On matchings in stochastic Kronecker graphs

Justyna Banaszak*

Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Poznań, Poland

just.banaszak@gmail.com

Submitted: Nov 5, 2015; Accepted: Sep 29, 2016; Published: Oct 14, 2016
Mathematics Subject Classifications: 05C80; 05C70; 05C40

Abstract

The stochastic Kronecker graph is a random structure whose vertex set is a hypercube and the probability of an edge depends on the structure of its ends. We prove that when a.a.s. the stochastic Kronecker graph becomes connected it a.a.s. contains a perfect matching.

Let $n \in \mathbb{N}$, $N = 2^n$, and let $0 \leqslant \alpha, \beta, \gamma \leqslant 1$, where $\gamma \leqslant \alpha$ be some constants. Denote by $\mathbf{P}$ a symmetric matrix

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 0 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 0 \end{array} \\ \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \end{array},$$

where 0's and 1's are labels of rows and columns of $\mathbf{P}$. A stochastic Kronecker graph $\mathcal{K}(n, \mathbf{P})$ is a random graph with vertex set $V = \{0,1\}^n$, the set of all binary sequences of length n, where the probability that two vertices $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in V$ are adjacent is given by

$$p_{u,v} = \prod_{i=1}^{n} \mathbf{P}[u_i, v_i].$$

This model was introduced by Leskovec, Chakrabarti, Kleinberg and Faloutsos in [4]. They showed empirically that the desired graph properties of real world networks hold in this model (see [5], [6]).

In [2] Mahdian and Xu showed that the threshold for connectivity of stochastic Kronecker graph is $\beta + \gamma = 1$. They also proved that if α, β, γ are such that $\mathcal{K}(n, \mathbf{P})$ is asymptotically almost surely (a.a.s.) connected, and moreover $\gamma \leqslant \beta \leqslant \alpha$, then the graph has a.a.s. a constant diameter. Horn and Radcliffe [7] studied the emergence of the giant

*Partially supported by grant NCN Maestro 2012/06/A/ST1/00261.

component and verified that a.a.s. it appears in $\mathcal{K}(n, \mathbf{P})$ as soon as $(\alpha+\beta)(\beta+\gamma) > 1$. Kang, Karoński, Koch and Makai [8] showed that the degree distribution of $K(n, \mathbf{P})$ does not obey a power-law and determined the threshold for the existence of some small subgraphs.

In this paper we denote by $d(v, u)$ the Hamming distance between two vertices v and u and by $w(v)$ the weight of a vertex $v = (v_1, \dots, v_n)$ the number of 1's in its label

$$w(v) = \sum_{i=1}^{n} v_i.$$

Note that the expected degree of a vertex v with weight $w = w(v)$ is

$$\mathbb{E}(\deg(v)) = \sum_{i=0}^{w} \binom{w}{i} \alpha^i \beta^{w-i} \sum_{j=0}^{n-w} \binom{n-w}{j} \gamma^j \beta^{n-w-j} = (\alpha+\beta)^w (\beta+\gamma)^{n-w}. \tag{1}$$

We start with stating the connectivity result. As we have already mentioned Mahdian and Xu [2] showed that if $\beta+\gamma > 1$, then a.a.s. $\mathcal{K}(n, \mathbf{P})$ is connected while for $\beta+\gamma < 1$ a.a.s. it contains isolated vertices. Later their result was supplemented (in much larger generality by Radcliffe and Young [9]). From their result we can derive the following observation concerning the connectivity of $\mathcal{K}(n, \mathbf{P})$ at the threshold, i.e. when $\beta + \gamma = 1$.

Theorem 1.

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{K}(n, \mathbf{P}) \text{ is connected}) = \begin{cases} 0 & \text{if } \beta+\gamma = 1,\ \beta \neq 1 \text{ or } \beta = 1,\ \alpha = \gamma = 0 \\ 1 & \text{if } \beta = 1,\ \alpha > 0 \text{ and } \gamma = 0. \end{cases}$$

We shall show that the threshold for the emergence of a perfect matching in $\mathcal{K}(n, \mathbf{P})$ is basically the same as the connectivity threshold. Our main result can be stated as follows.

Theorem 2.

