A note on log-concave random graphs

Alan Frieze*and Tomasz Tkocz
Department of Mathematical Sciences,
Carnegie Mellon University,
Pittsburgh PA15213,
U.S.A.

June 21, 2018

Abstract

We establish a threshold for the connectivity of certain random graphs whose (dependent) edges are determined by the uniform distributions on generalized Orlicz balls, crucially using their negative correlation properties. We also show the existence of a unique giant component for such random graphs.

1 Introduction

Probabilistic combinatorics is today a thriving field bridging the classical area of probability with modern developments in combinatorics. The theory of random graphs, pioneered by Erdős-Rényi [2], [3] has given us numerous insights, surprises and techniques and has been used to count, to establish structural properties and to analyze algorithms. There are by now several texts [1], [6], [4] that deal exclusively with the subject. The most heavily studied models being $G_{n,m}$ and $G_{n,p}$. Both have vertex set $[n]$ and in the first we choose m random edges and in the second we include each possible edge independently with probability p.

Let X be a random vector in $[0,\infty)^{\binom{n}{2}}$ with a log-concave down-monotone density f, that is (i) $\log f$ is concave and (ii) $f(x) \geqslant f(y)$ if $x \leqslant y$ (coordinate-wise). For $0 < p < 1$, let $G_{X,p}$ be a random graph with vertices $1,\ldots,n$ and edges determined by X: for $1 \leqslant i < j \leqslant n$, $\{i,j\}$ is an edge if and only if $X_{\{i,j\}} \leqslant p$. Such *log-concave random graphs* were introduced by Frieze, Vempala and Vera in [5]. For instance, when X is uniform on $[0,1]^{\binom{n}{2}}$, $G_{X,p}$ is the random graph $G_{n,p}$. The paper [5] introduced a surprising connection between random graphs and convex geometry.

*Research supported in part by NSF Grant DMS1661063

It studied, among other things, the connectivity of $G_{X,p}$ and found a logarithmic gap for the threshold. There is no gap when $G_{X,p}$ is defined by uniform sampling from a "well-behaved" regular simplex[1] and we extend this case to *Generalized Orlicz Balls* GOBs: that is sets of the form $\{x \in \mathbb{R}^d : \sum_{i=1}^d f_i(|x_i|) \leqslant 1\}$ for some nondecreasing lower semicontinuous convex functions $f_1, \ldots, f_d : [0,\infty) \to [0,\infty]$ with $f_i(0) = 0$, which are not identically 0 or $+\infty$ on $(0,\infty)$.
The key property of Orlicz balls is *negative correlation*. We say that a random vector X in $\mathbb{R}^d$ has negatively correlated coordinates if for any disjoint subsets I, J of $\{1,\ldots,d\}$ and nonnegative numbers s_i, t_j, we have

$$\mathbf{P}(\forall i \in I\ |X_i| > s_i, \forall j \in J\ |X_j| > t_j) \leqslant \mathbf{P}(\forall i \in I\ |X_i| > s_i)\mathbf{P}(\forall j \in J\ |X_j| > t_j).$$

It was shown in [7] that this property holds for random vectors uniformly distributed on GOBs (see also [8] for a first such result treating two coordinates and [9] for a simpler proof of the general result).
Notation: Throughout the paper we will let $\sigma_{\min}$ and $\sigma_{\max}$ be defined by

$$\sigma_{\min}^2 = \min_{1\leqslant i<j\leqslant n} \mathbb{E}X_{i,j}^2 \qquad \text{and} \qquad \sigma_{\max}^2 = \max_{1\leqslant i<j\leqslant n} \mathbb{E}X_{i,j}^2.$$

Our result concerning connectivity is the following theorem.

