple random walks on critical structures in high dimension. In particular, recent progress show that similar results hold for lattice trees.

References

[1] G. Ben Arous, M. Cabezas and A. Fribergh, *Scaling limit for the ant in a simple labyrinth*, arxiv 1609.03979
[2] G. Ben Arous, M. Cabezas and A. Fribergh, *Scaling limit for the ant in high-dimensional abyrinths*, arxiv 1609.03980

(Arguably) Hard on Average Constraint Satisfaction Problems

David Gamarnik

Many combinatorial optimization problems defined on random instances such as random graphs, exhibit an apparent gap between the optimal value, which can be estimated by non-constructive means, and the best values achievable by fast (polynomial time) algorithms. As a canonical example consider the problem of finding a largest clique in the random graph $\mathbb{G}(n,p)$. It is known that the largest clique has an asymptotic size $2\log_{1/p} n$, while the best (Greedy) algorithm can only construct a clique with size $\log_{1/p}(n)$. The multiplicative factor 2 gap remains open since Karp posed this as a challenge in 1976 [8]. Similarly, for sparse random graphs $\mathbb{G}(n,d/n)$ with edge probability d/n, a largest independent set has size $2(\log d/d)n$ asymptotically as $n\to\infty$, and then $d\to\infty$, in this order, while the best (again Greedy) algorithm finds an independent set with size only $(\log d/d)n$, namely factor 2 smaller than the optimum.

As a third example, consider a random instance of the K-SAT problem. The K-SAT problem is the canonical NP-complete problem on n boolean variables, $x_1,\ldots,x_n$ defined by constructing a conjuction $C_1\wedge\cdots\wedge C_m$ of m clauses C_j, each of which is a disjuction of the form $C_j = a_{j1}\vee\cdots\vee a_{jK}$, and each a_{ij} is one of the n variables $x_1,\ldots,x_n$ or one of their n negations $\bar{x}_1,\ldots,\bar{x}_n$. Suppose an instance of a K-SAT problem is constructed uniformly at random (in some appropriate sense). It is known that when the clause density $m/n = d$ is approximately $2^K\log 2 - O(1)$, the instance admits a satisfying assignment [1],[4]. However, the best algorithmic result can only construct a satisfying assignment when $d\le (2^K/K)\log K$, namely factor $K/\log K$ smaller that the best possible [3].

Similar gaps abound in many modern models of high dimensional statistics and machine learning. For example, consider the regression model of the form $Y = X\beta^* + W$, where X is $n\times p$ has i.i.d. standard Gaussian entries, W has i.i.d. $N(0,\sigma)$ Gaussian entries, and β^* is some unknown regression vector to be recovered from observed X and Y. It is known that if β^* is k-sparse, namely it has at most k non-zero entries and every non-zero entry is of constant order, then, provided that n is roughly at least $2k\log p$, the vector β^* can be reconstructed by means of convex relaxation techniques such as LASSO or Dantzig Selector (under some additional assumptions on k,n,p and σ), while these techniques provably fail when $n\le 2k\log p$. At the same time, the brute force search can reproduce β^* when $n\ge 2k\log p/\log(1+k/\sigma^2)$. Below this threshold, however, reconstructing β^* is impossible information theoretically. In particular, when σ is of constant order, the gap between the best algorithmically tractable technique and the best result achievable information theoretically is of the logarithmic $(\log(1+k/\sigma^2))$ order.

Through a combined effort of mathematicians, computer scientists and statistical physicists, it became apparent that a potential and in some cases a provable obstruction for designing algorithms bridging the algorithmic gap illustrated by several examples above, is an intricate geometry of nearly optimal solutions, in particular the presence of a certain Overlap Gap Property (OGP). This property

can be defined roughly as follows. Most of the models above can be described in general as the problem of minimizing some cost function $F(\sigma, Z)$ over some decision variables σ where Z is a random "disturbance" creating the random objective function F. In the case of cliques in $\mathbb{G}(n,p)$, for example, $\sigma \in \{0,1\}^n$ is an encoding of a clique, and $Z \in \{0,1\}^{\binom{n}{2}}$ encodes the status of each edge (open/close). In the case of the regression problem $W = (X, Y)$ and $F(\beta, W) = \|Y - X\beta\|_2$, and the minimization $\min_\beta \|Y - X\beta\|_2$ is over all k-sparse vectors $\beta \in \mathbb{R}^p$. We note that in this example, the disturbance W depends on some hidden signal β^* through $Y = X\beta^* + W$.

Roughly speaking, the model exhibits the OGP if for every two nearly optimal solutions σ_1, σ_2, namely solutions satisfying $\sigma_j \approx \min_\sigma F(\sigma, W), j = 1, 2$, either $\langle \sigma_1, \sigma_2 \rangle \approx 0$ or $\langle \sigma_1, \sigma_2 \rangle \approx \|\sigma_1\|_2 \approx \|\sigma_2\|_2$. Namely, every two nearly optimal solutions are either nearly orthogonal or nearly identical. The precise form of the OGP varies from problem to problem.

In this talk several theorems were presented which demonstrated that the onset of the OGP nearly coincides with the (apparent) onset of algorithmic hardness, and furthermore, the presence of the OGP implies a provable failure of a certain class of so-called local algorithms. Specific results concerned the problem of finding a largest cut of a random sparse hypergraph (joint work with Chen, Panchenko and Rahman [2]), the problem of finding a largest submatrix of a random matrix (joint work with Li [5]), and the sparse regression problem discussed above (joint work with Zadik [6],[7]).

References

[1] D. Achlioptas and Y.Peres, *The threshold for random k-SAT is* $2^k ln2 - O(k)$, Proc. 45th Ann. Symposium on the Theory of Computing (2003).

[2] W.-K. Chen, D. Gamarnik, D. Panchenko and M. Rahman, *Suboptimality of local algorithms for a class of max-cut problems.* arXiv preprint arXiv:1707.05386 (2017).

[3] A. Coja-Oghlan, *A better algorithm for random k-sat*, SIAM Journal on Computing **39** (2010), no. 7, 2823-2864.

[4] A. Coja-Oghlan, *The asymptotic k-sat threshold*, Proceedings of the 46th Annual ACM Symposium on Theory of Computing, ACM, 2014, pp. 804-813.

[5] D. Gamarnik and Q. Li, *Finding a large submatrix of a gaussian random matrix*, arXiv preprint arXiv:1602.08529 (2016).

[6] D. Gamarnik and I. Zadik, *High-dimensional regression with binary coefficients. Estimating squared error and a phase transition*, Proceedings of Conference on Learning Theory, 2017.

[7] D. Gamarnik and I. Zadik, *High-dimensional linear regression. Algorithmic barriers and a local search algorithm*, Submitted (2017).

[8] R. M. Karp, *The probabilistic analysis of some combinatorial search algorithms*, Algorithms and complexity: New directions and recent results **1** (1976), 1-19.

The scaling limit of critical random graphs having i.i.d. power-law degrees

Christina Goldschmidt
(joint work with Guillaume Conchon-Kerjan)

This abstract is based on the paper in preparation [4]. Let $n \geq 1$ and fix a degree sequence $\mathbf{d} = (d_1, d_2, \ldots, d_n)$ such that $d_1, d_2, \ldots, d_n \geq 1$ and $\sum_{i=1}^n d_i$ is even. Assume that the set $\mathcal{G}(\mathbf{d})$ of graphs with vertex-set $[n] := \{1, 2, \ldots, n\}$ and such that vertex i has degree d_i is non-empty. Let G_n be an element of $\mathcal{G}(\mathbf{d})$ chosen uniformly at random. The standard method of generating such a graph is via the *configuration model* (see van der Hofstad [9] for a detailed description and precise references for the general results quoted below). Assign vertex i a number d_i of half-edges and then pair the half-edges uniformly at random. In general, this generates a multigraph M_n (i.e. a graph with self-loops and multiple edges), but if we condition it to be simple, it has the same law as G_n. An important feature of the configuration model is that we may generate the pairing edge-by-edge in an order that is convenient for analysis.