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{K}(n, \mathbf{P}) \text{ contains a perfect matching}) = \begin{cases} 0 & \text{if } \beta+\gamma \leqslant 1 \text{ and } \beta \neq 1 \\ 1 & \text{if } \beta+\gamma > 1 \text{ or } \beta = 1. \end{cases}$$

Proof. Let $\beta + \gamma \leqslant 1$ and $\beta \neq 1$. In the proof of Theorem 1, Radcliffe and Young have shown that a.a.s. $\mathcal{K}(n, \mathbf{P})$ contains an isolated vertex and so a.a.s. it does not contain a perfect matching.

Let $\beta = 1$. Then every vertex $v = (v_1, \dots, v_n) \in V(\mathcal{K}(n, \mathbf{P}))$ is with probability $\beta^n = 1$ connected to $\bar{v} = (1-v_1, \dots, 1-v_n)$. Thus, with probability 1, $\mathcal{K}(n, \mathbf{P})$ contains a perfect matching.

Now let us consider the most interesting case, $\beta + \gamma > 1$. The main idea of our argument is the following. We shall choose a dense bipartite subgraph $\mathcal{H}$ of $\mathcal{K}(n, \mathbf{P})$ and show that it contains a perfect matching by verifying Hall's condition.

To this end for a given odd number t let $H = H(n, t)$ denote a graph with a vertex set $\{0,1\}^n$, and edges between the pairs of vertices of Hamming distance $d(u, v) = t$. Denote by V_1 and V_2 the subsets of vertices of H of odd and even weights respectively. Since t is odd, all the edges in H must have one end in V_1 and the other in V_2, i.e. H is bipartite. Now let $\mathcal{H} = \mathcal{H}(t, n, \mathbf{P})$ be a subgraph of $\mathcal{K}(n, \mathbf{P})$, which contains only the edges between the vertices of Hamming distance t, i.e. $\mathcal{H}$ contains only those edges of $\mathcal{K}(n, \mathbf{P})$ which belong to H. Clearly, $\mathcal{H}$ is bipartite. We shall show that for

$$t = 2\left\lceil \frac{\beta}{2(\beta+\gamma)} n \right\rceil + 1,$$

which is basically the value of t which maximizes the expected number of edges in $\mathcal{H}$, the random bipartite graph $\mathcal{H}$ a.a.s. fulfills Hall's condition, and so a.a.s. it contains a perfect matching.

In order to do that we show first that the underlying bipartite graph H has good expanding properties. Let us first introduce some notation. For two subsets W and U of the vertex set of H, let $e_H(W, U)$ denote the number of edges with one end in W and another in U. Let $\overline{W}$ denote the complement of W in the vertex set of H. By $\mathrm{Vol}(W)$ we denote the sum of vertex degrees of W.

Let us recall that a graph G is edge-transitive, if for any two edges $e^1, e^2 \in E(G)$ there exists a graph automorphism $F : V(G) \to V(G)$, which transforms e^1 into e^2. The following result of Chung [1] (Theorem 7.1) is crucial for our argument.

Theorem 3. *Let G be an edge-transitive graph with diameter D. Then for every $W \subseteq V(G)$, such that $\mathrm{Vol}(W) \leqslant \frac{\mathrm{Vol}(V(G))}{2}$,*

$$\frac{e_G(W, \overline{W})}{\mathrm{Vol}(W)} \geqslant \frac{1}{2D}.$$

In order to apply this result we need to check if H is edge transitive and has small diameter.

Lemma 4. *$H\left(n, 2\left\lceil \frac{\beta}{2(\beta+\gamma)} n \right\rceil + 1\right)$ is edge-transitive and its diameter can be bounded from above by a constant D which depends only on constants β and γ but not on n.*

Proof. Clearly, for $i \in [n]$ the function $\tau_i : \{0,1\}^n \to \{0,1\}^n$ that maps $(v_1, \ldots, v_i, \ldots, v_n)$ to $(v_1, \ldots, 1 - v_i, \ldots, v_n)$ is an automorphism of H. Also, for any permutation $\sigma : [n] \to [n]$, the map $\mathrm{Aut}(\sigma) : \{0,1\}^n \to \{0,1\}^n$ that maps the vertex $(v_1, \ldots, v_i, \ldots, v_n)$ to $(v_{\sigma(1)}, \ldots, v_{\sigma(i)}, \ldots, v_{\sigma(n)})$ is an automorphism of H. We show that the group generated by all automorphisms of the above two kinds acts transitively on edges of H.

