Uniwersytet im. Adama Mickiewicza w Poznaniu
Wydział Matematyki i Informatyki

Justyna Banaszak

ON THE STRUCTURE OF KRONECKER GRAPHS

O strukturze grafów Kroneckera

Rozprawa doktorska w dziedzinie nauk matematycznych
w dyscyplinie matematyka

Promotor: prof. dr hab. Tomasz Łuczak

Poznań, 2019

I would like to thank professor Tomasz Łuczak for his priceless help, support and patience throughout my work on this dissertation. All the discussions with him were inspiring.

I would also like to thank my family, especially my husband, who supported me in every way and my newborn son, who was sleeping more than I expected.

Abstract

In the thesis we study the asymptotic structure of Kronecker graphs. Kronecker graphs were introduced in computer science as possible models of small-world networks and have been studied by a number of combinatorialists and probabilists as an interesting model of random graph in which the neighbourhoods of different vertices are correlated.

Kronecker graph is a graph on vertex set $\{0,1\}^n$, where the probability that two vertices are connected depends on the number of positions in their labels on which they have common zeros, common ones, and different values. We study the behaviour of Kronecker graphs as n tends to infinity.

In the work we examine several properties of Kronecker graphs such as k-connectivity, the existence of a perfect matching, and constant diameter. The results are partially based on my paper published in *Electronic Journal of Combinatorics* ([1]) and on the article published in *Discrete Mathematics*, by Tomasz Łuczak and myself ([2]).

The thesis starts with basic definitions concerning graph theory and probabilistic tools we use throughout the work. Then we define Kronecker graphs and survey known results on this model of random graphs. In Chapter 3 we prove some useful facts regarding the neighbourhoods of vertices in Kronecker graphs generalising results of Kang, Karoński, Koch, and Makai. The main result of the next chapter, Theorem 4.1, states that the thresholds for connectivity and for the existence of perfect matching basically coincide. In Chapter 5 we show that above the connectivity threshold, with probability tending to 1 as n tends to infinity, the Kronecker graphs are δ-connected, where δ denotes the minimum degree of the graph (Theorem 5.1). In the last part of the thesis we study the diameter of Kronecker graphs and prove that just above the connectivity threshold their diameter, with probability tending to 1 as n tends to infinity, is bounded from above by a constant (Theorem 6.1).

Contents

Chapter 1

Notation and basic facts

1.1 Graphs

A graph is one of the most common and frequently used structure in combinatorics. Although graphs appeared already in a work of Leonard Euler on Königsberg bridges from 1736, graph theory as a separate field of mathematics emerged only in the last century. In this section we introduce some basic definitions and notations concerning graphs we use in the thesis. Most of them follow standard literature of the subject such as, for instance, Diestel's monography [11].

By a **graph** G we mean an ordered pair $G = (V, E)$, where V is a non-empty set and E consists of 2-element subsets of V, i.e. $E \subseteq \{\{u, v\} : \ u, v \in V\}$. The set V is the **vertex set** of $G = (V, E)$ and its elements are called **vertices**, while the elements of E are **edges** of the graph. If $\{u, v\} \in E$ we say that u and v are neighbours and that u, v are **adjacent**. We denote it by $u \sim v$. To simplify the notation we often denote an edge $\{u, v\}$ by uv.

The **degree** of a vertex v in graph G, denoted by $\deg_G(v)$, is the number of vertices adjacent to v. The minimum degree of G, denoted by $\delta(G)$, and the maximum degree of G, which we write as $\Delta(G)$, are the minimum and maximum of the degrees of vertices in G, respectively. If $\delta(G) = \Delta(G) = d$, i.e. all vertices of G have the same degree, then it is called d-**regular**. For a vertex set $S \subseteq V$, the **volume** of S, denoted by $\mathrm{Vol}(S)$, is the sum of vertex degrees of S.

If $S \subseteq V$, then we put $\overline{S} = V \setminus S$. The set of vertices from $\overline{S}$ adjacent to vertices from S is called **the neighbourhood** of S and is denoted by $N_G(S)$. If S is a singleton, i.e. $S = \{v\}$, we write $N_G(\{v\}) = N_G(v)$. The set of edges between two disjoint vertex sets $S, T \subseteq V$ is denoted by $E_G(S, T)$, and by $e_G(S, T) = |E_G(S, T)|$ we mean its size. By $E_G(S)$ we mean the set of edges with both ends in S.

We say that $G = (V, E)$ is a **bipartite** graph, if there exists a partition $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$ such that every edge $e \in E$ has one end in V_1 and another in V_2. In such a case we often write $G = (V_1 \cup V_2, E)$.

A **subgraph** of a graph $G = (V, E)$ is a graph $H = (V_H, E_H)$ such that $V_H \subseteq V$ and $E_H \subseteq E$. For a vertex set $S \subseteq V$, a subgraph $G[S]$ **induced by** S is a graph $G[S] = (S, E_S)$ such that $E_S = \{\{u, v\} \in E : \ u, v \in S\}$.

We say that $G = (V, E)$ is **connected** if we cannot partition its vertex set into two non-empty parts V_1 and V_2 such that $e_G(V_1, V_2) = 0$. A **component** of G is a maximal connected subgraph of G. For $k \in \mathbb{N}$ we say that G is k-**edge connected** if after removing any $k - 1$ edges from E it remains connected.

All the following definitions refer to graph $G = (V, E)$.

A **path** P between two vertices u, v of a graph G is a connected subgraph of G in which both u and v have degree one and all the other vertices of P have degree two. The number of edges in a path is called its **length**. The **distance** between two vertices u, v is the length of a shortest path between u and v. It is denoted by $\mathrm{dist}_G(u, v)$. The set of vertices within distance k from a vertex v is called the k-th **neighbourhood** of v (denoted by $N_G^k(v)$). The **diameter** $\mathrm{diam}(G)$ of a graph G is the longest distance between pairs of its vertices.

A k-**factor** in G is a k-regular subgraph of G on vertex set V.

A **matching** in a graph G is a set of edges $M \subseteq E$ such that each vertex in G belongs to at most one edge from M. We say that a matching M **saturates** a set $S \subseteq V$, when each vertex in S belongs to exactly one edge from M. A **perfect matching** in a graph is a matching saturating all its vertices.

For a graph $G = (V, E)$, a bijection $\sigma : V \to V$ is a **graph automorphism** if it preserves the edges of G, i.e. for every pair $u, v \in V$, $u \sim v$ if and only if $\sigma(u) \sim \sigma(v)$. We say that $G = (V, E)$ is **edge-transitive**, if for any two edges $e, e' \in E$ there exists a graph automorphism $\sigma : V \to V$, which transforms e into e'.

Finally, if it is clear which graph G we have in mind, we often omit lower index G in all graph notation (for example for a vertex degree we put $\deg(v) = \deg_G(v)$).

1.2 Random graphs

Although "randomly constructed graphs" have been analysed by researchers, in particular by epidemiologists, since the beginning of the XXth century, a systematic study of random graphs started, basically, with series of papers Paul Erdős and Alfréd Rényi published in late fifties and early sixties 1960 [12], [13], [14], [15], [16]. Nowadays random graph theory is a well-established part of combinatorics – for an overview of the field see, for instance, monographies [5], [18], [21].

In this work we use the following, fairly general, definition of a random graph. Let $n \in \mathbb{N}$ and P be a symmetric $n \times n$ matrix, where each entry $P[u, j] \in [0, 1]$. A **random**

graph $G_{n,P}$ is a graph with vertex set $[n] = \{1, 2, \ldots, n\}$, where each pair of vertices $i, j \in [n]$ is an edge with probability $P[i, j]$. We say that P is a **probability matrix** of graph $G_{n,P}$.