We study this model in the setting where the degrees are random variables $D_1, D_2, \ldots, D_n$ which are independent and identically distributed and such that

(1) $\mathbb{E}[D_1^2] = 2\mathbb{E}[D_1]$;
(2) $k^{\alpha+2}\mathbb{P}(D_1 = k) \to c$ as $k \to \infty$, for some constant $c > 0$ and $\alpha \in (1, 2)$.

If $\sum_{i=1}^n D_i$ is odd, simply throw away the last half-edge when we do the pairing. We observe that in this setting $\mathbb{P}(\mathcal{G}(\mathbf{D}) \neq \emptyset)$ converges to a strictly positive constant as $n \to \infty$ so that, for sufficiently large n, conditioning the multigraph to be simple makes sense. Condition (1) says that the graph is critical. Condition (2) entails that the degrees have finite variance but infinite third moment. The *finite* third moment setting has been extensively studied, in particular by Joseph [10] and Riordan [11] (for the component sizes) and Dhara, van der Hofstad, van Leeuwaarden and Sen [5] (for the metric structure). There the scaling limit is (up to constants) the same as that obtained for the critical Erdős–Rényi model in [1] (building on work of Aldous [2], who showed that the component sizes may be encoded as the lengths of excursions above the running minimum of a Brownian motion with parabolic drift).

Let us assume conditions (1) and (2). Write $C_1^n, C_2^n, \ldots$ for the connected components of G_n listed in decreasing order of size, and $|C_1^n|, |C_2^n|, \ldots$ for those sizes. Let $(L_t, t \geq 0)$ be a spectrally positive α-stable Lévy process with Laplace transform

$$\mathbb{E}[\exp(-\lambda L_t)] = \exp(c_\alpha t \lambda^\alpha), \quad \lambda \geq 0, t \geq 0,$$

where $c_\alpha = \frac{c\Gamma(2-\alpha)}{\mu\alpha(\alpha-1)}$. Define a new process $(\tilde{L}_t, t \geq 0)$ via the following martingale change of measure: for any $t \geq 0$ and any bounded measurable test function f,

$$\begin{aligned}&\mathbb{E}[f(\tilde{L}_s, 0 \leq s \leq t)] \\ &= \mathbb{E}\left[\exp\left(\frac{1}{\mu}\int_0^t (L_s - L_t)ds - \frac{c_\alpha t^{\alpha+1}}{(\alpha+1)\mu^{\alpha+1}}\right) f(L_s, 0 \leq s \leq t)\right].\end{aligned}$$

Then the following result is an extension (from M_n to G_n) and reformulation of a theorem due to Joseph [10].

Theorem 1. As $n \to \infty$,

$$n^{-\alpha/(\alpha+1)}(|C_1^n|, |C_2^n|, \ldots) \xrightarrow{d} (\gamma_1, \gamma_2, \ldots)$$

in ℓ^2, where $\gamma_1, \gamma_2, \ldots$ are the ordered lengths of the excursions of $\tilde{L}$ above its running minimum.

Our main result describes the metric space scaling limits of the components themselves. Consider the components $C_1^n, C_2^n, \ldots$ of G_n as measured metric spaces, by endowing each with the graph distance and by assigning mass $n^{-\alpha/(\alpha+1)}$ to each vertex. The limit spaces have an explicit description. First, observe that excursions of the stable Lévy process L may be used to encode a forest of *stable trees*, via the so-called height process H (see Duquesne and Le Gall [7]). We may use the absolute continuity relation between $\tilde{L}$ and L to define a height process $\tilde{H}$ corresponding to $\tilde{L}$: for any $t \geq 0$ and any bounded measurable test function g,

$$\mathbb{E}[g(\tilde{L}_s, \tilde{H}_s, 0 \leq s \leq t)]$$
$$= \mathbb{E}\left[\exp\left(\frac{1}{\mu}\int_0^t (L_s - L_t)ds - \frac{c_\alpha t^{\alpha+1}}{(\alpha+1)\mu^{\alpha+1}}\right) g(L_s, H_s, 0 \leq s \leq t)\right].$$

Let us write

$$R_t = \tilde{L}_t - \inf_{0 \leq s \leq t} \tilde{L}_s, \quad t \geq 0.$$

Write $\varepsilon_1, \varepsilon_2, \ldots$ for the excursions of R above 0, in decreasing order of length, and note that these are in one-to-one correspondence with the excursions $h_1, h_2, \ldots$ of $\tilde{H}$ above 0 (again listed in decreasing order of length). In particular, both ε_i and h_i have length γ_i, for $i \geq 1$. Let $(\mathcal{T}_i, d_i, \mu_i)$ be the measured $\mathbb{R}$-tree encoded in the standard way by h_i, with canonical projection $p_i : [0, \gamma_i] \to \mathcal{T}_i$.

Conditionally on R, consider now a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}_+$ of intensity $\frac{1}{\mu}\mathbb{1}_{\{0 \leq x \leq R_t\}} dtdx$. (Equivalently, this is a rate 1 Poisson point process in the region under the graph of the process R and above the horizontal axis.) Suppose that for $i \geq 1$, a number m_i of points fall under the excursion ε_i. If $m_i \geq 1$, let

$$(s_1^{(i)}, x_1^{(i)}), \ldots, (s_{m_i}^{(i)}, x_{m_i}^{(i)})$$

be the points themselves. Then for $1 \leq k \leq m_i$, let

$$t_k^{(i)} = \inf\left\{t \geq s_k^{(i)} : \varepsilon_i(t) = x_k^{(i)}\right\}.$$

Finally, for $i \geq 1$, if $m_i = 0$, let $\mathcal{C}_i = (\mathcal{T}_i, d_i, \mu_i)$; if $m_i \geq 1$, let $\mathcal{C}_i$ be the measured metric space obtained from $(\mathcal{T}_i, d_i, \mu_i)$ by identifying the pairs of points

$$\left(p_i(s_1^{(i)}), p_i(t_1^{(i)})\right), \ldots, \left(p_i(s_{m_i}^{(i)}), p_i(t_{m_i}^{(i)})\right).$$

For a measured metric space (M, d, μ), write aM as a shorthand for (M, ad, μ). Let $\mathfrak{M}$ be the set of sequences of measure-preserving isometry classes of compact measured metric spaces, endowed with the product Gromov–Hausdorff–Prokhorov topology. Then we have the following scaling limit.

Theorem 2. As $n \to \infty$,

$$n^{-(\alpha-1)/(\alpha+1)}(C_1^n, C_2^n, \ldots) \xrightarrow{d} (\mathcal{C}_1, \mathcal{C}_2, \ldots)$$

in $\mathfrak{M}$.

Closely related results have recently been proved by Dhara, van der Hofstad, van Leeuwaarden and Sen [6] (for component sizes) and by Bhamidi, Dhara, van der Hofstad and Sen [3] (for metric structure). See also the abstract of the talk by Nicolas Broutin in this volume, concerning his forthcoming work with Thomas Duquesne and Minmin Wang.

