Threshold functions for small subgraphs: an analytic approach

Gwendal Collet [1,2]

Discrete Mathematics & Geometry, T.U. Wien
Wien, Austria

Élie de Panafieu [3,2]

MathDyn, Nokia Bell Labs
France

Danièle Gardy [4,2]

David Lab., University of Versailles Saint-Quentin
Versailles, France

Bernhard Gittenberger [5,2]

Discrete Mathematics & Geometry, T.U. Wien
Wien, Austria

Vlady Ravelomanana [6,2]

IRIF, University Paris 7
Paris, France

Abstract

We revisit the problem of counting the number of copies of a fixed graph in a random graph or multigraph, including the case of constrained degrees. Our approach relies heavily on analytic combinatorics and on the notion of *patchwork* to describe the possible overlapping of copies.

Keywords: random graphs, subgraphs, analytic combinatorics, generating functions.

1 Introduction

Since the introduction of the random graph models $G(n, m)$ and $G(n, p)$ by Erdős and Rényi [8] in 1960, one of the most studied parameters is the number X_F of subgraphs isomorphic to a given graph F. By the asymptotic equivalence between $G(n, p)$ and $G(n, m)$, results from one model can be rigorously translated into the other one. Erdős and Rényi derived the threshold for $\{X_F > 0\}$ when F is a *strictly balanced* graph (see definition next page), and Bollobás [3] generalized their result to any graph F. Ruciński [15] proved that X_F is asymptotically normal beyond the threshold, and follows a Poisson law at the threshold iff F is strictly balanced. Then Janson, Oleszkiewicz and Ruciński [12] developed a moment-based method for estimating $\mathbb{P}(X_F \geq (1+\varepsilon)\mathbb{E}(X_F))$. The notion of *strongly balanced graphs*, introduced by Ruciński and Vince in [16], plays a key role in obtaining the results mentioned above.

Recently, there has been an increasing interest in the study of constrained random graphs, such as given degree sequences or regular graphs; the number of given subgraphs in such structures has been also studied. E.g., Wormald [18] proved that the number of short cycles in these structures asymptotically follows a Poisson distribution; using a multi-dimensional saddle-point approach,

[1] Partial support by the FWF (grant SFB F50-02).
[2] Email: `gwendal.collet@tuwien.ac.at`, `elie.de_panafieu@nokia.com`, `daniele.gardy@uvsq.fr`, `gittenberger@dmg.tuwien.ac.at`, `vlad@irif.fr`
[3] Partial support by the FWF (grant F5004), Amadeus program and PEPS *HYDrATA*.
[4] Partial support by the projects Amadeus 33697ZK (2015–16), PICS *ACCA* (2017–19), and ANR-MOST *MetaConc* (2015–19).
[5] Partial support by the FWF (grant SFB F50-03) and the ÖAD grant Amadée F01/2015.
[6] Partial support by the projects Amadeus 33697ZK (2015–16), PICS *ACCA* (2017–19), and *Combinatorics in Paris* (2014–17).

McKay [13] studied the structure of a random graph with given degree sequence, including the probability of a given subgraph or induced subgraph.

Our goal is to revisit (part of) these results through analytic combinatorics and extensive use of generating functions (g.f.). Ours is not the first paper that approaches graph problems with these tools. Early such work was by McKay and Wormald (see, e.g., [14] for the enumeration of graphs with a specified degree sequence); an important development was the study of planar graphs by Giménez and Noy [9], followed by several papers in the same direction; see also a recent paper by Drmota, Ramos and Rué [7] about the limiting distribution of the number of copies of a subgraph in subcritical graphs.

In the rest of this section, we give formal definitions of our model and the objects we are interested in. Then we address the problem of evaluating the number of subgraphs in Section 2; finally some of those results are extended to graphs and multigraphs with degree constraints in Section 3. Due to space constraints, the proofs are in the ArXiV paper of the same name [5].