Example. In the most frequently studied **binomial model** of a random graph $G(n, p)$, for each $i, j \in [n]$, we have $P[i, j] = p$, i.e. each pair of distinct vertices is connected by an edge independently with the same probability p. This model of a random graph was introduced in [19] and is basically equivalent to the uniform model of random graph studied in the seminal papers of Erdős and Rényi.

1.3 Asymptotic notation

As we study the random graph model, we need some asymptotic notation.

We say that a sequence of events $(\mathcal{A}_n)$ occurs **asymptotically almost surely** (a.a.s.) if $\lim_{n\to\infty} \mathbb{P}(\mathcal{A}_n) = 1$.

Let $f : \mathbb{N} \to \mathbb{R}_+$ and $g : \mathbb{N} \to \mathbb{R}_+$ be two functions. We say that

- $f(n) = O(g(n))$ if there exist constants $c > 0$ and $n_0 > 0$ such that
$$\forall_{n \geq n_0} f(n) \leq cg(n)\,.$$
- $f(n) = o(g(n))$ if for every constant $c > 0$ there exists $n_0 > 0$ such that
$$\forall_{n \geq n_0} f(n) \leq cg(n)\,.$$
- $f(n) = \Theta(g(n))$ if there exist constants $c_1 > 0$, $c_2 > 0$, $n_0 > 0$ such that
$$\forall_{n \geq n_0} c_1 g(n) \leq f(n) \leq c_2 g(n)\,.$$

1.4 Tensor product

Let A be a matrix of size $m \times n$ and B be a matrix of size $p \times q$. The **Kronecker (tensor) product** of A and B is the matrix

$$C = A \otimes B = \begin{pmatrix} A[1,1]B & A[1,2]B & \cdots & A[1,n]B \\ A[2,1]B & A[2,2]B & \cdots & A[2,n]B \\ \vdots & \vdots & \ddots & \vdots \\ A[m,1]B & A[m,2]B & \cdots & A[m,n]B \end{pmatrix}$$

of size $mp \times qn$, where for $i \in [m] = \{1, 2, \ldots, m\}$, $j \in [n] = \{1, 2, \ldots, n\}$, $k \in [p] = \{1, 2, \ldots, p\}$, $l \in [q] = \{1, 2, \ldots, q\}$,

$$C[(i-1)p + k, (j-1)q + l] = A[i,j]B[k,l]\,.$$

We define the n-fold Kronecker product of a matrix A as

$$A^{\otimes n} = \underbrace{A \otimes A \otimes \ldots \otimes A}_{n \text{ times}} .$$

1.5 Tools

A necessary and sufficient condition for the existence of a perfect matching in a bipartite graph G is stated in Hall's theorem (see for instance [11]).

Theorem 1.1 (Hall). *A bipartite graph $G = (W \cup U, E)$ contains a matching saturating W if and only if for every $S \subseteq W$:*

$$|N_G(S)| \geq |S| . \tag{1.1}$$

The following theorem is a direct consequence of Theorem 1.1.

Theorem 1.2. *Let $G = (W \cup U, E)$ be a bipartite graph, where $|W| = |U|$. G contains no perfect matching if and only if there exists a set S, where $|S| < |W|/2$, such that either $S \subseteq W$ or $S \subseteq U$ and*

$$|N_G(S)| = |S| - 1 .$$

Proof. Suppose that G contains a perfect matching. Then, by Hall's theorem, $|N_G(S)| \geq |S|$ whenever $S \subseteq W$ or $S \subseteq U$.

Now suppose G does not have a perfect matching. Let R be the smallest set which is contained in either W or in U for which (1.1) does not hold. Without loss of generality, suppose $R \subseteq W$. Assume $|N_G(R)| < |R| - 1$. Then we can delete any $|R| - |N_G(R)| - 1$ vertices from R to obtain a set smaller than R which also satisfies (1.1). Since R is the smallest set satisfying (1.1) we infer that $|N_G(R)| = |R| - 1$. Moreover the set $R' = U \setminus N_G(R)$ does not have neighbours in R, i.e. $N_G(R') \subseteq W \setminus R$, so $|N_G(R')| \leq |W| - |R|$, while $|R'| = |W| - |R| + 1$. Hence $|R'|$ also does not satisfy (1.1). Since $|R'| + |R| = |U| + 1$, as R is the smallest set which does not satisfy (1.1), we have $|R| \leq |U|/2 = |W|/2$.

■

The following theorem was proven by Fan Chung in [10] (Theorem 7.1). It basically says that the edge-transitive graphs with small diameter are good expanders.

Theorem 1.3. *Let $G = (V, E)$ be an edge-transitive graph of diameter D. Then for every $S \subseteq V$, such that $\mathrm{Vol}(S) \leq \frac{\mathrm{Vol}(V)}{2}$ we have*

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

The following probabilistic facts are used extensively throughout the thesis. For their proofs see for example [18].

Theorem 1.4 (Markov's Inequality). *Let X be a non-negative random variable. Then for all $t > 0$*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}\,.$$

Theorem 1.5 (The First Moment Method). *Let X be a non-negative, integer-valued random variable. Then*

$$\mathbb{P}(X > 0) \leq \mathbb{E}X\,.$$

Let X be a random variable. If X has the binomial distribution with parameters n and p, we denote it by $X = \mathrm{Bi}(n,p)$. The following two inequalities say that under some specified condition, the binomial random variable is concentrated around its expectation. Their proofs can be found for example in [21].

Theorem 1.6 (Chernoff's Inequality). *If $X \in \mathrm{Bi}(n,p)$, then, for $t \geq 0$,*

$$\mathbb{P}(X \geq \mathbb{E}X + t) \leq \exp\left(-\frac{t^2}{2(\mathbb{E}X + t/3)}\right),$$
$$\mathbb{P}(X \leq \mathbb{E}X - t) \leq \exp\left(-\frac{t^2}{2\mathbb{E}X}\right).$$

As a direct consequence we get the following inequalities

$$\begin{aligned}\mathbb{P}(X \geq 2\mathbb{E}X) &\leq \exp\left(-\frac{3}{8}\mathbb{E}X\right),\\ \mathbb{P}(X \leq \mathbb{E}X/2) &\leq \exp\left(-\frac{1}{8}\mathbb{E}X\right).\end{aligned} \tag{1.2}$$

Theorem 1.7 (Talagrand's Inequality). *Suppose that $Z_1, Z_2, \dots, Z_N$ are independent random variables taking their values in some sets $\Lambda_1, \Lambda_2 \dots, \Lambda_N$, respectively. Suppose further that $X = f(Z_1, Z_2, \dots, Z_N)$, where $f : \Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_N \to \mathbb{R}$ is a function such that, for some constants c_k, $k = 1, 2, \dots, N$, and some function ψ, the following two conditions hold:*

1. *If $z, z' \in \Lambda = \prod_{i=1}^{N} \Lambda_i$ differ only in the k-th coordinate, then $|f(z) - f(z')| \leq c_k$.*

2. *If $z \in \Lambda$ and $r \in \mathbb{R}$ with $f(z) \geq r$, then there exists a set $J \subseteq \{1, 2, \dots, N\}$ with $\sum_{i \in J} c_i^2 \leq \psi(r)$, such that for all $y \in \Lambda$ with $y_i = z_i$ when $i \in J$, we have $f(y) \geq r$.*

Then, for every $r \in \mathbb{R}$ and $t \geq 0$,

$$\mathbb{P}(X \leq r - t)\mathbb{P}(X \geq r) \leq \exp\left(-\frac{t^2}{4\psi(r)}\right).$$

In particular, if m is a median of X, then for every $t \geq 0$,

$$\mathbb{P}(X \leq m - t) \leq 2\exp\left(-\frac{t^2}{4\psi(m)}\right)$$

and

$$\mathbb{P}(X \geq m + t) \leq 2\exp\left(-\frac{t^2}{4\psi(m+t)}\right).$$

The following is a direct consequence of Theorem 1.7.