We observe that the spanning trees $(\mathcal{T}_i, d_i, \mu_i), i \geq 1$ are measure-changed stable trees. Indeed, if $\tilde{\mathbb{e}}$ is an excursion of R conditioned to have length x and $\mathbb{e}$ is an excursion of $(L_t - \inf_{0\leq s\leq t} L_s, t \geq 0)$ conditioned to have length x (and thus encoding a stable tree of size x), we have

$$\mathbb{E}[f(\tilde{\mathbb{e}}(t), 0 \leq t \leq x)] = \frac{\mathbb{E}\left[\exp\left(\frac{1}{\mu}\int_0^x \mathbb{e}(u)du\right) f(\mathbb{e}(t), 0 \leq t \leq x)\right]}{\mathbb{E}\left[\exp\left(\frac{1}{\mu}\int_0^x \mathbb{e}(u)du\right)\right]}.$$

This relation together with detailed knowledge of the stable trees from the literature entail that the distributional properties of the limit spaces $\mathcal{C}_1, \mathcal{C}_2, \ldots$ are particularly tractable; these properties will be further explored in the paper [8] in preparation.

Acknowledgement. The research by CG is supported by EPSRC Fellowship EP/N004833/1.

References

[1] L. Addario-Berry, N. Broutin and C. Goldschmidt, *The continuum limit of critical random graphs*, Probability Theory and Related Fields **152** (2012), 367–406.
[2] D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, Annals of Probability **25** (1997), 812–854.
[3] S. Bhamidi, S. Dhara, R. van der Hofstad and S. Sen, *Universality for critical heavy-tailed network models: metric structure of maximal components*, `arXiv:1703.07145` (2017).
[4] G. Conchon-Kerjan and C. Goldschmidt, *Stable graphs: the metric space scaling limits of critical random graphs with i.i.d. power-law degrees*, in preparation.
[5] S. Dhara, R. van der Hofstad, J. van Leeuwaarden and S. Sen, *Critical window for the configuration model: finite third moment degrees*, `arXiv:1605.02868` (2016).
[6] S. Dhara, R. van der Hofstad, J. van Leeuwaarden and S. Sen, *Heavy-tailed configuration models at criticality*, `arXiv:1612.00650` (2016).
[7] T. Duquesne and J.-F. Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque **281** (2002), vi+147.
[8] C. Goldschmidt, B. Haas and D. Sénizergues, *Stable graphs: distributions and line-breaking constructions*, in preparation.
[9] R. van der Hofstad, *Random graphs and complex networks. Vol I.*, Cambridge University Press, Cambridge (2017), xvi+321.
[10] A. Joseph, *The component sizes of a critical random graph with given degree sequence*, Annals of Applied Probability, **24** (2014), 2560–2594.
[11] O. Riordan, *The phase transition in the configuration model*, Combinatorics, Probability and Computing **21** (2012), 265–299.

Mixing times of random walks on dynamic configuration models

Hakan Güldaş
(joint work with Luca Avena, Remco van der Hofstad, Frank den Hollander)

The mixing time of a random walk, with or without backtracking, on a random graph generated according to the configuration model on n vertices, is known to be of order $\log n$. In this talk we investigate what happens when the random graph becomes *dynamic*, namely, at each unit of time a fraction α_n of the edges is randomly rewired. Under mild conditions on the degree sequence, guaranteeing that the graph is locally tree-like, we show that for every $\varepsilon \in (0,1)$ the ε-mixing time of random walk without backtracking grows like $\sqrt{2\log(1/\varepsilon)/\log(1/(1-\alpha_n))}$ as $n \to \infty$, provided that $\lim_{n\to\infty} \alpha_n (\log n)^2 = \infty$. The latter condition corresponds to a regime of fast enough graph dynamics. Our proof is based on a randomised stopping time argument, in combination with coupling techniques and combinatorial estimates. The stopping time of interest is the first time that the walk moves along an edge that was rewired before, which turns out to be close to a strong stationary time.

Scale-free percolation

Markus Heydenreich
(joint work with Tim Hulshof, Joost Jorritsma)

Many real-world networks, such as WWW, social, financial, neural, or biological networks, exhibit a number of fairly general patterns:

- the length of a smallest path between two vertices is small w.r.t. the system size (*small world*),
- the degrees of vertices exhibit a power law (*a scale-free network*),
- vertices that are geographically close are likely to be connected (*geometric clustering*),
- vertices with high degree are likely to be connected even if far away from each other (*hierarchies*).

It is a challenge to find good mathematical network models that are rich enough to capture these properties but sufficiently simple to be amenable to a rigorous analysis.

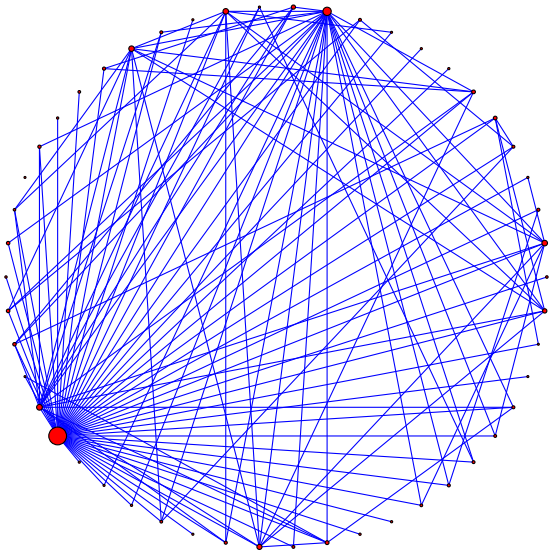
Scale-free percolation, as introduced by Deijfen, van der Hofstad, and Hooghiemstra (2013), is an excellent candidate that meets all of the above criteria. It denotes a percolation model on $\mathbb{Z}^d$, where two lattice points x and y are connected by an edge with probability

$$p_{x,y} = 1 - \exp\left\{-\lambda \frac{W_x W_y}{|x-y|^\alpha}\right\}$$

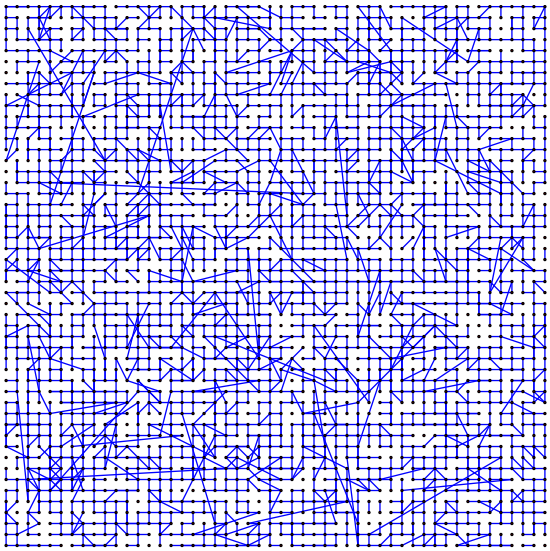
where $\lambda > 0$ is a percolation parameter, W_x and W_y are i.i.d. edge weights with power law distribution