Theorem 1.1 *Let $X = (X_{i,j})_{1\leqslant i<j\leqslant n}$ be a log-concave random vector in $[0,\infty)^{\binom{n}{2}}$ with a down-monotone density and negatively correlated coordinates.*

(a) For every $\delta \in (0,1)$, there are constants c_1 and c_2 dependent only on δ such that for $p < c_1\sigma_{\min}\frac{\log n}{n}$, we have

$$\mathbf{P}(G_{X,p} \text{ has isolated vertices}) > 1 - c_2 n^{-\delta}.$$

(b) For every $\delta \in (0,1)$, there are constants C_1 and C_2 dependent only on δ such that for $p > C_1\sigma_{\max}\frac{\log n}{n}$, we have

$$\mathbf{P}(G_{X,p} \text{ is connected }) > 1 - C_2 n^{-\delta}.$$

We will also discuss the existence of a giant component for smaller values of p.
Notation: Let

$$M = \max_T \sup_{y\in[0,\infty)^{|T|}} \max_{(i,j)\notin T} \mathbb{E}(X_{i,j}^2 | X_T = y), \tag{1}$$

where the first maximum is over all nonempty subsets T of the index set $\{(i,j),\ 1 \leqslant i < j \leqslant n\}$ and we denote $X_T = (X_{i,j})_{(i,j)\in T}$.
For our theorem on the existence of a giant component we need to have $M = O(1)$. For a GOB, $\left\{x \in \mathbb{R}^{\binom{n}{2}} : \sum_{1\leqslant i<j\leqslant n} f_{i,j}(|x_{i,j}|) \leqslant 1\right\}$, this is justified by the following assumption: we let

$$a_{i,j} = \sup\{t > 0 : f_{i,j}(t) \leqslant 1\}.$$

[1] A regular simplex $\left\{x \in \mathbb{R}^d : a \cdot x \leqslant 1\right\}$ for some $a \geqslant 0$ if $a_i/a_j \leqslant K$ for some not too large K.

Now our assumptions on the $f_{i,j}$ imply that the $a_{i,j}$'s are finite. Furthermore, $M \leqslant \max_{i,j} a_{i,j}^2$ and so our assumption here is that $\max_{i,j} a_{i,j}$ is bounded by an absolute constant.

Theorem 1.2 *Let $X = (X_{i,j})_{1\leqslant i<j\leqslant n}$ be a log-concave random vector in $[0,\infty)^{\binom{n}{2}}$ with a down-monotone density. Assume that $M = O(1)$. There are constants c_1 and c_2 such that for every $\beta > 1$, we have*

(i) If $p < \frac{c_1\sigma_{\min}}{n}$, then

$$\mathbf{P}(G_{X,p} \text{ has a component of order } \geqslant \beta \log n) < \frac{12}{n^{\beta-1}}.$$

(ii) If $p > \frac{c_2 M \log\left(\frac{M}{\sigma_{\min}}\right)}{n}$, then

$$\mathbf{P}(G_{X,p} \text{ has a component of order } \in [\beta \log n, n/2]) < \frac{1}{n^{\beta-1}}$$

and

$$\mathbf{P}(G_{X,p} \text{ has a unique giant component of order } > n/2) > 1 - \frac{5\beta}{\log n} - \frac{1}{n^{\beta-1}}.$$

Note that we have dropped the assumption of negative correlation.

2 Connectivity: Proof of Theorem 1.1

Part (b) is part of Theorem 2.1 of [5]. For (a), we adapt the standard second moment argument used for the Erdös-Rényi model. For $1 \leqslant i \leqslant n$, let Y_i be equal to 1 if the vertex i is isolated and 0 otherwise. Let $Y = Y_1 + \ldots + Y_n$ be the number of isolated vertices. We have,

$$\mathbf{P}(G_{X,p} \text{ has isolated vertices}) = \mathbf{P}(Y > 0) \geqslant \frac{(\mathbb{E}Y)^2}{\mathbb{E}Y^2}.$$