Corollary 1.8. *Let $I_1, I_2, \ldots, I_N$ be independent random variables taking their values in set $\{0,1\}$ and let $X = f(I_1, I_2, \ldots, I_N)$, where $f : \{0,1\}^N \to \mathbb{N}$. Suppose that for some constant $c \in \mathbb{N}$*

1. *If X_i denote $X_i = X - I_i$ for $i = 1, \ldots, N$, then for every $i \in 1, \ldots, N$, $|X - X_i| \leq 1$.*
2. *If for some $x \in \{0,1\}^N$ and $k \in \mathbb{N}$, $f(x) \geq k$, then there exists a set $J \subseteq \{1, 2, \ldots, N\}$ of size $|J| = ck$, such that for all $y \in \{0,1\}^N$ with $y_i = x_i$ when $i \in J$ we have $f(y) \geq k$.*

Then

$$\mathbb{P}\left(X < \frac{1}{2}\mathbb{E}X\right) \leq 2\exp\left(-\frac{\mathbb{E}X}{16c}\right).$$

Finally, we state two useful elementary facts we often use in the technical parts of the proofs.

Fact 1.9. *For every $x \in \mathbb{R}$ we have*

$$1 + x \leq e^x. \tag{1.3}$$

For every integer n, k

$$\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k. \tag{1.4}$$

Fact 1.10. *Let $n \to \infty$ be an integer and let $a, b > 0$. Then*

$$\binom{n}{i} a^i b^{n-i}, \quad i \in \{0, 1, \ldots, n\}$$

is maximized for $i = \frac{a}{a+b}n + O(1)$. Consequently, there exists a function $f(n) = \frac{a}{a+b}n + O(1)$ such that

$$\binom{n}{f(n)} a^{f(n)} b^{n-f(n)} \geq (1 - o(1))\frac{(a+b)^n}{n+1}. \tag{1.5}$$

Since for every $i = \frac{a}{a+b}n + O(1)$ we have

$$\binom{n}{i+1} a^{i+1} b^{n-i-1} = \Theta\left(\binom{n}{i} a^i b^{n-i}\right),$$

for every constant A there exists a positive constant c such that

$$\binom{n}{f(n)} a^{f(n)} b^{n-f(n)} \geq c \frac{(a+b)^n}{n+1} \tag{1.6}$$

whenever

$$\left| f(n) - \frac{a}{a+b} n \right| \leq A\,.$$

Chapter 2

Kronecker graphs - the definition and known results

In this section we define Kronecker graph $\mathcal{K}(n, \mathbf{P})$ which is the main object studied in the thesis. We start with its definition and basic notions used to characterize its properties. Then we briefly survey the known properties of $\mathcal{K}(n, \mathbf{P})$.

2.1 Definition

Kronecker graph is a random graph, where vertices are binary vectors of length n, and the probability that two vertices u, v are connected depends on the number of positions, on which u and v differ, the number of positions on which they both have zeros and the number of positions on which they both have ones.

Now we define Kronecker graph more precisely. Let n be a natural number and let $\alpha, \beta, \gamma \in [0, 1]$ be constants. Denote by $\mathbf{P}$ a matrix:

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

where 0's and 1's are labels of rows and columns of $\mathbf{P}$. **Kronecker graph** $\mathcal{K}(n, \mathbf{P})$ is a graph with vertex set $\{0, 1\}^n$ and with probability matrix $\mathbf{P}^{\otimes n}$ (n-fold Kronecker product of $\mathbf{P}$). Equivalently, for two vertices $v = (v_1, v_2, \dots, v_n)$, $u = (u_1, u_2, \dots, u_n)$ of $\mathcal{K}(n, \mathbf{P})$, the probability that u, v are adjacent is given by

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

For technical reasons we allow self loops with weight 1/2 (which are counted once in the expected degree of a vertex). Throughout the thesis we denote by V and E the vertex and edge set of $\mathcal{K}(n, \mathbf{P})$. The size of the vertex set is denoted by N, i.e. $N = 2^n$.

For example, let $n = 4$, $u = (1,0,0,1)$ and $v = (1,1,0,0)$. The probability that u, v are adjacent is

$$p_{u,v} = \mathbf{P}[1,1]\mathbf{P}[0,1]\mathbf{P}[0,0]\mathbf{P}[1,0] = \alpha\beta^2\gamma\,.$$

We often represent a vertex v in $\mathcal{K}(n, \mathbf{P})$ as in the following Figure 2.1

Figure 2.1: A graphical representation of a vertex

where grey and white rectangles correspond to one and zero coordinates of v respectively.

Then, the probability that two vertices v, u, represented on Figure 2.2 below

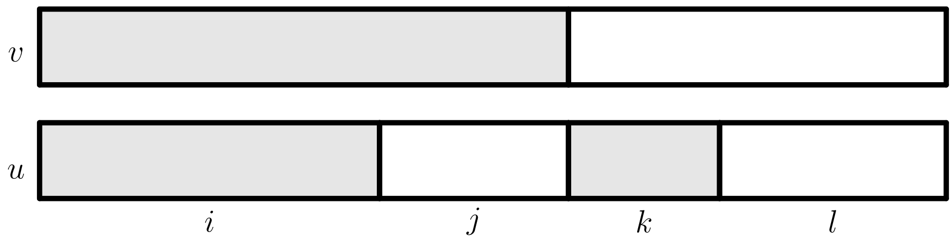

Figure 2.2: Zeros and ones in two vertices v and u of $\mathcal{K}(n, \mathbf{P})$

are connected by an edge in $\mathcal{K}(n, \mathbf{P})$ is given by

$$p_{u,v} = \alpha^i \beta^{j+k} \gamma^l\,,$$

where $i + j + k + l = n$.

Note that by the symmetry of the definition of the Kronecker graph, without loss of generality we may assume that $\alpha \geq \gamma$, since otherwise we can interchange the role of zeros and ones. **We will follow this rule throughout the thesis.**

The **weight** of a vertex $v = (v_1, v_2, \ldots, v_n)$, denoted by $\mathrm{w}(v)$, is the number of ones in its label, i.e.

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

The **Hamming distance** $\mathrm{d}(u, v)$ between vertices $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n)$ of $\mathcal{K}(n, \mathbf{P})$, is the number of coordinates on which u,v have different values, i.e.

$$\mathrm{d}(u, v) = \sum_{i=1}^{n} |v_i - u_i|\,.$$

We say that u, v form an i-**acquaintance** if $u \sim v$ and $\mathrm{d}(u, v) = n - i$, and they form an $[i, j]$-**acquaintance** when $u \sim v$ and they have i common ones and j common zeros.