$$P(W_x > w) \approx w^{-(\tau-1)}, \qquad w > 0,$$

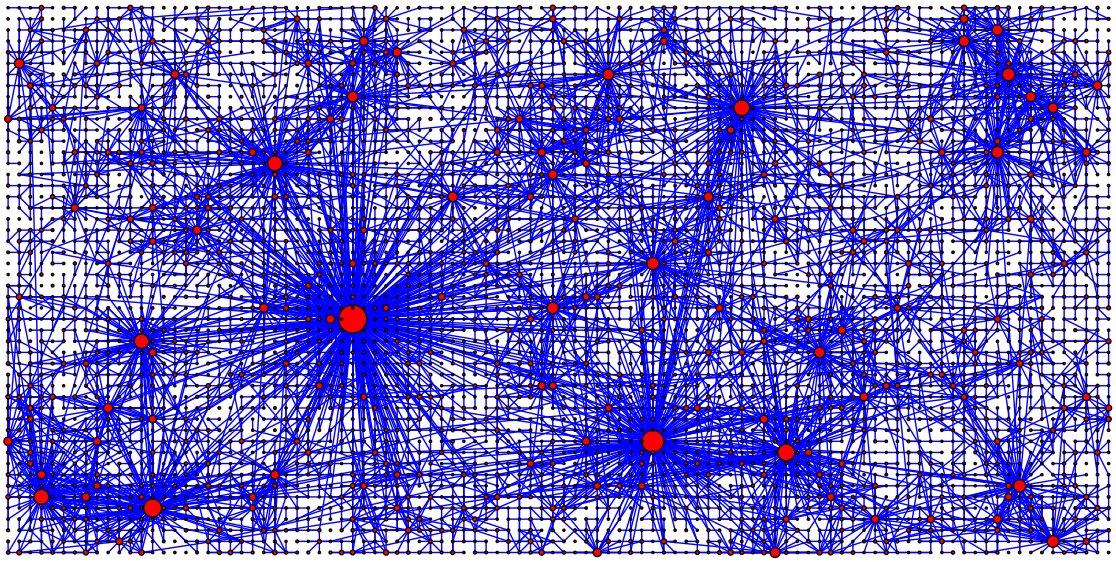
and $\alpha > 0$ denotes the exponent for the long-range connections. It arises as a marriage of the (non-spatial) Norros-Reittu model and the (non scale-free) long-range percolation, as illustrated in Figure 1.



(A) Norros-Reittu random graph, where $p_{x,y} = 1 - \exp\{-\lambda W_x W_y/n\}$ with $\tau = 1.95$.



(B) Long-range percolation, where $p_{x,y} = 1 - \exp\{-\lambda/|x-y|^{\alpha}\}$ with $\alpha = 3.9, \lambda = 0.9$.



(C) Scale-free percolation where $p_{x,y}$ as in Eq. () with $\alpha = 3.9, \tau = 1.95, \lambda = 0.1$.

FIGURE 1. Simulations of the Norros-Reittu random graph (A), long-range percolation (B), and scale-free percolation (C). The size of the vertices is drawn proportionally to their weights.

Bringmann, Keusch, and Lengler (2016) consider a variant of this model in finite domain and continuous space. Indeed, they prove that, for a certain parameter range, the main results from the model on infinite lattice carry over. Further properties have been obtained by Deprez, Hazra, and Wüthrich (2015).

We study structural properties in the regime when there is an infinite connected component. Interestingly, the various parameters introduced above can be condensed into a new parameter $\gamma = \alpha(\tau - 1)/d$, which governs the behaviour of the model. The first result concerns random walk properties on infinite clusters.

Theorem 1. *The following is true for scale-free percolation whenever the parameter λ is chosen such that a (unique) infinite cluster exists.*

(i) *If $\tau \leq 2$ and $\gamma \in (1, 2)$ or $\tau > 2$ and $\alpha \in (d, 2d)$, then the infinite cluster of scale-free percolation is* transient.

(ii) *In dimension $d = 1, 2$, if $\gamma > 2$ and $\alpha > 2d$, then the infinite cluster of scale-free percolation is* recurrent.

For $\gamma < 1$ or $\alpha < d$, every vertex has a.s. infinite degree, hence the notion transience vs. recurrence is not applicable to these cases. It is an open problem to investigate transience vs. recurrence in the regime (ii) of the theorem in dimension $d \geq 3$; this appears to be unresolved even for long-range percolation.

Furthermore, we formalise the notion of "hierarchies" by introducing a new object called *hierarchically clustered tree*, and prove that such trees exist whenever $\gamma \in (1, 2)$ and $\lambda > 0$. For details we refer to Definition 2.5 and Theorem 2.6 in Heydenreich, Hulshof, and Jorritsma (2017).

References

[1] K. Bringmann, R. Keusch, and J. Lengler. Average distance in a general class of scale-free networks with underlying geometry. Preprint arXiv:1602.05712 [cs.SI], 2016.

[2] K. Bringmann, R. Keusch, and J. Lengler. Geometric inhomogeneous random graphs. Preprint arXiv:1511.00576 [cs.DM], 2016.

[3] M. Deijfen, R. van der Hofstad, and G. Hooghiemstra. Scale-free percolation. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):817–838, 2013.

[4] M. Heydenreich, T. Hulshof, and J. Jorritsma. Structures in supercritical scale-free percolation. *Ann. Applied Probab.*, 27(4):2569–2604, 2017.

[5] P. Deprez, R. S. Hazra, and M. V. Wüthrich. Inhomogeneous long-range percolation for real-life network modeling. *Risks*, 3(1):1, 2015.

Slightly Subcritical Hypercube Percolation

Tim Hulshof

(joint work with Asaf Nachmias)

We study bond percolation on the hypercube $\{0, 1\}^m$ in the slightly subcritical regime where $p = p_c(1-\varepsilon_m)$ and $\varepsilon_m = o(1)$ but $\varepsilon_m \gg 2^{-m/3}$ and study the clusters of largest volume and diameter. We establish that with high probability the largest component has cardinality $\Theta\left(\varepsilon_m^{-2} \log(\varepsilon_m^3 2^m)\right)$, that the maximal diameter of all

clusters is $(1+o(1))\varepsilon_m^{-1}\log(\varepsilon_m^3 2^m)$, and that the maximal mixing time of all clusters is $\Theta\left(\varepsilon_m^{-3}\log^2(\varepsilon_m^3 2^m)\right)$.

These results hold in different levels of generality, and in particular, some of the estimates hold for various classes of graphs such as high-dimensional tori, expanders of high degree and girth, products of complete graphs, and infinite lattices in high dimensions.

Edge exchangeable random graphs

Svante Janson

Edge exchangeable random (multi)graphs were introduced by Crane and Dempsey [6, 7]. An equivalent model, using somewhat different formulations, was given by Broderick and Cai [4] and Campbell, Cai and Broderick [5].

The idea is that we have a fixed (labelled) vertex set, and add a sequence of edges (regarded as pairs of vertices). Repetions are allowed, so we construct a multigraph. The sequence of edges is supposed to be exchangeable. By De Finetti's theorem, this is equivalent to the following:

Let V be a finite or infinite set, and let μ be a deterministic or random probability measure on the edges of the complete graph on V.

(1) Given μ, take N i.i.d. edges with distribution μ.
(2) Delete all isolated vertices.

There are some similarities with the (in my opinion more important) vertex exchangeable random graphs (see e.g. [2, 3, 9, 1, 11]) with a discrete type space $\mathbb{N}$, but the two models are quite different. For example, the edge exchangeable graphs have at most one vertex of each type.

Example 1. Let (q_i) be a probability distribution on $\mathbb{N}$. For each edge, just pick its two endpoints independently with this distribution. Thus $\mu(ij) = q_i q_j$. Cf. similar "rank 1" cases of vertex exchangeable graphs, with $W(x,y) = \phi(x)\phi(y)$.

Example 2. Pittel [10] considered a random multigraph process with a fixed vertex set $[n]$ and N edges added one by one, with an edge ij added with probability proportional to $(d_i+\alpha)(d_j+\alpha)$, where d_i is the current degree of i. (Slightly modified for loops.) Here $\alpha > 0$ is a fixed parameter. Equivalently: choose vertices with probability proportional to $d_i+\alpha$. Then join the first two vertices to an edge, then the next two, and so on. Thus, the vertices are chosen according to a Pólya urn process, starting with α balls of each colour (= vertex). The sequence of vertices is exchangeable, and thus so is the sequence of edges. Hence, this is an edge exchangeable random multigraph.