Model and definitions. Most of the following definitions come from [8] and [3]. A graph G is a pair $(V(G), E(G))$, where $V(G)$ denotes the set of labeled vertices, and $E(G)$ the set of edges. Each edge is a unoriented pair of distinct vertices. An (n,m)-graph is a graph with n vertices, labeled from 1 to n, and m edges. A graph F is a *subgraph* of G if $V(F) \subset V(G)$ and $E(F) \subset E(G)$; we write $F \subset G$. Two graphs F, G are *isomorphic* if there exists a bijection from $V(F)$ to $V(G)$ that induces a bijection between $E(F)$ and $E(G)$. An F*-graph* is a graph isomorphic to F, an F*-subgraph* of G is a subgraph of G that is an F-graph, and $G[F]$ denotes the number of subgraphs of G that are F-graphs. Given a graph family $\mathcal{F}$, an $\mathcal{F}$*-graph* is an F-graph for some $F \in \mathcal{F}$. The *density* $d(G)$ of a graph G is the ratio between its numbers of edges and of vertices. A graph is *strictly balanced* if its density is larger than the density of its strict subgraphs. The *essential density* $d^\star(G)$ of G is the highest density of its subgraphs $d^\star(G) = \max_{H \subset G} d(H)$. To any graph family $\mathcal{F}$, we associate the generating function $F(z,w) = \sum_{n,m \geq 0} \mathcal{F}_{n,m} w^m \frac{z^n}{n!}$, where $\mathcal{F}_{n,m}$ denotes the number of (n,m)-graphs isomorphic to a graph from $\mathcal{F}$.

2 Number of subgraphs in a random graph

Theorem 2.1 *The number of (n,m)-graphs with one distinguished $\mathcal{F}$-subgraph is*

$$n![z^n w^m] F\left(z, \frac{w}{1+w}\right) e^z (1+w)^{\binom{n}{2}} \sim \binom{\binom{n}{2}}{m} F\left(n, \frac{m}{\binom{n}{2}}\right),$$

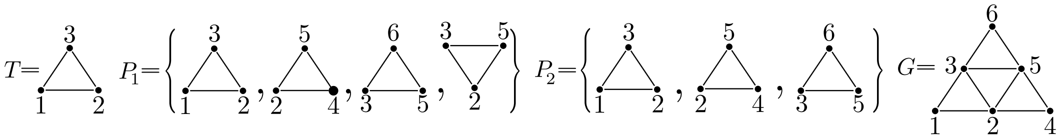

Fig. 1. A graph T and two T-patchworks P_1 and P_2 that reduce to the same graph $G(P_1) = G(P_2) = G$.

where the asymptotics holds when $F(z, w)$ is an entire function, n, $m \to +\infty$ s.t. $m = o(n^2)$, and $F(nz, mw/\binom{n}{2})/F(n, m/\binom{n}{2})$ converges uniformly on any compact set to an analytic function.

Let H denote a densest subgraph of F and $\mathcal{F}$ the family of the H-graphs. Dividing both sides of the expression in Theorem (2.1) by the total number of (n, m)-graphs gives a new proof for the following classical result of [8,3].

Corollary 2.2 *Denote by $\ell^\star$ and $d^\star$ the number of edges and density of a densest subgraph of F, and consider a random (n, m)-graph G with $m = \mathcal{O}(n^\alpha)$ for some fixed $0 < \alpha < 2$. Then*

$$\mathbb{E}(G[F]) = \mathcal{O}(n^{\ell^\star(\alpha - 2 + 1/d^\star)}).$$

Thus, for any $\alpha < 2 - 1/d^\star$, $G[F] = 0$ almost surely.

Given a graph F, an *F-patchwork* P is a set of distinct F-graphs $\{F_1, \ldots, F_t\}$ that might share vertices and edges, and s.t. the pair $(\cup_{i=1}^t V(F_i), \cup_{i=1}^t E(F_i))$ is a graph, denoted by $G(P)$. This notion is illustrated in Figure 1. Let $\mathrm{Patch}_{F,n,m,t}$ denote the number of F-patchworks composed of t F-graphs, and s.t. $G(P)$ is an (n, m)-graph: the g.f. of F-patchworks is $\mathrm{Patch}_F(z, w, u) = \sum_{n,m,t \geq 0} \mathrm{Patch}_{F,n,m,t}\, u^t w^m \frac{z^n}{n!}$.

Theorem 2.3 *The number $\mathrm{SG}^F_{n,m,t}$ of (n, m)-graphs that contain exactly t F-subgraphs is*

$$\mathrm{SG}^F_{n,m,t} = n![z^n w^m u^t]\, \mathrm{Patch}_F\left(z, \frac{w}{1+w}, u - 1\right) e^z (1 + w)^{\binom{n}{2}}.$$

For a general graph F, although we have no explicit expression for the g.f. of F-patchworks, partial information is enough to address some interesting problems. The following theorem was first derived by [3].