Note that the probability that two vertices which have i common ones and j common zeros form an $[i, j]$-acquaintance is given by

$$p_{u,v} = \alpha^i \beta^{n-i-j} \gamma^j.$$

A **graph property** $\mathcal{P}$ is a subfamily of the family of all graphs. A graph property $\mathcal{P}$ is **increasing** if for every graph G, whenever $G \in \mathcal{P}$ and G' is a graph obtained from G by adding an edge, we have also $G' \in \mathcal{P}$. A graph property $\mathcal{P}$ is **decreasing** if for every graph G, whenever $G \in \mathcal{P}$ and G'' is a graph obtained from G by removing an edge, we have $G'' \in \mathcal{P}$.

We examine the behaviour of $\mathcal{K}(n, \mathbf{P})$ when $n \to \infty$.

Let $\mathcal{P}$ be an increasing graph property. If there exists a function $f : [0, 1] \times [0, 1] \times [0, 1] \to \mathbb{R}$, increasing on each coordinate, such that for every $\alpha, \beta, \gamma \in [0, 1]$,

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{K}(n, \mathbf{P}) \in \mathcal{P}) = \begin{cases} 0 & \text{if } f(\alpha, \beta, \gamma) < 1 \\ 1 & \text{if } f(\alpha, \beta, \gamma) > 1, \end{cases}$$

then we say that $f(\alpha, \beta, \gamma) = 1$ is the **threshold** for property $\mathcal{P}$.

Analogously, we define the threshold for a decreasing graph property.

2.2 Results

Kronecker graphs were introduced by computer scientists in papers [3], [8], [9], [17], [24]. Using computer simulations the authors applied the Kronecker graphs to study some real world networks such as Internet Autonomous Systems with parameters $\alpha = 0.98$, $\beta = 0.58$, $\gamma = 0.06$, Citation Graphs for High-Energy Physics Theory and As-Route-Views: a data set consisting of a single snapshot of connectivity in Internet Autonomous Systems. They also argued that Kronecker graphs exhibit phase transition.

The first work which rigorously studied Kronecker graphs was presented by Mahdian and Xu at a conference in 2007 and published later in [25]. Their main result concerns the connectivity threshold for $\mathcal{K}(n, \mathbf{P})$.

Theorem 2.1. *The connectivity threshold for $\mathcal{K}(n, \mathbf{P})$ is $\beta + \gamma = 1$ i.e. $\mathcal{K}(n, \mathbf{P})$ is a.a.s. connected for $\beta + \gamma > 1$, and is a.a.s. disconnected whenever $\beta + \gamma < 1$.*

Note that as we assume $\alpha \geq \gamma$, this condition is equivalent to $\alpha + \beta > 1$ and $\beta + \gamma > 1$.

The result was later developed by Radcliffe and Young [26], who studied generalised model of Kronecker graphs, i.e., when the starting matrix $\mathbf{P}$ is of dimension $k \times k$ (for any integer $k \geq 2$). For the special case of our standard model with $k = 2$, their result can be stated as follows.

Theorem 2.2.

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

Mahdian and Xu [25] investigated also the emergence of the giant component (i.e. the component of $\mathcal{K}(n,\mathbf{P})$ of size $\Theta(N)$) and the diameter of $\mathcal{K}(n,\mathbf{P})$ under additional assumption that $\gamma < \beta < \alpha$. It should be emphasised, however, that in this range of parameters the properties of $\mathcal{K}(n,\mathbf{P})$ are much easier to study, as the probability that an edge exists increases when in the label of one of its ends we change zeros to ones in any positions.

Later the properties of the giant component were studied for all sets of parameters α, β, γ by Horn and Radcliffe [20]. The authors showed that, in the general case, significantly more complicated than the one when we assume $\gamma < \beta < \alpha$, the threshold for the emergence of the giant component in $\mathcal{K}(n,\mathbf{P})$ is $(\alpha+\beta)(\beta+\gamma) = 1$.

Kang *et al.* [22] showed that, surprisingly, contrary to general beliefs and results of numerical simulations, the vertex degree distribution of $\mathcal{K}(n,\mathbf{P})$ is not power-law for any set of parameters α, β, γ. In the same paper the authors found the thresholds for the emergence of certain classes of small subgraphs such as stars, trees, and cycles.

Chapter 3

Neighbours

In this chapter we examine the size and the structure of the neighbourhood of vertices in Kronecker graphs. Similar results were proved in [22] in the case $\alpha = \gamma$. We generalise them using similar techniques, and supplement them with a number of other properties we shall use later in the thesis.

The explicit formula for the expected degree of a vertex $v \in V$ of weight $w = \mathrm{w}(v)$ can be easily seen to be as follows (see also [25]).

$$\mathbb{E}(\deg(v)) = \sum_{i=0}^{w} \binom{w}{i} \alpha^i \beta^{w-i} \sum_{j=0}^{n-w} \binom{n-w}{j} \beta^j \gamma^{n-w-j} = (\alpha+\beta)^w (\beta+\gamma)^{n-w} . \tag{3.1}$$

Let us examine more closely the largest terms of the binomial sums above. The first sum reaches maximum for $i = \frac{\alpha}{\alpha+\beta} w + O(1)$ and the second reaches maximum for $j = \frac{\beta}{\beta+\gamma}(n-w) + O(1)$, that is for the number of vertices u, which have roughly $\frac{\alpha}{\alpha+\beta} w$ ones on the same positions as v and $\frac{\beta}{\beta+\gamma}(n-w)$ ones on the positions where v has zeros (see Figure 3.1).

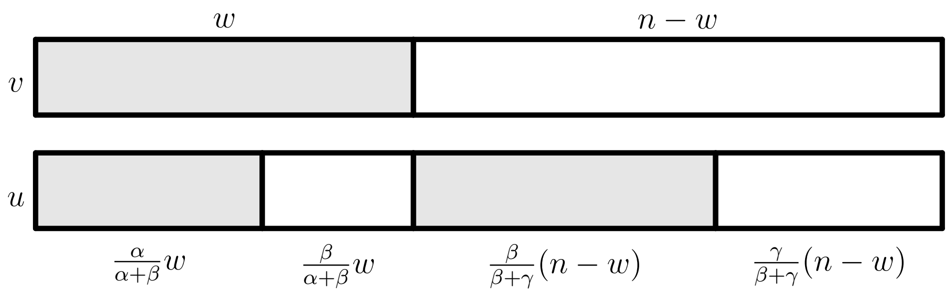

Figure 3.1: A typical neighbour u of a vertex v

Thus, by (1.5), the number of such neighbours u of v is at least

$$(1 - o(1)) \frac{(\alpha+\beta)^w (\beta+\gamma)^{n-w}}{(w+1)(n-w+1)} . \tag{3.2}$$

Now we study the structure of the typical neighbourhood of v more carefully. Our aim is to show that most of the neighbours of v, in fact $(1-o(1))\mathbb{E}\deg(v)$ of them, have weight close to $\widetilde{\mathrm{w}}$, where

$$\widetilde{\mathrm{w}}(v) = \frac{\alpha}{\alpha+\beta}\,\mathrm{w}(v) + \frac{\beta}{\beta+\gamma}(n-\mathrm{w}(v))\,.$$

To this end for a vertex $v \in V$ and a positive constant $\epsilon > 0$ we put

$$U(v) = \{u \in V:\ \mathrm{w}(u) = \widetilde{\mathrm{w}}(v)\}\,,$$

and

$$U_\epsilon(v) = \{u \in V:\ |\,\mathrm{w}(u) - \widetilde{\mathrm{w}}(v)| < \epsilon n\}\,.$$

The neighbours of v which lie in $U(v)$ are called **good neighbours** of v and those from $U_\epsilon(v)$ are **ϵ-good neighbours** of v. The set of all good [ϵ-good] neighbours of v is the **good** [**ϵ-good**] **neighbourhood** of v, denoted by $N^G(v)$ [$N^G_\epsilon(v)$]. The neighbours of v which lie outside $U(v)$ and $U_\epsilon(v)$ are called **bad** and **ϵ-bad neighbours** of v, respectively. Finally, the **bad** [**ϵ-bad**] **neighbourhood** of v, denoted by $N^B(v)$ [$N^B_\epsilon(v)$], is the set of all bad [ϵ-bad] neighbours of v.