Remarks:

(1) Exchangeability implies that conditioned on the final degree of each vertex, all possible edge sequences have the same probability. Hence, conditioned on the degree sequence, the random multigraph is the multigraph given by the configuration model.

(2) A standard result for Pólya urn processes shows that the vector $(d_i/2N)$ converges to a Dirichlet$(\alpha, \dots, \alpha)$ distribution as $N \to \infty$.
(3) The random sequence of vertices in the construction can be seen as a two-parameter Chinese restaurant process with parameters $(-\alpha, n\alpha)$. A Chinese restaurant process with other parameters yields a similar edge exchangeable random multigraph (on a number of vertices growing to ∞).

Simple graph version. We can merge multiple edges and ignore loops, and thus obtain a random simple graph. This gives an increasing sequence of simple graphs.

Let G_m be the resulting simple graph with m edges.

Example 3. If $P(ij) \sim ((i \vee j)!)^{-4}$, then a.s. $G_m = K_n$ when $m = \binom{n}{2}$, for all large n. Thus $G_n \to$ the graphon 1 a.s. as $n \to \infty$.

Example 4. There exists a distribution μ of edges on $V = \mathbb{N}$ such that a.s. the sequence G_n is dense in the space of graph limits, i.e., for every graph limit (graphon), there exists a subsequence G_{m_i} converging to it. An example of everything? Or of nothing?

See further [8].

References

[1] Christian Borgs, Jennifer T. Chayes, Henry Cohn & Nina Holden. Sparse exchangeable graphs and their limits via graphon processes. Preprint, 2016. `arXiv:1601.07134v1`
[2] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós & Katalin Vesztergombi. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing, *Advances in Math.* **219** (2008), 1801–1851.
[3] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós & Katalin Vesztergombi. Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)* **176** (2012), no. 1, 151–219.
[4] Tamara Broderick and Diana Cai. Edge-exchangeable graphs and sparsity. Preprint, 2016. `arXiv:1603.06898v1`
[5] Trevor Campbell, Diana Cai and Tamara Broderick. Exchangeable trait allocations. Preprint, 2016. `arXiv:1609.09147v1`
[6] Harry Crane and Walter Dempsey. Edge exchangeable models for network data. Preprint, 2016. `arXiv:1603.04571v3`
[7] Harry Crane and Walter Dempsey. Relational exchangeability. Preprint, 2016. `arXiv:1607.06762v1`
[8] Svante Janson, On edge exchangeable random graphs. *J. Statistical Physics*, to appear.
[9] László Lovász, *Large Networks and Graph Limits*. American Mathematical Society, Providence, RI, 2012.
[10] Boris Pittel. On a random graph evolving by degrees. *Adv. Math.* **223** (2010), no. 2, 619–671.
[11] Victor Veitch & Daniel M. Roy. The class of random graphs arising from exchangeable random measures. Preprint, 2015. `arXiv:1512.03099`

Weighted distances in the configuration model

Júlia Komjáthy
(joint work with E. Adriaans, E. Baroni, R. van der Hofstad)

In this talk we study weighted distances in the configuration model with empirical degree distribution that follows a power-law with exponent $\tau \in (2,3)$. We assign independent and identically distributed (i.i.d.) weights to the edges of the graph. We investigate the weighted distance (the length of the shortest weighted path) between two uniformly chosen vertices, called typical distances.

Let us denote the edge-weight distribution by σ and let $d_\sigma(u_n, v_n)$ denote the weighted distance between two uniformly chosen vertices u_n, v_n in the graph on n vertices with respect to i.i.d. σ-distributed edge-weights. Let $F_\sigma^{(-1)}$ denote the generalised inverse of the distribution function of σ. The underlying age-dependent branching process approximating the local neighborhoods of vertices is found to produce infinitely many individuals in finite time – called explosive branching process – for all $\tau \in (2,3)$, precisely when

$$\sum_{k=1}^{\infty} 2F_\sigma^{(-1)}(\exp(-\mathrm{e}^k)) < \infty.$$

When this condition holds, we show that typical distances in the weighted configuration model converge in distribution to a bounded random variable, that is,

$$d_\sigma(u_n, v_n) \xrightarrow{d} V$$

for some proper random variable V, where $\xrightarrow{d}$ denotes convergence in distribution [1]. When the underlying branching process is not explosive, we show that

$$d_\sigma(u_n, v_n) / \sum_{k=1}^{\log\log n/|\log(\tau-2)|} 2F_\sigma^{(-1)}(\exp(-(\tau-2)^{-k})) \xrightarrow{\mathbb{P}} 1,$$

where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability [2]. These results holds for any (not necessary continuous) edge-weight distributions. This sequence tends to infinity with the amount of vertices, and, by choosing an appropriate weight distribution, can be tuned to be any growing function that is $O(\log\log n)$, where n is the number of vertices in the graph.

The proof techniques contain degree-dependent percolation on the configuration model and branching process approximations.

References

[1] E. Baroni, R. van der Hofstad, and J. Komjáthy. *Nonuniversality of weighted random graphs with infinite variance degrees*, Journal of Applied Probability, **54**(1): (2017)146-164,

[2] E. Adriaans, and J. Komjáthy. *Weighted distances in scale-free configuration models*, accepted in Journal of Statistical Physics, Special Issue: Complex Networks. (2018)

Percolation on random planar triangulations: a Boltzmann approach

Grégory Miermont
(joint work with Olivier Bernardi, Nicolas Curien)

Various aspects of Bernoulli percolation on random planar maps have been studied in recent years [1, 2, 11, 7, 12, 6, 8, 9]. Let T_n be a uniformly chosen random rooted triangulation of the 2-sphere with n vertices, where triangulations are considered up to orientation-preserving homeomorphisms that preserve the root. Angel and Schramm have shown [3] that T_n converges locally in distribution to a fundamental limiting object, the Uniform Infinite Planar Triangulation (UIPT) T_∞. We consider Bernoulli site (or bond) percolation on T_∞, where every vertex (or edge) is independently declared open with probability p and closed otherwise, conditionally given the realisation of T_∞. Using a Markovian exploration of T_∞ coined as the peeling process, Angel [1] showed that the Bernoulli site percolation threshold, above which the probability that the open cluster of the root of T_∞ has a positive probability to be infinite, is $p_c^{\mathrm{site}}(T_\infty) = 1/2$. Angel and Curien [2] later identified, using similar methods, the bond percolation threshold as $p_c^{\mathrm{bond}}(H_\infty) = (2\sqrt{3}-1)/11$, for a related half-planar version H_∞ of T_∞. As opposed to the "dynamical" approach of the peeling process, the work [7] uses a "fixed" combinatorial decomposition (inspired by [4]) and known enumeration results on triangulations to study the scaling limit of percolation cluster conditioned on having a large boundary. All the above works focused, in a sense, on the geometry of one percolation interface, hence studied the geometry of the outer boundary of a large percolation cluster. Here we develop a method that is aimed at studying the *full cluster* of the origin in a finite map.