Although it is a rather easy observation let us prove it more formally. Let $e^1 = \{u^1, v^1\}$, $e^2 = \{u^2, v^2\}$ be two edges of H. For $i \in \{1, 2\}$, there exist precisely t positions j such that $u^i_j \neq v^i_j$. Let $I_i \subseteq [n]$ be the set of those positions (for $i \in \{1, 2\}$). Let ϕ be a permutation of $[n]$ such that $\phi(I_1) = I_2$ and $\phi^* = \mathrm{Aut}(\phi)$ be the automorphism of H induced by ϕ. This automorphism is uniquely defined for fixed permutation ϕ. Note that the pairs $\{\phi^*(u^1), \phi^*(v^1)\}$ and $\{u^2, v^2\}$ differ on the same positions, i.e. $\phi^*(u^1)_j \neq \phi^*(v^1)_j$ if and only if $u^2_j \neq v^2_j$. Define $\psi : \{0,1\}^n \to \{0,1\}^n$ by putting

$$\psi(x)_j = \begin{cases} x_j & \text{if } \phi^*(u^1)_j = u^2_j \\ 1 - x_j & \text{otherwise.} \end{cases}$$

Clearly $\psi(\phi^*(u^1)) = u^2$. Moreover $\psi(x)_j = x_j$ iff $\phi^*(u^1)_j = u^2_j$ and it happens iff $\phi^*(v^1)_j = v^2_j$. Thus $\psi(\phi^*(v^1)) = v^2$ so $\psi \circ \phi^*$ is an automorphism of H which maps e^1 into e^2. Hence H is edge-transitive.

It remains to find an upper bound for the diameter of H. Let v, v' be two vertices of H such that $d(v, v')$ is even. We show that they are connected by a short path. We split our argument into several cases.

Case 1. $d(v, v') \leqslant \min\{2t, 2n - 2t\}$.

In this case there exists a vertex v'' which is adjacent to both v and v'. Indeed, to find v'' it is enough to change v on $d(v, v')/2$ positions on which v and v' differ and $t - d(v, v')/2$ positions on which they coincide.

Case 2. $d(v, v') > 2t$ (which is possible only if $\gamma > \beta$).

For each pair of such vertices v and v' there exists a vertex v'' adjacent to v such that $d(v'', v') = d(v, v') - t$. To get v'' we only need to change v on t positions on which v and v' differ. Applying this observation $2r$ times, where

$$2r \leqslant \left\lceil \frac{n - 2t}{t} \right\rceil + 1 = \left\lceil \frac{n}{t} - 2 \right\rceil + 1 \leqslant \frac{n}{\frac{\beta}{\beta+\gamma} n} = \frac{\beta + \gamma}{\beta}$$

one can construct a path $v v_1 \cdots v_{2r}$ in H such that for every $1 \leqslant i \leqslant 2r$, we have $d(v_i, v') = d(v_{i-1}, v') - t$ and $d(v_{2r}, v') \leqslant 2t$. Notice that in this case $2t < n$, so $2t < 2n - 2t$ and thus $d(v_{2r}, v') \leqslant \min\{2t, 2n - 2t\}$. As $d(v_{2r}, v')$ is even, one can connect vertices v_{2r} and v' by a path of length two using the argument from Case 1.

Case 3. $2n - 2t < d(v, v') \leqslant 2t$ (which is possible only if $\beta > \gamma$).

For each such v and v' there exist a path $v v_1 v_2$ such that $d(v_2, v') = d(v, v') - 2(n - t)$. To obtain v_1 from v, we need to change all $n - d(v, v')$ positions on which v, v' do not differ and $t - n + d(v, v')$ among other positions. Then, $d(v, v_1) = n - d(v, v') + t - n + d(v, v') = t$. To obtain v_2 from v_1, we need to change all $n - d(v, v')$ positions on which v, v' do not differ, all $n - t$ positions on which v and v_1 are the same and $2t - 2n + d(v, v') > 0$ other positions. Then, indeed $d(v_1, v_2) = n - d(v, v') + n - t + 2t - 2n + d(v, v') = t$ and $d(v_2, v') = 2t - 2n + d(v, v')$.