The following two lemmata show that for every constant $\epsilon > 0$ most of the neighbours of v are ϵ-good.

Lemma 3.1. *Let v be a vertex of weight $w = \mathrm{w}(v)$. Then*

$$\mathbb{E}|N^G(v)| \geq (1-o(1))\,\frac{(\alpha+\beta)^w(\beta+\gamma)^{n-w}}{n^2}\,.$$

Proof. We estimate $N^G(v)$ by the number of all $[\frac{\alpha}{\alpha+\beta}w, \frac{\gamma}{\beta+\gamma}(n-w)]$ - acquaintances of v which belong to it. Thus we get

$$\begin{aligned}\mathbb{E}|N^G(v)| &\overset{(3.2)}{>} (1-o(1))\,\frac{(\alpha+\beta)^w}{w+1}\,\frac{(\beta+\gamma)^{n-w}}{n-w+1}\\ &> (1-o(1))\,\frac{(\alpha+\beta)^w(\beta+\gamma)^{n-w}}{n^2}\,.\end{aligned}$$

■

Lemma 3.2. *Let $\epsilon > 0$ be a constant and let v be a vertex of $\mathcal{K}(n, \mathbf{P})$ of weight $w = \mathrm{w}(v)$. Then*

$$\mathbb{E}|N^G_\epsilon(v)| = (1-o(1))\,\mathbb{E}\,(\deg(v))\,. \tag{3.3}$$

More precisely

$$\mathbb{E}|N^B_\epsilon(v)| \leq \exp\left(-\frac{\epsilon^2}{20}n\right)(\alpha+\beta)^w(\beta+\gamma)^{n-w}\,, \tag{3.4}$$

so

$$\mathbb{E}|N^G_\epsilon(v)| \geq \left(1-\exp\left(-\frac{\epsilon^2}{20}n\right)\right)(\alpha+\beta)^w(\beta+\gamma)^{n-w}\,. \tag{3.5}$$

Proof. Let v be a vertex of weight w. Define

$$A = \left\{ i \in [w] : \left| i - \frac{\alpha}{\alpha+\beta} w \right| > \frac{\epsilon}{2} n \right\},$$

and

$$B = \left\{ j \in [n-w] : \left| j - \frac{\gamma}{\beta+\gamma} (n-w) \right| > \frac{\epsilon}{2} n \right\}.$$

Let $S(v)$ be the set of all $[i,j]$ - acquaintances of v, where $i \in A$ or $j \in B$. Note first that $N_\epsilon^B(v) \subseteq S(v)$. Indeed, if $u \notin S(v)$, then u has respectively $i \notin A$ ones on positions where v has ones, and $j \notin B$ zeros on positions where v has zeros. Thus,

$$\mathrm{w}(u) = i + n - w - j < \frac{\alpha}{\alpha+\beta} w + n - w - \frac{\gamma}{\beta+\gamma}(n-w) + \epsilon n = \widetilde{\mathrm{w}}(v) + \epsilon n$$

and

$$\mathrm{w}(u) = i + n - w - j > \frac{\alpha}{\alpha+\beta} w + n - w - \frac{\gamma}{\beta+\gamma}(n-w) - \epsilon n = \widetilde{\mathrm{w}}(v) - \epsilon n\,.$$

Consequently, $N_\epsilon^B(v) \subseteq S(v)$.

Now observe that

$$\begin{aligned}
\mathbb{E}|S(v)| &< \sum_{i \in A} \sum_{j=0}^{n-w} \alpha^i \beta^{w-i} \beta^{n-w-j} \gamma^j + \sum_{i=0}^{w} \sum_{j \in B} \alpha^i \beta^{w-i} \beta^{n-w-j} \gamma^j \\
&= (\alpha+\beta)^w (\beta+\gamma)^{n-w} \times \\
&\times \left(\sum_{i \in A} \left(\frac{\alpha}{\alpha+\beta} \right)^i \left(\frac{\beta}{\alpha+\beta} \right)^{w-i} + \sum_{j \in B} \left(\frac{\beta}{\beta+\gamma} \right)^{n-w-j} \left(\frac{\gamma}{\beta+\gamma} \right)^j \right).
\end{aligned}$$

By Chernoff's inequality (Theorem 1.6)

$$\begin{aligned}
\sum_{i \in A} \left(\frac{\alpha}{\alpha+\beta} \right)^i \left(\frac{\beta}{\alpha+\beta} \right)^{w-i} &= \mathbb{P}\left(\left| \mathrm{Bi}\left(w, \frac{\alpha}{\alpha+\beta} \right) - \frac{\alpha}{\alpha+\beta} w \right| > \frac{\epsilon}{2} n \right) \\
&< \exp\left(-\frac{\epsilon^2}{10} n \right),
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j \in B} \left(\frac{\beta}{\beta+\gamma} \right)^{n-w-j} \left(\frac{\gamma}{\beta+\gamma} \right)^j &= \mathbb{P}\left(\left| \mathrm{Bi}\left(n-w, \frac{\gamma}{\beta+\gamma} \right) - \frac{\gamma}{\beta+\gamma}(n-w) \right| > \frac{\epsilon}{2} n \right) \\
&< \exp\left(-\frac{\epsilon^2}{10} n \right).
\end{aligned}$$

Thus

$$\mathbb{E}|N_\epsilon^B(v) \le \mathbb{E}|S(v)| < \exp\left(-\frac{\epsilon^2}{20} n \right) (\alpha+\beta)^w (\beta+\gamma)^{n-w} = \exp\left(-\frac{\epsilon^2}{20} n \right) \mathbb{E}(\deg(v))\,,$$

and, consequently,

$$\mathbb{E}|N_\epsilon^G(v)| = \mathbb{E}(\deg(v)) - \mathbb{E}|N_\epsilon^B(v)| \geq \left(1 - \exp\left(-\frac{\epsilon^2}{20}n\right)\right)\mathbb{E}(\deg(v))\,.$$

■

Chapter 4

Perfect Matchings

In this chapter we study the existence of a perfect matching in $\mathcal{K}(n, \mathbf{P})$ and show that the threshold for the emergence of a perfect matching is the same as the connectivity threshold i.e. $\beta + \gamma = 1$. We also investigate the existence of a perfect matching on this threshold. The main result of this chapter, proved by the author in [1], can be stated as follows.

Theorem 4.1.

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

Observe that, by Theorem 2.2, there exists only one set of parameters, when $\beta = 1$ and $\alpha = \gamma = 0$, for which $\mathcal{K}(n, \mathbf{P})$ a.a.s. contains a perfect matching and at the same time is a.a.s. not connected; in fact it is easy to see that in this case with probability one $\mathcal{K}(n, \mathbf{P})$ is just a perfect matching. For any other choice of parameters α, β, γ, $\mathcal{K}(n, \mathbf{P})$ is either a.a.s. connected and contains a perfect matching, or is a.a.s. disconnected and contains no perfect matchings.