We consider a *critical Boltzmann triangulation*, that is a random finite planar triangulation T chosen with probability proportional to $z_0^{\#\mathrm{triangles}}$, where $z_0 = 432^{-1/4}$ is the maximal value for which this definition makes sense. Under this law, the probability that T has n triangles decays polynomially in n. We then endow T with a Bernoulli site or bond percolation model with parameter $p \in [0,1]$, and consider the origin cluster $\mathfrak{C}(p)$. The cluster $\mathfrak{C}(p)$ is a random planar map which also has a Boltzmann distribution, in the sense that there is a sequence $(q_k)_{k>0}$ of non-negative numbers depending on the parameter p, such that the probability that $\mathfrak{C}(p)$ is equal to any map $\mathfrak{m}$ is proportional to the product over all faces f of $\mathfrak{m}$ of $q_{\deg(f)}$.

By enumerating triangulations with boundaries according to both the boundary length and the number of vertices/edges on the boundary, using a generating function approach in the vein of [13, 5], we find a phase transition at the values $p_c = 1/2$ for site-percolation, and $p_c = (2\sqrt{3}-1)/11$ for bond-percolation. This phase transition manifests itself in at least three ways:

(a) the probability that the cluster $\mathfrak{C}(p)$ has n vertices decays exponentially in n for $p < p_c$ and polynomially for $p \geq p_c$, as $n^{-20/7}$ for $p = p_c$ and $n^{-5/2}$ for $p > p_c$,

(b) the probability that the percolation interface surrounding $\mathfrak{C}(p)$ has length ℓ decays exponentially in ℓ for $p \neq p_c$ and polynomially as $n^{-10/3}$ for $p = p_c$,
(c) the asymptotic form of the sequence $(q_k)_{k>0}$ is different for $p < p_c$, $p = p_c$ and $p > p_c$.

By using (a) and comparison techniques between T and the UIPT T_∞, we give an alternative proof of the fact that the values of p_c correspond to the percolation thresholds $p_c(T_\infty)$ of site and bond percolation on T_∞ as defined above. We also prove the exponential decay of the tail distribution for the size of $\mathfrak{C}(p)$ in the subcritical regime. The result (b) indicates that the critical cluster $\mathfrak{C}(p_c)$ conditioned to have many vertices will have some faces of polynomially large degrees. The result (c) allows us to show that the critical cluster $\mathfrak{C}(p_c)$ is a *non-regular critical Boltzmann map* in the sense of Le Gall and Miermont [10]. It strongly suggests that the rescaled critical percolation cluster conditioned to have n vertices converges in law toward the so-called *stable map* of parameter $\frac{7}{6}$. This conjectural limit leads us to make several additional conjectures on the distribution and geometry of $\mathfrak{C}(p_c)$, in relation to recent results of Gorny, Maurel-Segala and Singh [8].

References

[1] O. Angel. Growth and percolation on the uniform infinite planar triangulation. *Geom. Funct. Anal.*, 13(5):935–974, 2003.
[2] O. Angel and N. Curien. Percolations on infinite random maps, half-plane models. *Ann. Inst. H. Poincaré Probab. Statist.*, 51(2):405–431, 2014.
[3] O. Angel and O. Schramm. Uniform infinite planar triangulation. *Comm. Math. Phys.*, 241(2-3):191–213, 2003.
[4] G. Borot, J. Bouttier, and E. Guitter. A recursive approach to the O(N) model on random maps via nested loops. *J. Phys. A: Math. Theor.*, 45, 2012.
[5] M. Bousquet-Mélou and A. Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *J. Combin. Theory Ser. B*, 96(5):623 – 672, 2006.
[6] N. Curien. Peeling random planar maps. *available at https://www.math.u-psud.fr/ curien/.*
[7] N. Curien and I. Kortchemski. Percolation on random triangulations and stable looptrees. *Probab. Theory Related Fields (to appear).*
[8] M. Gorny, E. Maurel-Segala, and A. Singh. The geometry of a critical percolation cluster on the UIPT. *arxiv.1701.01667*, 2017.
[9] E. Gwynne and J. Miller. Convergence of percolation on uniform quadrangulations with boundary to SLE_6 on $\sqrt{8/3}$-Liouville quantum gravity. arxiv:1701.05175
[10] J.-F. Le Gall and G. Miermont. Scaling limits of random planar maps with large faces. *Ann. Probab.*, 39(1):1–69, 2011.
[11] L. Ménard and P. Nolin. Percolation on uniform infinite planar maps. *Electron. J. Probab.*, 19(78):1–27, 02 2014.
[12] L. Richier. Universal aspects of critical percolation on random half-planar maps. *Electron. J. Probab.*, 20:1–45, 2015.
[13] W. Tutte. A census of slicings. *Canad. J. Math.*, 14:708–722, 1962.

Semidefinite programs on random graphs

Andrea Montanari
(joint work with Zhou Fan)

Given a graph $G = (V, E)$, $|V| = n$, the minimum bisection problem requires to partition its vertex set in two subset of equal cardinality as to minimize the number of edges across the partition. Up to rescalings, this is equivalent to solving the following integer programming problem

$$\begin{aligned} &\text{maximize } \langle \boldsymbol{\sigma}, \boldsymbol{A}_G \boldsymbol{\sigma} \rangle \,, \\ &\text{subject to } \boldsymbol{\sigma} \in \{+1, -1\}^n \quad \langle \boldsymbol{\sigma}, \mathbf{1} \rangle = 0 \,, \end{aligned}$$

where $\boldsymbol{A}_G$ is the adjacency matrix of G. Denote by $\mathsf{OPT}(G)$ the value of this problem. It was proven in [1] that, if G is an Erdös-Renyi random graphs with average degree d, $G \sim \mathcal{G}(n, d/n)$, then with high probability $\mathsf{OPT}(G)/n = 2\mathsf{P}_*\sqrt{d} + o_n(1) + o(\sqrt{d})$, where $\mathsf{P}_* \approx 0.763166726566547$ is the ground state energy of the Sherrington-Kirkpatrick model, which can be computed using Parisi's formula (the evaluation reported here is due to Manuel J. Schmidt).

What are polynomial-time computable upper bounds on $\mathsf{OPT}(G)$? Namely would like a quantity $\mathsf{UB}(G)$ that can be computed in polynomial time and such that: (i) $\mathsf{OPT}(G) \le \mathsf{UB}(G)$ for any graph; (ii) $\mathsf{OPT}(G) \approx \mathsf{UB}(G)$ for random graphs. A classical such upper bound is obtained by solving the following semidefinite program (with $\boldsymbol{A}_G^{\rm cen} = \boldsymbol{A}_G - (d/n)\mathbf{1}\mathbf{1}^{\mathsf{T}}$)

$$\begin{aligned} &\text{maximize } \langle \boldsymbol{A}_G^{\rm cen}, \boldsymbol{X} \rangle \,, \\ &\text{subject to } \boldsymbol{X} \succeq 0, \quad X_{ii} = 1 \; \forall i \,. \end{aligned}$$

In [2, 3], we prove several the value of this program, denoted by $\mathsf{SDP}(G)$. In particular [3], if $G \sim \mathcal{G}(n, d/n)$, then, with high probability, for $d \ge 1$,

$$2\sqrt{d}\left(1 - \frac{1}{d+1}\right) - o_n(1) \le \frac{1}{n}\mathsf{SDP}(G) \le 2\sqrt{d}\left(1 - \frac{1}{2d}\right) + o_n(1) \,.$$

References

[1] A. Dembo, A. Montanari, and S. Sen, *Extremal cuts of sparse random graphs,* The Annals of Probability 45.2 (2017): 1190-1217.