Thus, if we show that $\mathbb{E}Y^2 \leqslant (1+\varepsilon)(\mathbb{E}Y)^2$, then $\mathbf{P}(Y > 0) \geqslant 1 - \varepsilon$. Clearly,

$$\mathbb{E}Y^2 = \sum_k \mathbb{E}Y_k^2 + \sum_{k\neq l} \mathbb{E}Y_kY_l = \sum_k \mathbb{E}Y_k + \sum_{k\neq l} \mathbf{P}(Y_k = 1 = Y_l) = \mathbb{E}Y + \sum_{k\neq l} \mathbf{P}(Y_k = 1 = Y_l)$$

and our goal is to show that

$$\mathbb{E}Y \leqslant \frac{\varepsilon}{2}(\mathbb{E}Y)^2 \qquad \text{and} \qquad \sum_{k\neq l} \mathbf{P}(Y_k = 1 = Y_l) \leqslant \left(1 + \frac{\varepsilon}{2}\right)(\mathbb{E}Y)^2.$$

From the negative correlation of coordinates of X as well as an elementary inequality $\mathbf{P}(A) \leqslant \mathbf{P}(A \cap B) + 1 - \mathbf{P}(B)$, we get

$$\begin{aligned}\mathbf{P}(Y_k = 1 = Y_l) &= \mathbf{P}(\forall i \neq k\ X_{ik} > p, X_{il} > p, X_{kl} > p)\\ &\leqslant \mathbf{P}(\forall i \neq k\ X_{ik} > p)\mathbf{P}(\forall i \neq k, l\ X_{il} > p)\\ &\leqslant \mathbf{P}(\forall i \neq k\ X_{ik} > p)\big[\mathbf{P}(\forall i \neq l\ X_{il} > p) + 1 - \mathbf{P}(X_{kl} > p)\big]\\ &= \mathbf{P}(Y_k = 1)\big[\mathbf{P}(Y_l = 1) + \mathbf{P}(X_{kl} \leqslant p)\big].\end{aligned}$$

By Lemma 3.5 from [5], $\mathbf{P}(X_{kl} \leqslant p) \leqslant \frac{p}{\sigma_{\min}}$ (recall that by the Prékopa-Leindler inequality, marginals of log-concave vectors are log-concave; clearly, marginals of down-monotone densities are down-monotone). Therefore,

$$\begin{aligned}\sum_{k\neq l} \mathbf{P}(Y_k = 1 = Y_l) &\leqslant \sum_{k\neq l} \mathbf{P}(Y_k = 1)\mathbf{P}(Y_l = 1) + \sum_{k\neq l} \mathbf{P}(Y_k = 1)\frac{p}{\sigma_{\min}}\\ &\leqslant \left(\sum_k \mathbf{P}(Y_k = 1)\right)^2 + \frac{np}{\sigma_{\min}} \sum_k \mathbf{P}(Y_k = 1)\\ &\leqslant \left(1 + \frac{np}{\sigma_{\min}\mathbb{E}Y}\right)(\mathbb{E}Y)^2 < \left(1 + \frac{c_1 \log n}{\mathbb{E}Y}\right)(\mathbb{E}Y)^2,\end{aligned}$$

so it suffices to take ε such that

$$\varepsilon \geqslant \frac{2}{\mathbb{E}Y} \qquad \text{and} \qquad \varepsilon \geqslant \frac{2c_1 \log n}{\mathbb{E}Y}.$$

By Lemma 3.1 from [5], $\mathbf{P}(Y_k = 1) \geqslant e^{-apn/\sigma_{\min}}$, for some universal constant a (the assumption $p < \frac{1}{4}\sigma_{\min}$ of that lemma is clearly satisfied if $p < c_1\sigma_{\min}\frac{\log n}{n}$), so

$$\mathbb{E}Y = \sum_k \mathbf{P}(Y_k = 1) \geqslant ne^{-apn/\sigma_{\min}} > n^{1-ac_1}.$$