Theorem 2.4 *Let F denote a strictly balanced graph of density d, with ℓ edges and $\mathfrak{a}$ automorphisms, and assume $m \sim cn^{2-1/d}$ for some positive constant c. The number of F-subgraphs in a random (n, m)-graph G follows a Poisson limit law of parameter $\lambda = (2c)^\ell/\mathfrak{a}$.*

3 Small subgraphs in graphs with degree constraints

We consider now (n, m, D)-graphs, i.e., (n, m)-graphs where all vertices have their degree in the set D, which contains at least two integers. We restrict our study to the case where m goes to infinity with n in such a way that $\frac{2m}{n}$ has a limit in $]\min(D), \max(D)[$. Since the sum of the degrees is twice the number of edges, if $\frac{2m}{n}$ reaches one of those bounds, the corresponding (n, m, D)-graphs are regular (a case already treated in the literature) while if $\frac{2m}{n}$ is outside the interval, there exist no (n, m, D)-graphs. Finally, to shorten the theorems, we assume $\gcd(d - \min(D) \mid d \in D) = 1$. The g.f. of the set D is $\Delta(x) = \sum_{d \in D} \frac{x^d}{d!}$, and we define $\chi = \chi_{\frac{m}{n}}$ as the unique positive solution (see Note IV.46 of [10]) of $\frac{\chi \Delta'(\chi)}{\Delta(\chi)} = \frac{2m}{n}$. As observed by, e.g., [1,6], *multigraphs* are easier to analyze than graphs when considering degree constraints. A multigraph G is a pair $(V(G), E(G))$ where $V(G)$ denotes the set of labeled vertices, and $E(G)$ the set of labeled oriented edges, each edge is an oriented pair of vertices, and loops and multiple edges are allowed. The definitions on graphs extend naturally to multigraphs. Given a multigraph family $\mathcal{F}$, let $\mathcal{F}_{n,m,(d_0,d_1,\dots)}$ denote the number of (n, m)-multigraphs with d_j vertices of degree j, for all $j \geq 0$, that are isomorphic to some multigraph from $\mathcal{F}$. We associate to the family $\mathcal{F}$ the g.f.

$$F(z, w, (\delta_0, \delta_1, \dots)) = \sum_{n,m,d_0,d_1,\dots} \mathcal{F}_{n,m,(d_0,d_1,\dots)} \left(\prod_{s \geq 0} \delta_s^{d_s} \right) \frac{w^m}{2^m m!} \frac{z^n}{n!}.$$

Theorem 3.1 *The number of (n, m, D)-multigraphs where one $\mathcal{F}$-subgraph is distinguished is*

$$n! 2^m m! [z^n w^m] \sum_{j \geq 0} (2j)! [x^{2j}] F\left(z, w, (\Delta(x), \Delta'(x), \Delta''(x), \dots)\right) e^{z\Delta(x)} \frac{w^j}{2^j j!}.$$

If

$$\frac{F\left(\frac{nz}{\Delta(x\chi)}, \frac{w(x\chi)^2}{2m}, (\Delta(x\chi), \Delta'(x\chi), \dots)\right)}{F\left(\frac{n}{\Delta(\chi)}, \frac{\chi^2}{2m}, (\Delta(\chi), \Delta'(\chi), \dots)\right)}$$

converges uniformly on any compact set to an analytic function, and $\mathrm{MG}_{n,m,D}$ *denotes the total number of (n, m, D)-multigraphs, then the asymptotics of the number of multigraphs with one distinguished subgraph is*

$$\mathrm{MG}_{n,m,D}\, F\left(\frac{n}{\Delta(\chi)}, \frac{\chi^2}{2m}, (\Delta(\chi), \Delta'(\chi), \dots)\right).$$

Corollary 3.2 *Denote by $\ell^\star$ and $d^\star$ the number of edges and density of a densest subgraph of the multigraph F, and consider a random (n, m, D)-multigraph G; then $\mathbb{E}(G[F]) = \mathcal{O}(n^{\ell^\star(1/d^\star - 1)})$.*

As stated earlier, we consider random (n, m, D)-multigraphs with a number m of edges growing linearly with the number n of vertices: χ has a finite positive limit, and the condition of the following theorem is satisfied only for a cycle. In a future extension of this work, we plan to consider the case where $\frac{2m}{n}$ goes to infinity (when D is infinite). In this more general setting, other subgraphs will appear, but the condition should remain as stated here.