4.1 Proof

Let us start with the simplest case when $\beta + \gamma \le 1$ and $\beta \neq 1$. In [26] the authors showed that for this set of parameters $\mathcal{K}(n, \mathbf{P})$ a.a.s. contains an isolated vertex (it is also an immediate consequence of the first moment method and (3.1) applied to a set of all vertices of weight one). Consequently, in this case a.a.s. $\mathcal{K}(n, \mathbf{P})$ is disconnected and contains no perfect matchings.

Let $\beta = 1$. For a vertex $v = (v_1, v_2, \dots, v_n)$ of $\mathcal{K}(n, \mathbf{P})$ we put $\bar{v} = (1-v_1, 1-v_2, \dots, 1-v_n)$. The probability that v and $\bar{v}$ are connected by an edge is

$$p_{v,\bar{v}} = \beta^n = 1.$$

Thus, with probability 1, $\mathcal{K}(n, \mathbf{P})$ contains a perfect matching.

Now we consider the last, most interesting case, when $\beta + \gamma > 1$. For a given odd number $t \in [n]$, let H_t denote a graph with a vertex set $\{0,1\}^n$, where we connect by an edge all pairs of vertices which lie at Hamming distance t. Notice that H_t is a bipartite graph. Indeed, denote by $V_1(H_t)$ and $V_2(H_t)$ the subsets of vertices of H_t of odd and even weights respectively. Since t is odd, all edges of H_t have one end in $V_1(H_t)$ and other in $V_2(H_t)$.

Denote by $\mathcal{H}_t = \mathcal{H}_t(n, \mathbf{P})$ the subgraph of $\mathcal{K}(n, \mathbf{P})$, which contains only those edges of $\mathcal{K}(n, \mathbf{P})$ which join two vertices of Hamming distance t. $\mathcal{H}_t$ is a random subgraph of H_t so it is bipartite. Let us set

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

Note that t is an odd integer close to $\frac{\beta}{\beta+\gamma} n$.

In order to simplify the notation in this and the next chapter we put $H = H_t$, $\mathcal{H} = \mathcal{H}_t$, $V_1 = V_1(H)$, and $V_2 = V_2(H)$.

We are going to use expanding properties of the deterministic graph H to show that the Hall's condition is a.a.s. satisfied for sets V_1 and V_2 in $\mathcal{H}$, so a.a.s. $\mathcal{H}$ contains a perfect matching. We start with showing that Theorem 1.3 implies that H is a good expander.

Lemma 4.2. *Let T be a subset of the vertex set of H such that*

$$|T| \le |V|/2 = 2^{n-1}.$$

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

$$e_H(T, \overline{T}) \ge c|T|\binom{n}{t}\,.$$

Proof. Since we want to deduce Lemma 4.2 from Theorem 1.3, we show first that H is an edge-transitive graph of a diameter bounded by a constant which does not depend on n.

Clearly, for $i \in [n]$ the function $\tau_i : \{0,1\}^n \to \{0,1\}^n$ which maps $(v_1, \dots, v_i, \dots, v_n)$ into $(v_1, \dots, 1 - v_i, \dots, v_n)$ is an automorphism of H. Also, for any permutation $\sigma : [n] \to [n]$, the map $\mathrm{Aut}(\sigma) : \{0,1\}^n \to \{0,1\}^n$ which maps $(v_1, \dots, v_i, \dots, v_n)$ into $(v_{\sigma(1)}, \dots, v_{\sigma(i)}, \dots, v_{\sigma(n)})$ is an automorphism of H. We show that for every two edges e^1, e^2 of H there is a composition of two automorphisms of the above kinds which transforms e^1 into e^2. Although it is a rather easy observation, let us prove it 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 $u^i_j \ne v^i_j$. 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. 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, $\psi(x)_j = x_j$ if and only if $\phi^*(u^1)_j = u^2_j$, and it happens if and only if $\phi^*(v^1)_j = v^2_j$. Thus $\psi(\phi^*(v^1)) = v^2$ so $\psi \circ \phi^*$ is the desired automorphism of H which maps e^1 into e^2. Hence H is edge-transitive.

Now we bound the diameter of H. Let v, u be two vertices of H. Our aim is to show that they are connected by a short path. We split our argument into several cases.

Case 1. $\mathrm{d}(v, u)$ *is even and* $\mathrm{d}(v, u) \le \min\{2t, 2n - 2t\}$.

In this case there exists a vertex v' which is adjacent to both v and u. Indeed, to find v' it is enough to change v on $\mathrm{d}(v, u)/2$ positions on which v and u differ and $t - \mathrm{d}(v, u)/2$ positions on which they coincide (see Figure 4.1). Notice that v' is a neighbour of both v and u.

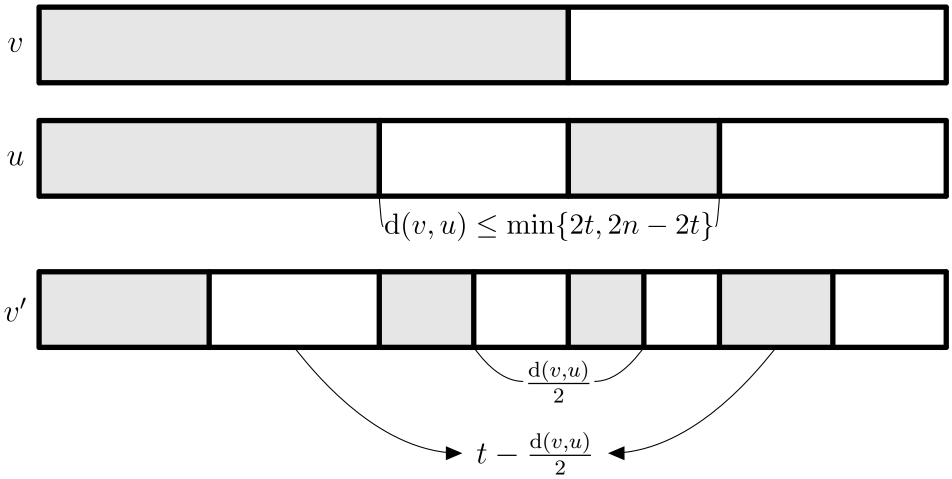

Figure 4.1: Neighbour v' of u and v

Case 2. $\mathrm{d}(v, u)$ *is even and* $\mathrm{d}(v, u) > 2t$ *(which is possible only if* $n > 2t$ *i.e.* $\gamma > \beta$*).*

For each pair of such vertices v and u there exists a vertex v' adjacent to v such that $\mathrm{d}(v', u) = \mathrm{d}(v, u) - t$. To get v' we need to change v on t positions on which v and u differ (see Figure 4.2).