[2] A. Montanari and S. Sen, *Semidefinite programs on sparse random graphs and their application to community detection,* Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. ACM, 2016

[3] Z. Fan and A. Montanari, *How well do local algorithms solve semidefinite programs?* Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing. ACM, 2017

Information Flow on Networks

Elchanan Mossel

We started by reviewing the process of information flow on trees, see e.g [3, 1], then we discussed a new variant of information flow on DAGs studied together with Anuran Makur and Yury Polyanskiy. Interestingly in information flow on DAGs, it suffices for the network to grow at logarithmic rate to remember the bit. We discussed various bounds and phase transitions. In the second part of the talk, based on [2] we introduced information flow processes which model data that is generated in a hierarchal fashion. For such models we can prove separation between deep and shallow algorithms appropriately defined.

References

[1] E. Mossel. Survey: Information flow on trees. In J. Nestril and P. Winkler, editors, *Graphs, Morphisms and Statistical Physics. DIMACS series in discrete mathematics and theoretical computer science*, pages 155–170. 2004.
[2] E. Mossel. Deep learning and hierarchal generative models, 2016. arXiv preprint arXiv:1612.09057.
[3] E. Mossel and Y. Peres. Information flow on trees. *Ann. Appl. Probab.*, 13(3):817–844, 2003.

Branching Brownian motion with decay of mass and the non-local Fisher-KPP equation

Sarah Penington
(joint work with Louigi Addario-Berry, Julien Berestycki)

Branching Brownian motion with decay of mass is a model of competition for resources in a spatially-structured population, introduced in [1]. The model is based on a one-dimensional branching Brownian motion (BBM). We assign a mass to each particle in the BBM; the mass of a particle decays at rate proportional to the total mass of particles in a window of radius μ centred at the location of the particle, for some fixed $\mu > 0$. More precisely, we let $N(t)$ denote the number of particles in the BBM at time t and let $(X_i(t), i \leq N(t))$ denote the locations of the particles. For $t \geq 0$, $x \in \mathbb{R}$, let

$$\zeta(t,x) = \frac{1}{2\mu} \sum_{\{i:|X_i(t)-x|\in(0,\mu)\}} M_i(t).$$

Then for $i \leq N(t)$, let

$$M_i(t) = \exp\left(-\int_0^t \zeta(s, X_{i,t}(s))ds\right),$$

where $X_{i,t}(s)$ is the location of the ancestor of $X_i(t)$ at time s.

In joint work with Louigi Addario-Berry [1], we showed that in a weak sense, at large times t, the front location in this model is $\Theta(t^{1/3})$ behind the location of the rightmost particle in the BBM. For $m > 0$, we let $d(t,m) = \inf\{x > 0 : \zeta(t,x) < m\}$ and $D(t,m) = \sup\{x : \zeta(t,x) > m\}$. Then using results on consistent

maximal displacement for branching Brownian motion [3, 4], we showed in [1] that for $m \in (0,1)$, almost surely,

$$\limsup_{t\to\infty} \frac{\sqrt{2}t - d(t,m)}{t^{1/3}} \geq c^* \quad \text{and} \quad \liminf_{t\to\infty} \frac{\sqrt{2}t - D(t,m)}{t^{1/3}} \leq c^*,$$

where $c^* = 3^{4/3}\pi^{2/3}/2^{7/6}$.

In recent joint work with Louigi Addario-Berry and Julien Berestycki [2], we showed that at large times, the local mass density on a bounded time interval is well approximated by a solution of the non-local Fisher-KPP equation

$$\tfrac{\partial u}{\partial t} = \tfrac{1}{2}\Delta u + u(1 - \phi * u),$$

where $\phi(y) = (2\mu)^{-1}1_{|y|\leq\mu}$. This allowed us to show that for μ sufficiently small, the local mass density in the BBM with decay of mass converges to one behind the front. The behaviour of the local mass density behind the front is an open question for large μ.

References

[1] L. Addario-Berry, S. Penington. *The front location in branching Brownian motion with decay of mass*, Annals of Probability **45(6A)** (2017), 3752–3794.

[2] L. Addario-Berry, J. Berestycki, S. Penington. *Branching Brownian motion with decay of mass and the non-local Fisher-KPP equation*, arXiv preprint arXiv:1712.08098 (2017).

[3] B. Jaffuel. *The critical barrier for the survival of branching random walk with absorption*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **48(4)** (2012), 989–1009.

[4] M. Roberts. *Fine asymptotics for the consistent maximal displacement of branching Brownian motion*, Electron. J. Probab. **20(28)** (2015), 1–26.

Mixing time and cutoff for random walk on random directed graphs

Justin Salez

(joint work with Charles Bordenave, Pietro Caputo)

A finite ergodic Markov chain exhibits cutoff if its distance to equilibrium remains close to 1 over a certain number of iterations and then abruptly drops to 0 on a much shorter time scale. Originally discovered in the context of card shuffling [1], this remarkable phenomenon is now rigorously established for many reversible chains. In [2] we consider the non-reversible case of random walks on sparse directed graphs, for which even the equilibrium measure is far from being understood. We work under the configuration model, allowing both the in-degrees and the out-degrees to be freely specified. We establish the cutoff phenomenon, determine its precise window and prove that the cutoff profile approaches a universal shape. We also provide a detailed description of the equilibrium measure.

References

[1] D. Aldous and P. Diaconis. *Shuffling cards and stopping times*, American Mathematical Monthly (1986), 333–348.
[2] C. Bordenave, P. Caputo and J. Salez. *Random walk on sparse random digraphs*, Probability Theory and Related Fields, to appear.

Optimal subgraph structures

Clara Stegehuis
(joint work with Remco van der Hofstad and Johan S. H. van Leeuwaarden)

Subgraphs reveal information about the geometry and functionalities of complex networks. We count the number of times a small connected graph occurs as an induced subgraph (graphlet counting) in an inhomogeneous random graph or an erased configuration model with power-law degrees with infinite variance. We define an optimization problem that finds for any subgraph the most likely degrees of the nodes that together span the subgraph. We find that every subgraph occurs typically between vertices with specific degree ranges. Furthermore, these degree ranges only take on four different orders of magnitude. In this way, we can count and characterize all subgraphs.

References

[1] R. van der Hofstad, J.S.H. van Leeuwaarden, C. Stegehuis, *Optimal subgraph structures in scale-free configuration models*, arXiv:1709.03466 (2017).

No exceptional words for Bernoulli percolation

Augusto Teixeira
(joint work with Pierre Nolin, Vincent Tassion)

Benjamini and Kesten in [5] introduced in 1995 the problem of embedding infinite binary sequences into a Bernoulli percolation configuration, known as *percolation of words*. We give a positive answer to their Open Problem 2: almost surely, all words are seen for site percolation on $\mathbb{Z}^3$ with parameter $p = \frac{1}{2}$. We also extend this result in various directions.

References

[1] Benjamini, I. Kesten, H., *Percolation of arbitrary words in* $\{0,1\}^{\mathbf{N}}$, Ann. Prob. **23** (1995), 1024–1060.

Resource sharing with logarithmic weights

Amandine Véber
(joint work with Philippe Robert)

We shall consider a communication network in which J clients share a single server. If client j has x jobs pending, it receives a fraction of the capacity of the server which is proportional to $\ln(1+x)$. In contrast with the case of a linear allocation of resources (i.e., proportional to x), when the total number of jobs tends to infinity several timescales interact in a fine way to shape the asymptotic behaviour of the system. In particular, the numbers of jobs corresponding to each client may then evolve on different temporal scales and have very different orders of magnitude.

Extending this result to a different graph of interference between the clients (i.e., star-shaped instead of complete) reveals how dependent on the interaction network the stability of the global system is.

References

[1] P. Robert and A. Véber, *A stochastic analysis of resource sharing with logarithmic weights*, Ann. Applied Probab. **25** (2015), 2626–2670.
[2] P. Robert and A. Véber, *A scaling analysis of a star network with logarithmic weights*, Preprint (2017).