Theorem 3.3 *Let F denote a strictly balanced (k, ℓ)-multigraph with $\mathfrak{a}$ automorphisms. Assuming that m goes to infinity with n in such a way that $\frac{1}{\mathfrak{a}} \frac{n^k}{(2m)^\ell} \frac{\chi^{2\ell}}{\Delta(\chi)^k} \prod_{v \in V(F)} \left(\frac{d}{d\chi}\right)^{\deg(v)} \Delta(\chi)$ has a positive limit, denoted by λ, then the number of F-subgraphs in a random (n, m, D)-multigraph follows a Poisson limit law of parameter λ.*

There are $2^m m!$ ways to orient and label the edges of a graph with m edges: each graph matches $2^m m!$ multigraphs. Conversely, consider a multigraph family $\mathcal{F}$, stable by multigraph automorphisms, where each multigraph has m edges, and that contains neither loops nor multiple edges. Then $\mathcal{F}$ can be partitioned into sets of sizes $2^m m!$, each corresponding to a graph. Thus, as proven by [6], counting graphs with degree constraints can be achieved by removing loops and double edges from multigraphs with degree constraints. The following theorem describes the small subgraphs of (n, m, D)-graphs, when $m = \mathcal{O}(n)$. It has been derived in the particular case of regular graphs by [2] and [17], and of graphs with degrees 1 or 2 by [4].

Theorem 3.4 *Consider a random (n, m, D)-graph G that satisfies the conditions stated at the beginning of the section. Then any connected graph that is neither a tree nor a unicycle is asymptotically almost surely not a subgraph of G. Denoting by C_j a cycle of length j: $G[C_3], \ldots, G[C_k]$ are asymptotically independent Poisson random variables of mean $\frac{1}{2j}\left(\frac{1}{2m/n}\frac{\chi^2 \Delta''(\chi)}{\Delta(\chi)}\right)^j$ for each $3 \leq j \leq k$.*

References

[1] Bender, E.A. and E. R. Canfield, *The asymptotic number of labeled graphs with given degree sequences*, JCT A **24 (3)** (1978), 296–307.

[2] Bollobás, B., *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European J. of Combinatorics **1** (1980), 311–316.

[3] Bollobás, B., *Threshold functions for small subgraphs*, Math. Proc. of the Cambridge Phil. Soc. **9** (1981), 197–206.

[4] Broutin, N. and É. de Panafieu, *Limit law for number of components of fixed sizes of graphs with degree one or two*, Arxiv (2014).

[5] Collet, G., É. de Panafieu, D. Gardy, B. Gittenberger and V. Ravelomanana, *Threshold functions for small subgraphs: an analytic approach*, Arxiv (2017).

[6] de Panafieu, É. and L. Ramos, *Graphs with degree constraints*, Proc. of Analytic Algorithmics and Combinatorics (Analco'16) (2016).

[7] Drmota, M., L. Ramos and J. Rué, *Subgraph statistics in subcritical graph classes*, To appear in *Random Structures and Algorithms.*

[8] Erdős, P. and A. Rényi, *On the evolution of random graphs*, Math. Institute of the Hungarian Academy of Sciences, **5** (1960), 17–61.

[9] Giménez, O. and M. Noy, *Asymptotic enumeration and limit laws of planar graphs*, Journal of the AMS, **22 (2)** (2009), 309–329.

[10] Flajolet, P. and R. Sedgewick, "Analytic Combinatorics", Cambridge U. Press, 2009.

[11] Janson, S., *Poisson convergence and Poisson processes with applications to random graphs*, Stochastic Process. Appl. **26** (1987), 1–30.

[12] Janson, S., K. Oleszkiewicz and A. Ruciński, *Upper tails for subgraph counts in random graphs*, Israel J. of Mathematics, **142** (2004), 61–92.

[13] McKay, B. D., *Subgraphs of Dense Random Graphs with Specified Degrees*, Combinatorics, Probability and Computing (2011), **20 (3)**, 413–433.

[14] McKay, B. and N. Wormald, *Asymptotic enumeration by degree sequence of graphs of high degree*, European J. of Combinatorics **11 (6)** (1990), 565–580.

[15] Ruciński, A., *When are small subgraphs of a random graph normally distributed?*, Prob. Th. and Related Fields, **78** (1988), 1–10.

[16] Ruciński, A. and A. Vince, *Strongly Balanced Graphs and Random Graphs*, J. of Graph Theory **10** (1986), 251–264.

[17] Wormald, N. *The asymptotic distribution of short cycles in random regular graphs*, JCT, **B 31** (1981), 168–182.

[18] Wormald, N. C., *Models of random regular graphs*, Surveys in Combinatorics **276**, Cambridge U. Press 239–298 (1999).