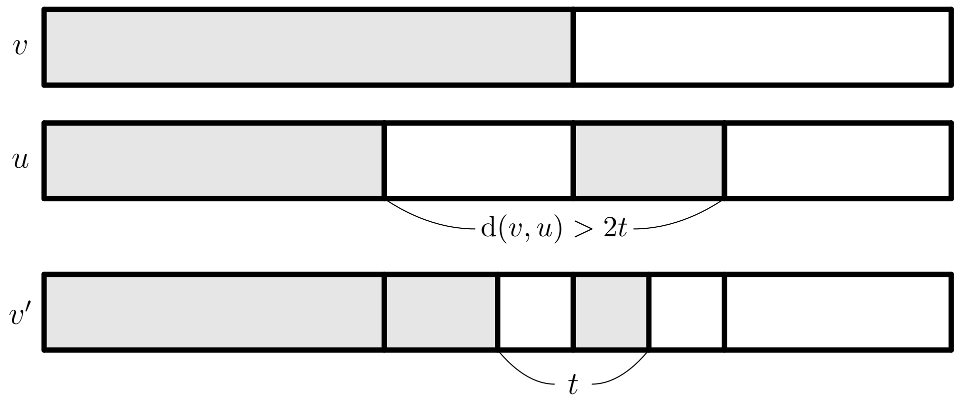

Figure 4.2: Vertex v' adjacent to v such that $\mathrm{d}(v', u) = \mathrm{d}(v, u) - t$

Applying this operation $2r$ times, where

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

one can construct a path $vv^1 \cdots v^{2r}$ in H such that for every $1 \leq i \leq 2r$, we have $\mathrm{d}(v^i, u) = \mathrm{d}(v^{i-1}, u) - t$ and $\mathrm{d}(v^{2r}, u) \leq 2t$. Notice that in this case $2t < n$, so $2t < 2n - 2t$ and thus $\mathrm{d}(v^{2r}, u) \leq \min\{2t, 2n - 2t\}$. As $\mathrm{d}(v^{2r}, u)$ is even, one can connect vertices v^{2r} and u by a path of length two using the argument from Case 1. Thus v, u are connected by a path of length at most

$$\frac{\beta + \gamma}{\beta} + 2\,.$$

Case 3. $\mathrm{d}(v, u)$ *is even and* $2n - 2t < \mathrm{d}(v, u) \leq 2t$ *(which is possible only if* $n < t$ *i.e.* $\beta > \gamma$*).*

For each such v and u there exist a path vv^1v^2 in H such that $\mathrm{d}(v^2, u) = d(v, u) - 2(n - t)$. To obtain v^1 from v, we need to change all $n - \mathrm{d}(v, u)$ positions on which v, u do not differ and $t - n + \mathrm{d}(v, u)$ positions among other places. Then, $\mathrm{d}(v, v^1) = n - \mathrm{d}(v, u) + t - n + \mathrm{d}(v, u) = t$.

To obtain v^2 from v^1, we need to change all $n - \mathrm{d}(v, u)$ positions on which v, u do not differ, all $n - t$ positions on which v and v^1 are the same and $2t - 2n + \mathrm{d}(v, u) > 0$ other positions (see Figure 4.3).

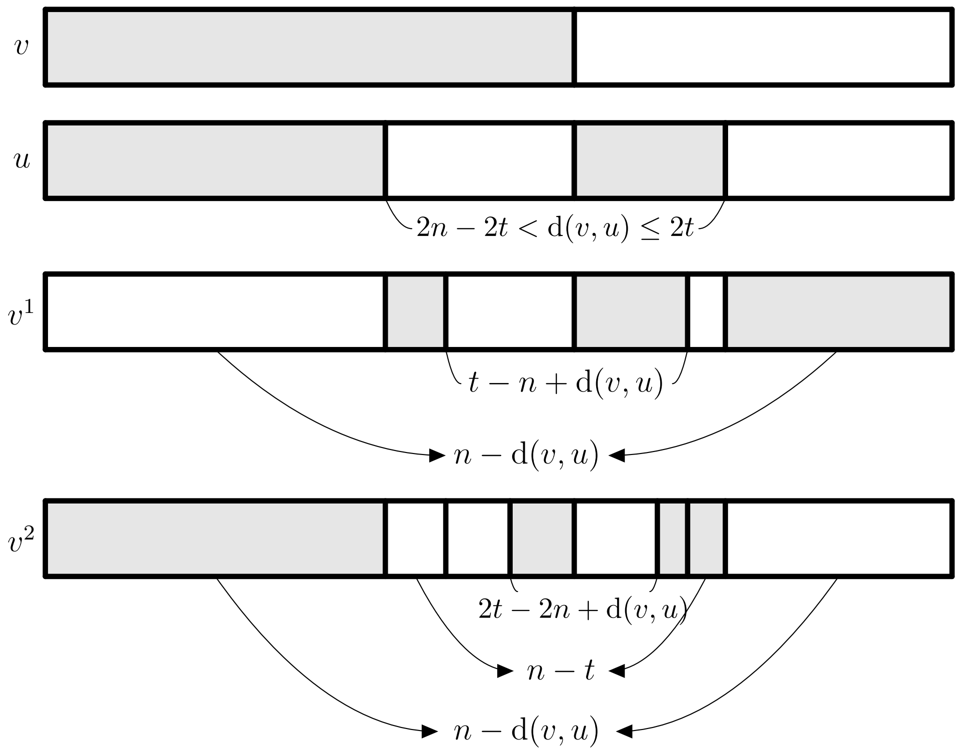

Figure 4.3: Vertices v^1 and v^2 such that $v \sim v^1 \sim v^2$ and $\mathrm{d}(v^2,u) = \mathrm{d}(v,u) - 2(n-t)$

Then, indeed $\mathrm{d}(v^1,v^2) = n - \mathrm{d}(v,u) + n - t + 2t - 2n + \mathrm{d}(v,u) = t$ and $\mathrm{d}(v^2,u) = \mathrm{d}(v,u) - 2(n-t)$.

Arguing in the same way we find a path $vv^1 \cdots v^{2s}$ where

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

such that for every $i \le s$, $\mathrm{d}(v^{2i},u) = \mathrm{d}(v^{2i-2},u) - 2(n-t)$ and $\mathrm{d}(v^{2s},u) \le 2n-2t$. Now, since $2n-2t < 2t$, $\mathrm{d}(v^{2s},u) \le \min\{2t, 2n-2t\}$, and $\mathrm{d}(v^{2s},u)$ is even, we can apply Case 1 to connect v^{2s} and u by a path of length two. Thus, v and u are connected by a path of length at most

$$\frac{2(\beta+\gamma)}{\gamma} + 2\,.$$

Case 4. $\mathrm{d}(v,u)$ *is odd.*

First we connect v to any neighbour v'. Since v' differs with v on t positions, $\mathrm{d}(v',u)$ is even and we can apply one of the already analyzed cases $1,2,3$ to v', u, to obtain a path of length at most

$$\max\left\{\frac{\beta+\gamma}{\beta}, \frac{2(\beta+\gamma)}{\gamma}\right\} + 2\,.$$

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

$$D \leq \max\left\{\frac{\beta+\gamma}{\beta}, \frac{2(\beta+\gamma)}{\gamma}\right\} + 3.$$

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

$$\mathrm{Vol}(T) = \binom{n}{t}|T| \leq \binom{n}{t}\frac{|V|}{2} = \frac{\mathrm{Vol}(V)}{2}.$$

By Theorem 1.3 we get

$$\frac{e_H(T,\overline{T})}{\mathrm{Vol}(T)} \geq \frac{1}{2D}.$$

Since D is bounded from above by some positive constant c, which depends only on β and γ, we have

$$e_H(T,\overline{T}) \geq \frac{1}{2D}\,\mathrm{Vol}(T) \geq c|T|\binom{n}{t}.$$

■

Now we go back to the random graph $\mathcal{H}$. Recall that $\mathcal{H}$ is a bipartite graph with a bipartition $V_1 \cup V_2$, $|V_1| = |V_2|$. Let $\mathcal{A}_i$, $i \in \{1,2\}$ be the event that there exists a subset $S \subseteq V_i$ such that $|N_{\mathcal{H}}(S)| = |S| - 1$ and $|S| \leq |V_1|/2 = 2^{n-2}$. By Theorem 1.2, if $\mathcal{H}$ does not contain a perfect matching, either $\mathcal{A}_1$ or $\mathcal{A}_2$ occurs. Since $\mathbb{P}(\mathcal{A}_1) = \mathbb{P}(\mathcal{A}_2)$,