Thus, $\varepsilon = c_2 n^{ac_1 - 1}\log n$ will suffice. □

3 Giant Component: Proof of Theorem 1.2

Lemma 3.1 *Let $X = (X_{i,j})_{1\leqslant i<j\leqslant n}$ be a log-concave random vector in $[0,\infty)^{\binom{n}{2}}$ with a down-monotone density. There are universal constants a and b such that for $S, T \subset \{(i,j),\ 1 \leqslant i < j \leqslant n\}$ and $p > 0$, we have*

$$\mathbf{P}(\forall s \in S\ X_s > p, \forall t \in T\ X_t \leqslant p) \leqslant e^{-ap|S|/M}\left(\frac{bp}{\sigma_{\min}}\right)^{|T|}.$$

Proof Fix disjoint sets $S, T \subset \{(i,j),\ 1 \leqslant i < j \leqslant n\}$ (if they are not disjoint, the probability in question is 0) and $y \in [0,\infty)^{|T|}$. Let f be the density of (X_S, X_T). The conditional density of the vector X_S given $X_T = y$,

$$f_{X_S|X_T}(x|y) = \frac{f(x,y)}{\int f(x',y)\mathrm{d}x'}$$

is down-monotone and log-concave. Therefore, by Lemma 3.1 from [5],

$$\mathbf{P}(\forall s \in S\ X_s > p | X_T = y) \leqslant e^{-ap|S|/M}.$$

We denote the density of X_T by f_{X_T} and get

$$\begin{aligned}
\mathbf{P}(\forall s \in S\ X_s > p, \forall t \in T\ X_t \leqslant p) &= \int_{[0,p]^{|T|}} \mathbf{P}(\forall s \in S\ X_s > p | X_T = y) f_{X_T}(y)\mathrm{d}y \\
&\leqslant \int_{[0,p]^{|T|}} e^{-ap|S|/M} f_{X_T}(y)\mathrm{d}y \\
&= e^{-ap|S|/M}\mathbf{P}(\forall t \in T\ X_t \leqslant p) \\
&\leqslant e^{-ap|S|/M}\left(\frac{bp}{\sigma_{\min}}\right)^{|T|},
\end{aligned}$$

where the final inequality follows directly from Lemma 3.2 of [5]. □

With this lemma in hand, we can prove Theorem 1.2.

Proof Let Z_k be the number of components of order k (that is, on k vertices) in $G_{X,p}$. As for the Erdös-Rényi model, looking at a spanning tree for each component and bounding the corresponding in-out edge probabilities using Lemma 3.1 yields

$$\begin{aligned}
\mathbb{E}Z_k &\leqslant \binom{n}{k} k^{k-2} e^{-apk(n-k)/M}\left(\frac{bp}{\sigma_{\min}}\right)^{k-1} \\
&\leqslant \left(\frac{en}{k}\right)^k k^{k-2} e^{-apk(n-k)/M}\left(\frac{bp}{\sigma_{\min}}\right)^{k-1} \\
&= \frac{\sigma_{\min}}{bp}\frac{1}{k^2}\left[\frac{eb}{\sigma_{\min}} pn e^{-\frac{ap}{M}(n-k)}\right]^k.
\end{aligned}$$

If $p = \frac{M}{a}\frac{c}{n}$, with c being a constant (chosen soon), this becomes

$$\mathbb{E}Z_k \leqslant \frac{e}{A}\frac{1}{c}\frac{n}{k^2}\left[Ace^{-c}e^{ck/n}\right]^k,$$

where we put $A = \frac{eb}{a}\frac{M}{\sigma_{\min}}$.

Case 1. If c is a small constant, say $c \leqslant \frac{1}{eA}$ (equivalently, $p \leqslant \frac{\sigma_{\min}}{e^2 b}\frac{1}{n}$), then we bound $e^{-c}e^{ck/n}$ crudely by 1 and get that

$$\mathbb{E}Z_k \leqslant \frac{e}{A}\frac{1}{c}\frac{n}{k^2}(Ac)^k \leqslant en(Ac)^{k-1} \leqslant e^2 ne^{-k}.$$

Thus,

$$\mathbb{E}\left(\sum_{k\geqslant\beta\log n} Z_k\right) \leqslant e^2 n \cdot \sum_{k\geqslant\beta\log n} e^{-k} \leqslant e^2 ne^{-\beta\log n}\frac{1}{1-e^{-1}} = \frac{e^3}{e-1}\frac{1}{n^{\beta-1}} < \frac{12}{n^{\beta-1}}.$$

By the first moment method, this gives (i).