The phase transition in bounded-size Achlioptas processes

Lutz Warnke
(joint work with Oliver Riordan)

In this talks we discuss the percolation phase transition in Achlioptas processes, which have become a key example for random graph processes with dependencies between the edges (see, e.g., [1, 6, 14, 9, 15]). Starting with an empty graph on n vertices, in each step two potential edges are chosen uniformly at random. One of these two edges is then added to the evolving graph according to some rule, where the choice may only depend on the sizes of the components containing the four endvertices. For the widely studied class of bounded-size rules (where all component sizes larger than some constant K are treated the same), the location and existence of the percolation phase transition is nowadays well-understood. However, despite many partial results during the last decade (see, e.g., [5, 21, 11, 13, 4, 15, 3, 2, 16, 8, 17, 18]), our understanding of the finer details of the phase transition has remained incomplete, in particular concerning the size of the largest component.

Our main results (from `arXiv:1704.08714` and `arXiv:1706.00283`) resolve the finite-size scaling behaviour of percolation in all bounded-size Achlioptas processes. We show that for any such rule the phase transition is qualitatively the same as that of the classical Erdős–Rényi random graph process in a very precise sense: the width of the 'critical window' (or 'scaling window') is the same, and so is the asymptotic behaviour of the size of the largest component above and below this

window, as well as the tail behaviour of the component size distribution throughout the phase transition. In particular, when $\varepsilon = \varepsilon(n) \to 0$ as $n \to \infty$ but $\varepsilon^3 n \to \infty$, we show that, with probability tending to 1 as $n \to \infty$, the size of the largest component after i steps satisfies

$$L_1(i) \sim \begin{cases} C\varepsilon^{-2}\log(\varepsilon^3 n) & \text{if } i = t_c n - \varepsilon n, \\ c\varepsilon n & \text{if } i = t_c n + \varepsilon n, \end{cases}$$

where $t_c, C, c > 0$ are rule-dependent constants (in the Erdős–Rényi case we have $t_c = C = 1/2$ and $c = 4$). These and our related results for the component size distribution settle a number of conjectures and open problems from [21, 20, 11, 7, 13, 3, 12, 8]. In the language of mathematical physics, they establish that all bounded-size Achlioptas processes fall in the same 'universality class' (we do not expect this to be true for general Achlioptas processes). Such strong results (which fully identify the phase transition of the largest component and the critical window) are known for very few random graph models.

Our proof deals with the edge–dependencies present in bounded-size Achlioptas processes via a mixture of combinatorial multi-round exposure arguments, the differential equation method, PDE theory, and coupling arguments. This eventually enables us to analyze the phase transition via branching process arguments, see [19, 10] for the details.

References

[1] D. Achlioptas, R.M. D'Souza, and J. Spencer. Explosive percolation in random networks. *Science* **323** (2009), 1453–1455.

[2] S. Bhamidi, A. Budhiraja, and X. Wang. The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs. *Probab. Theory Related Fields* **160** (2014), 733–796.

[3] S. Bhamidi, A. Budhiraja, and X. Wang. Bounded-size rules: The barely subcritical regime. *Combin. Probab. Comput.* **23** (2014), 505–538.

[4] S. Bhamidi, A. Budhiraja, and X. Wang. Aggregation models with limited choice and the multiplicative coalescent. *Random Struct. Alg.* **46** (2015), 55–116.

[5] T. Bohman and D. Kravitz. Creating a giant component. *Combin. Probab. Comput.* **15** (2006), 489–511.

[6] T. Bohman. Emergence of connectivity in networks. *Science* **323** (2009), 1438–1439.

[7] C. Borgs and J. Spencer. Personal communication, EURANDOM workshop Probability and Graphs (2014).

[8] M. Drmota, M. Kang, and K. Panagiotou. Pursuing the Giant in Random Graph Processes. Preprint (2013).

[9] S. Janson. Networking – Smoothly does it. *Science* **333** (2011), 298–299.

[10] S. Janson, O. Riordan, and L. Warnke. Sesqui-type branching processes. Preprint (2017). `arXiv:1706.00283`

[11] S. Janson and J. Spencer. Phase transitions for modified Erdős-Rényi processes. *Ark. Math.* **50** (2012), 305–329.

[12] M. Kang, W. Perkins, and J. Spencer. Personal communication (2012).

[13] M. Kang, W. Perkins, and J. Spencer. The Bohman–Frieze process near criticality. *Random Struct. Alg.* **43** (2013), 221–250.

[14] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science* **333** (2011), 322–324.

[15] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Ann. Appl. Probab.* **22** (2012), 1450–1464.
[16] O. Riordan and L. Warnke. Achlioptas processes are not always self-averaging. *Physical Review E* **86** (2012), 011129.
[17] O. Riordan and L. Warnke. The evolution of subcritical Achlioptas processes. *Rand. Struct. & Algor.* **47** (2015), 174–203.
[18] O. Riordan and L. Warnke. Convergence of Achlioptas processes via differential equations with unique solutions. *Combin. Probab. Comput.* **25** (2016), 154–171.
[19] O. Riordan and L. Warnke. The phase transition in bounded-size Achlioptas processes. Preprint (2017). `arXiv:1704.08714`
[20] J. Spencer. Potpourri. *J. Comb.* **1** (2010), 237–264.
[21] J. Spencer and N.C. Wormald. Birth control for giants. *Combinatorica* **27** (2007), 587–628.

Phase Transitions in High Dimensional Linear Regression

Ilias Zadik
(joint work with David Gamarnik)

In this talk we consider the sparse high dimensional linear regression problem $Y = X\beta^* + W$ under the setting where $X \in \mathbb{R}^{n\times p}, W \in \mathbb{R}^n$ have independent standard Gaussian entries and β^* is known to be k-sparse. The goal is recovering the support of the vector $\beta^* \in \mathbb{R}^p$ from observed X and Y with high probability (w.h.p.) with respect to the randomness coming from X, W. Note that n corresponds to the sample size and p to the number of features.

The vast literature on the topic implies the existence, under certain assumptions, of an intriguing asymptotic gap between the best known necessary sample size n^* for information-theoretically recovering the support of the vector β^* and the best known sufficient sample size n_{alg} for efficiently recovering the vector β^* (which is for example the sample size that techniques such as LASSO or Orthogonal Matching Pursuit need to provably work in this setting). This gap naturally proposes two questions. Supposing the sample size n satisfies $n^* < n < n_{\text{alg}}$, does there exist enough information to recover β^* and if yes, can we construct also an efficient algorithm that recovers it?

We will present in this talk several new results studying these questions. Firstly we will present an all-or-nothing type of result answering completely the first question. Secondly we will present a result showing that for sample size satisfying $n^* < n < cn_{\text{alg}}$, for some $c > 0$ the problem satisfies a property called Overlap Gap Property, which originates in statistical physics and is known to provide evidence of algorithmic difficulty. Finally we will show that if $n > Cn_{\text{alg}}$ for some $C > 0$ the problem does not satisfy this Overlap Gap Property and furthermore the "local" structure of the problem is smooth enough so that even a very simple algorithm based on local search can recover the vector β^*.

References

[1] David Gamarnik and Ilias Zadik. *High dimensional linear regression with binary coefficients: Mean squared error and a phase transition.* Conference on Learning Theory (COLT), 2017.
[2] David Gamarnik and Ilias Zadik. *Sparse High-Dimensional Regression. Algorithmic Barriers and a Local Search Algorithm* https://arxiv.org/abs/1711.04952

Reporter: Matteo Sfragara