Arguing in the same way we find a path $v v_1 \cdots v_{2s}$ of length

$$2s \leqslant \left\lceil \frac{n - (2n - 2t)}{n - t} \right\rceil + 1 = \left\lceil \frac{n}{n - t} - 2 \right\rceil + 1 \leqslant \frac{n}{\frac{\gamma}{\beta+\gamma} n - 2} \leqslant \frac{2(\beta + \gamma)}{\gamma}$$

such that for every $i \leqslant s$, $d(v_{2i}, v') = d(v_{2i-2}, v') - 2(n - t)$ and $d(v_{2s}, v') \leqslant 2n - 2t$. Now, since $d(v_{2s}, v') \leqslant \min\{2t, 2n - 2t\}$, and $d(v_{2s}, v')$ is even, we can apply Case 1 to connect v_{2s} and v' by a path of length two.

Consequently, we have shown that the diameter D of H is bounded from above by

$$D \leqslant 2\frac{\beta+\gamma}{\gamma}+3 \qquad \square$$

As a direct consequence of Theorem 3 and Lemma 4 we get the following result on expansion properties of H.

Lemma 5. *Let W be a subset of the vertex set of H such that*

$$|W| \leqslant |V(H)|/2 = 2^{n-1}.$$

Then there exists a constant $c = c(\beta,\gamma) > 0$ such that

$$e_H(W,\overline{W}) \geqslant c|W|\binom{n}{t}.$$

Proof. Let W, $|W| \leqslant 2^{n-1}$, be a set of vertices of H. Since H is an $\binom{n}{t}$-regular graph,

$$\mathrm{Vol}(W) = \binom{n}{t}|W| \leqslant \binom{n}{t}\frac{|V(H)|}{2} = \frac{\mathrm{Vol}(V(H))}{2}.$$

Since H is edge transitive, by Theorem 3 we get

$$\frac{e_H(W,\overline{W})}{\mathrm{Vol}(W)} \geqslant \frac{1}{2D},$$

where D is the diameter of H. By Lemma 4, D is bounded above by a constant, so for some positive constant c we have

$$e_H(W,\overline{W}) \geqslant \frac{1}{2D}\mathrm{Vol}(W) \geqslant c|W|\binom{n}{t} \qquad \square$$

Let us return to the random graph $\mathcal{H}$. Recall that $\mathcal{H}$ is a bipartite graph with a bipartition (V_1, V_2), where $|V_1| = |V_2| = 2^{n-1}$. We will use the Hall's condition, which states that a bipartite graph $G(V,U)$, $|V| = |U|$ does not have a perfect matching iff there exists a set $R \subseteq V$ or $R \subseteq U$ such that

$$|N_G(R)| < |R|, \tag{2}$$

where $N_G(R)$ is the set of all vertices adjacent in G to the vertices from R. Suppose G does not have a perfect matching. Let S be the smallest set $S \subseteq V$ or $S \subseteq U$ which satisfies (2). Without loss of generality, suppose $S \subseteq V$. Assume $|N_G(S)| < |S|-1$. Then we can delete any $|S|-|N_G(S)|-1$ vertices from S to obtain a set smaller than S which also satisfies (2). Since S is the smallest set satisfying (2) this situation is impossible, so $|N_G(S)| = |S|-1$. Moreover the set $S' = U \setminus N_G(S)$ does not have neighbours in S, i.e. $N_G(S') \subseteq V \setminus S$, so $|N_G(S')| \leqslant |V|-|S|$, while $|S'| = |U|-|S|+1 > |N_G(S')|$. Hence

$|S'|$ also satisfies (2). $|S'| + |S| = |U| + 1$, so as S is the smallest set which satisfies (2), $|S| \leqslant |U|/2$.