Case 2. Let c be a large constant, say such that $Ace^{-c/2} \leqslant \frac{1}{e}$ and $Ac \geqslant e^2$, which holds when, say $c \geqslant 4\log A$, provided that A is large enough, which leads to the assumption on p in (ii). Then for $k \leqslant n/2$, we have

$$\mathbb{E}Z_k \leqslant \frac{en}{Ac}(Ace^{-c/2})^k \leqslant ne^{-k-1}.$$

Thus,

$$\mathbb{E}\left(\sum_{\beta\log n\leqslant k\leqslant n/2} Z_k\right) \leqslant ne^{-1}\sum_{k\geqslant\beta\log n} e^{-k} \leqslant \frac{1}{e-1}\frac{1}{n^{\beta-1}} < \frac{1}{n^{\beta-1}}.$$

By the first moment method, this gives the first part of (ii).
To go about the second part and show that there is a giant component, we shall simply count the number of vertices on the small components and show that with high probability, there are strictly less n such vertices. The uniqueness of a giant component plainly follows from the fact that it has more than $n/2$ vertices, so there cannot be more than one such components. Fix $1 \leqslant k \leqslant \beta\log n$ and set $t = ne^{-k-1}$. For any positive integer $l \leqslant et+1$, we have

$$\begin{aligned}\mathbf{P}(Z_k \geqslant et) &\leqslant \mathbf{P}(Z_k(Z_k-1)\dots(Z_k-l+1) \geqslant et(et-1)\dots(et-l+1)) \\ &\leqslant \frac{\mathbb{E}Z_k(Z_k-1)\dots(Z_k-l+1)}{et(et-1)\dots(et-l+1)} \\ &\leqslant \frac{\mathbb{E}Z_k(Z_k-1)\dots(Z_k-l+1)}{(et-l+1)^l}.\end{aligned}$$

As for the upper bound for $\mathbb{E}Z_k$, looking at spanning trees for each l-tuple of distinct components of order k and bounding the corresponding in-out edge probabilities using

Lemma 3.1 yields

$$\begin{aligned}
&\mathbb{E}Z_k(Z_k-1)\dots(Z_k-l+1)\\
&\quad\leqslant \binom{n}{k}\binom{n-k}{k}\dots\binom{n-(m-1)k}{k}(k^{k-2})^l e^{-\frac{ap}{M}kl(n-kl)}\left(\frac{bp}{\sigma_{\min}}\right)^{(k-1)l}\\
&\quad\leqslant \left(\frac{en}{k}\right)^{kl}(k^{k-2})^l e^{-\frac{ap}{M}kl(n-kl)}\left(\frac{bp}{\sigma_{\min}}\right)^{(k-1)l}\\
&\quad= \left(\frac{e}{A}\frac{1}{c}\frac{n}{k^2}\left[Ace^{-c}e^{ckl/n}\right]^k\right)^l.
\end{aligned}$$

Provided that $kl \leqslant n/2$, under our assumption $c \geqslant 4\log A$, this is further upper bounded by $(t/k^2)^l$, which gives

$$\mathbf{P}(Z_k \geqslant ne^{-k}) = \mathbf{P}(Z_k \geqslant et) \leqslant \frac{1}{k^{2l}}\left(\frac{t}{et-l+1}\right)^l.$$

For $k \geqslant \frac{1}{2}\log n$, we choose $l=1$ and get

$$\mathbf{P}(Z_k \geqslant ne^{-k}) \leqslant \frac{1}{e}\frac{1}{(\frac{1}{2}\log n)^2}, \qquad k \geqslant \frac{1}{2}\log n.$$