Therefore if $\mathcal{H}$ does not have a perfect matching, there exists a set $S \subseteq V_1$ or $S \subseteq V_2$ such that $|N_{\mathcal{H}}(S)| = |S| - 1$ and $|S| \leqslant |V_1|/2 = 2^{n-2}$. Let $\mathcal{A}$ be the event that such a subset S exists in $\mathcal{H}$. Let $\mathcal{A}_1$ be the event that such a subset $S \subseteq V_1$ exists. Let $\mathcal{A}_2$ be the event that such a subset $S \subseteq V_2$ exists. Then $\mathbb{P}(\mathcal{A}_1) = \mathbb{P}(\mathcal{A}_2)$, hence

$$\mathbb{P}(\mathcal{H} \text{ does not contain a perfect matching}) = \mathbb{P}(\mathcal{A}) \leqslant 2\mathbb{P}(\mathcal{A}_1).$$

For two fixed sets $S \subseteq V_1$, $|S| \leqslant 2^{n-2} = N/4$ and $T \subseteq V_2$, $|T| = |S| - 1$, let $\mathcal{A}_{S,T}$ denote the event that T is the neighbourhood of S in the random graph $\mathcal{H}$. In order to estimate the probability of $\mathcal{A}_{S,T}$ we apply Lemma 5 to the set $W = S \cup T$. Clearly $|W| = 2|S| - 1 < N/2$.

For deterministic H we have

$$e_H(S, V_2 \setminus T) + e_H(T, V_1 \setminus S) = e_H(W, \overline{W}) \geqslant c\binom{n}{t}|W| = c\binom{n}{t}(2|S| - 1), \tag{3}$$

while from the regularity of H we get

$$e_H(S, \overline{S}) = e_H(S, V_2 \setminus T) + e_H(S, T) = \binom{n}{t}|S|, \tag{4}$$

and

$$e_H(T, \overline{T}) = e_H(T, V_1 \setminus S) + e_H(T, S) = \binom{n}{t}|T| = \binom{n}{t}(|S| - 1). \tag{5}$$

Adding (3) and (4) and subtracting (5), we obtain that in H,

$$e_H(S, V_2 \setminus T) \geqslant \frac{1}{2}\binom{n}{t}\left(|S| + 2c|S| - c - |S| + 1\right) \geqslant c'\binom{n}{t}|S|\,,$$

for some constant $c' > 0$. Thus if $\mathcal{A}_{S,T}$ occurs, $c'\binom{n}{t}|S|$ fixed pairs of vertices which are adjacent in H are not adjacent in $\mathcal{H}$. Observe that for each pair u, v of vertices with Hamming distance t, the probability that there exists an edge $\{u, v\}$ is at least $\beta^t\gamma^{n-t}$. Thus, the probability of $\mathcal{A}$ that Hall's condition fails for some set S, $|S| \leqslant 2^{n-2} = N/4$, is bounded from above by

$$\begin{aligned}\mathbb{P}(\mathcal{A}) &\leqslant 2 \sum_{\substack{S \subseteq V_1 \\ |S| \leqslant N/4}} \sum_{\substack{T \subseteq V_2 \\ |T| = |S| - 1}} \mathbb{P}(\mathcal{A}_{S,T}) \\ &\leqslant 2 \sum_{s=1}^{N/4} \binom{N/2}{s}\binom{N/2}{s-1}\left(1 - \beta^t\gamma^{n-t}\right)^{c's\binom{n}{t}} \\ &\leqslant 2 \sum_{s=1}^{N/4} N^{2s} \exp\left(-c's\binom{n}{t}\beta^t\gamma^{n-t}\right)\,.\end{aligned}$$

Since

$$\sum_{i=0}^{n} \binom{n}{i} \beta^i \gamma^{n-i} = (\beta + \gamma)^n$$

and t was chosen to correspond the largest term in the sum, so for n large enough

$$\binom{n}{t} \beta^t \gamma^{n-t} \geqslant (\beta + \gamma)^n / n \ .$$

Hence

$$\mathbb{P}(\mathcal{A}) \leqslant 2 \sum_{s=1}^{N/4} \left(2^{2n} \exp\left(-c'(\beta + \gamma)^n / n\right) \right)^s$$

and since the term in brackets is exponentially small, it is maximal when $s = 1$ that is

$$\mathbb{P}(\mathcal{A}) \leqslant 2 \sum_{s=1}^{N/4} \left(2^{2n} \exp\left(-c'(\beta + \gamma)^n / n\right) \right)^s \leqslant 2^n 2^{2n} \exp\left(-c'(\beta + \gamma)^n / n\right) = o(1).$$