For $k < \frac{1}{2}\log n$, we have $t = ne^{-k-1} > e^{-1}\sqrt{n}$, so choosing, say $l-1 = \lfloor e^{-1}\sqrt{n}\rfloor$ yields

$$\begin{aligned}
\mathbf{P}(Z_k \geqslant ne^{-k}) \leqslant \left(\frac{t}{et-\lfloor e^{-1}\sqrt{n}\rfloor}\right)^l = \left(\frac{1}{e-\frac{\lfloor e^{-1}\sqrt{n}\rfloor}{t}}\right)^l &\leqslant \left(\frac{1}{e-1}\right)^l\\
&\leqslant \left(\frac{1}{e-1}\right)^{e^{-1}\sqrt{n}}, \qquad k < \frac{1}{2}\log n.
\end{aligned}$$

Combining the last two estimates, the union bound gives that the probability of the event $E = \{\exists k \leqslant \beta\log n,\ Z_k \geqslant ne^{-k}\}$ is at most

$$\frac{4}{e}\frac{(\beta-1/2)\log n+1}{(\log n)^2} + \frac{\frac{1}{2}\log n}{(e-1)^{e^{-1}\sqrt{n}}} < \frac{5\beta}{\log n}$$

(we check that $\frac{\frac{1}{2}\log n}{(e-1)^{e^{-1}\sqrt{n}}} < \frac{2}{\log n}$ and simply bound $\frac{4}{e}\frac{(\beta-1/2)\log n+1}{(\log n)^2} \leqslant \frac{\frac{4}{e}\beta+\frac{2}{e}}{\log n}$). To finish, it remains to check that on E^c, there are few vertices on the small components. On E^c, we have

$$\sum_{k\leqslant\beta\log n} kZ_k \leqslant n\sum_{k\leqslant\beta\log n} ke^{-k} < n\sum_{k=1}^{\infty} ke^{-k} = n\frac{e}{(e-1)^2} < 0.93n.$$

□

Remark 3.2 It was shown in [9] that the negative correlation property holds in fact for random vectors with densities of the form $h(\sum f_i(x_i))$, where $h : [0,\infty) \to [0,\infty)$ is a nonincreasing log-concave function ($h = \mathbf{1}_{[0,1]}$ giving uniform densities on GOBs). For such densities, M is finite and can be bounded as for GOBs in terms of certain parameters depending on the functions f_i and h.

4 Conclusion and Open Questions

We have successfully generalised the results on the regular simplex in [5] to GOBs. The following questions seem most apposite.

Q1 What we prove in Theorem 1.2 does not rule out the possibility that in some range of p there is more than one giant component. Can the proof be tightened to rule this out?

Q2 What is the connectivity or giant component threshold for the intersection of two well-behaved regular simplices?

Q3 What is the connectivity or giant component threshold for the intersection of a *few* regular simplices with independent randomly chosen coefficients?

References

[1] B. Bollobás: Random Graphs, *Academic Press*, 1985.

[2] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* 6 (1959) 290-297.

[3] P. Erdős and A. Rényi: On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* 5 (1960) 17-61.

[4] A.M. Frieze and M. Karoński, An introduction to random graphs, *Cambridge University Press*, 2015.

[5] A.M. Frieze, S. Vempala and J. Vera, Logconcave random graphs. *Electron. J. Combin.* 17 (2010), no. 1, Research Paper 108, 31 pp.

[6] S. Janson, T. Łuczak and A Rucinski: Random Graphs, *Wiley-Interscience*, 2000.

[7] M. Pilipczuk, J.O. Wojtaszczyk, The negative association property for the absolute values of random variables equidistributed on a generalized Orlicz ball. *Positivity* 12 (2008), no. 3, 421–474.

[8] J.O. Wojtaszczyk, The square negative correlation property for generalized Orlicz balls. Geometric aspects of functional analysis, 305–313, *Lecture Notes in Math.*, 1910, Springer, Berlin, 2007.

[9] J.O. Wojtaszczyk, A simpler proof of the negative association property for absolute values of measures tied to generalized Orlicz balls. *Bull. Pol. Acad. Sci. Math.* 57 (2009), no. 1, 41–56.