Consequently, a.a.s. $\mathcal{H}$, and thus also $\mathcal{K}(n, \mathbf{P})$, contains a perfect matching. □

In the proof we have found a perfect matching in a bipartite subgraph $\mathcal{H}$ of $\mathcal{K}(n, \mathbf{P})$, containing only the edges joining vertices which are at Hamming distance

$$t = 2 \left\lceil \frac{\beta}{2(\beta + \gamma)} n \right\rceil + 1$$

in the hypercube. Note however that if we take k edge-disjoint subgraphs $\mathcal{H}_\ell$, for $\ell \in [k]$, containing the edges of $\mathcal{K}(n, \mathbf{P})$ joining vertices at Hamming distance

$$t = t(\ell) = 2 \left\lceil \frac{\beta}{2(\beta + \gamma)} n \right\rceil + 2\ell + 1.$$

respectively, we can mimic our argument to construct k edge-disjoint perfect matchings.

Thus, let k-PM denote the property, that a graph contains k edge-disjoint perfect matchings.

Theorem 6. *Let $k \in \mathbb{N}$, $k \geqslant 2$ be a constant.*

$$\lim_{n \to \infty} \mathbb{P}(\mathcal{K}(n, \mathbf{P}) \ \textit{has } k\textit{-PM property}) = \begin{cases} 0 & \textit{if} \quad \beta + \gamma \leqslant 1 \\ 1 & \textit{if} \quad \beta + \gamma > 1. \end{cases}$$

In particular

$$\lim_{n \to \infty} \mathbb{P}(\mathcal{K}(n, \mathbf{P}) \ \textit{contains } k\textit{-factor}) = \begin{cases} 0 & \textit{if} \quad \beta + \gamma \leqslant 1 \\ 1 & \textit{if} \quad \beta + \gamma > 1. \end{cases}$$

Note the difference between the cases $k = 1$ and $k \geqslant 2$ for $\beta = 1$ and $\gamma = 0$ when, as we have already observed, a.a.s. the minimum degree of $\mathcal{K}(n, \mathbf{P})$ is one.

Acknowledgements

I would like to thank Professor Tomasz Łuczak for stimulating discussion and valuable comments.

References

[1] F. Chung. *Spectral Graph Theory.* CBMS Conference on Recent Advances in Spectral Graph Theory Vol. 92, AMS Publications, 1997.

[2] M. Mahdian and Y. Xu. Stochastic Kronecker Graphs. *Random Structures and Algorithms*, 38(4): 453–466, 2011.

[3] P. Stanica. Good lower and upper bounds on binomial coefficients. *JIPAM. Journal of Inequalities in Pure and Applied Mathematics*, 2(3): 1–4, 2001.

[4] D. Chakrabarti, C. Faloutsos, J. Kleinberg and J. Leskovec. Realistic, mathematically tractable graph generation and evolution, using Kronecker multiplication. *Proceedings of the European Conference on Principles and Practice of Knowledge Discovery in Databases*: 133–145, 2005.

[5] D. Chakrabarti, C. Faloutsos, Z. Ghahramani, J. Kleinberg and J. Leskovec. Kronecker graphs: an approach to modeling networks. *Journal of Machine Learning Research*, 11: 985–1042, 2010.

[6] T. Kolda, A. Pinar and C. Seshadhri. An in-depth analysis of stochastic Kronecker graphs. *Journal of the ACM*, 60(2): 1–32, 2013.

[7] P. Horn and M. Radcliffe. Giant components in Kronecker graphs. *Random Structures and Algorithms*, 40(3): 385–397, 2012.

[8] M. Kang, M. Karoński, Ch.Koch and T.Makai. Properties of stochastic Kronecker graphs. *Journal of Combinatorics*, 4(6): 395–432, 2015.

[9] M. Radcliffe, S. Young. Connectivity and giant components of Stochastic Kronecker Graphs. *Journal of Combinatorics*, 6(4): 457–482, 2015.