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

For two fixed sets $S_1 \subseteq V_1$, $|S_1| \leq 2^{n-2} = N/4$ and $S_2 \subseteq V_2$, $|S_2| = |S_1| - 1$, let $\mathcal{A}_{S_1,S_2}$ denote the event that $S_2 = N_{\mathcal{H}}(S_1)$. Clearly

$$\mathbb{P}\left(\mathcal{A}_1\right) \leq \sum_{\substack{S_1 \subseteq V_1 \\ |S_1| \leq N/4}} \sum_{\substack{S_2 \subseteq V_2 \\ |S_2| = |S_1|-1}} \mathbb{P}\left(\mathcal{A}_{S_1,S_2}\right).$$

Now we bound from below the number of edges between S_1 and $V_2 \setminus S_2$ in the deterministic graph H. Note that if $S_2 = N_{\mathcal{H}}(S_1)$, these are precisely the edges which occur in H and do not occur in $\mathcal{H}$ (see Figure 4.4).

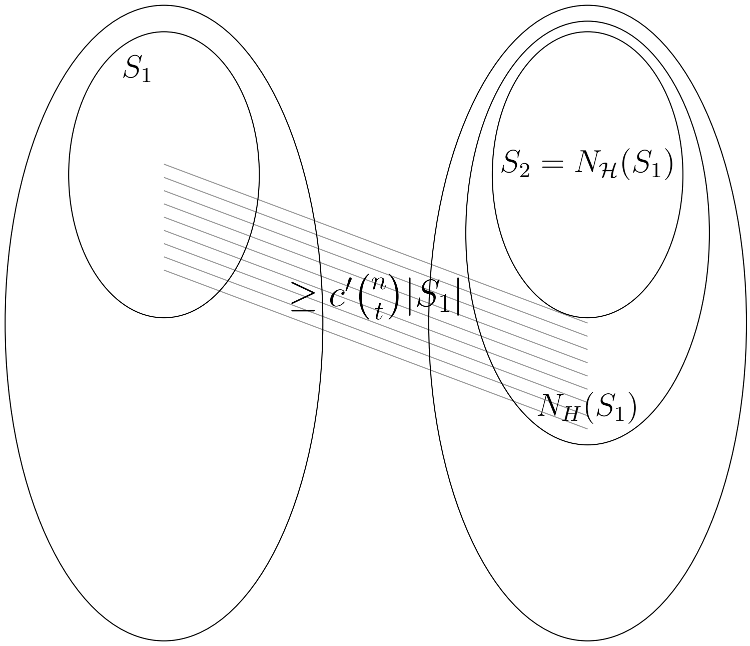

Figure 4.4: Edges in H between S_1 and $V_2 \setminus S_2$

We apply Lemma 4.2 to the set $T = S_1 \cup S_2$. Clearly $|T| = 2|S_1| - 1 < N/2$. We have

$$e_H(S_1, V_2 \setminus S_2) + e_H(S_2, V_1 \setminus S_1) = e_H(T, \overline{T}) \geq c\binom{n}{t}|T| = c\binom{n}{t}(2|S_1| - 1)\,, \qquad (4.1)$$

while from the regularity of H we get

$$e_H(S_1, \overline{S_1}) = e_H(S_1, V_2 \setminus S_2) + e_H(S_1, S_2) = \binom{n}{t}|S_1|\,, \qquad (4.2)$$

and

$$e_H(S_2, \overline{S_2}) = e_H(S_2, V_1 \setminus S_1) + e_H(S_1, S_2) = \binom{n}{t}|S_2| = \binom{n}{t}(|S_1| - 1)\,. \qquad (4.3)$$

Adding (4.1) and (4.2) and subtracting (4.3), we obtain that in H,

$$e_H(S_1, V_2 \setminus S_2) \geq \frac{1}{2}\binom{n}{t}\left(|S_1| + 2c|S_1| - c - |S_1| + 1\right) \geq c'\binom{n}{t}|S_1|\,,$$

for some positive constant c'. Thus if $\mathcal{A}_{S_1,S_2}$ occurs, $c'\binom{n}{t}|S_1|$ fixed pairs of vertices which are adjacent in H are not adjacent in $\mathcal{H}$.

Observe that for each pair u, v of vertices of Hamming distance t, the probability that there exists an edge $\{u, v\}$ is at least $\beta^t\gamma^{n-t}$. Thus, the probability of the event $\mathcal{A}_1$ is bounded from above by

$$\begin{aligned}
\mathbb{P}\left(\mathcal{A}_1\right) &\leq \sum_{\substack{S_1 \subseteq V_1 \\ |S_1| \leq N/4}} \sum_{\substack{S_2 \subseteq V_2 \\ |S_2| = |S_1|-1}} \mathbb{P}(\mathcal{A}_{S_1,S_2}) \\
&\leq \sum_{s=1}^{N/4} \binom{N/2}{s}\binom{N/2}{s-1}\left(1-\beta^t\gamma^{n-t}\right)^{c's\binom{n}{t}} \\
&\overset{(1.3)}{\leq} \sum_{s=1}^{N/4} N^{2s}\exp\left(-c's\binom{n}{t}\beta^t\gamma^{n-t}\right).
\end{aligned}$$

Since t is close to $\frac{\beta}{\beta+\gamma}n$, by (1.6) we have

$$\binom{n}{t}\beta^t\gamma^{n-t} \geq \frac{(\beta+\gamma)^n}{n^2}.$$

Thus

$$\mathbb{P}\left(\mathcal{A}_1\right) \leq \sum_{s=1}^{N/4}\left(2^{2n}\exp\left(-c'\frac{(\beta+\gamma)^n}{n^2}\right)\right)^s.$$

Since the term in brackets is exponentially small,

$$\mathbb{P}\left(\mathcal{A}_1\right) \leq \sum_{s=1}^{N/4}\left(2^{2n}\exp\left(-c'\frac{(\beta+\gamma)^n}{n^2}\right)\right)^s \leq 2^n 2^{2n}\exp\left(-c'\frac{(\beta+\gamma)^n}{n^2}\right) = o(1).$$

Consequently, with probability $1-o(1)$, $\mathcal{H}$ contains a perfect matching. Thus a.a.s. $\mathcal{K}(n,\mathbf{P})$, contains a perfect matching as well.

4.2 Property k-PM

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

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

respectively, we construct k edge-disjoint perfect matchings in $\mathcal{K}(n,\mathbf{P})$, where each is present in $\mathcal{K}(n,\mathbf{P})$ independently with probability $1-o(1)$.

Thus, let k-PM denote the property that a graph contains k edge-disjoint perfect matchings. Then as an immediate consequence of Theorem 4.1 we get the following result.

Theorem 4.3. *Let $k \in \mathbb{N}$, $k \geq 2$ be a constant.*

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

In particular

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

■

Note the difference between the cases $k = 1$ and $k \geq 2$ 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.

Chapter 5

Edge Connectivity

In this chapter we generalise Theorem 2.2 which stated that, except for the special case when $\beta = 1$ and $\alpha = \gamma = 0$, a.a.s. $\mathcal{K}(n, \mathbf{P})$ is connected provided it contains no isolated vertices. Let us recall that $\delta(G)$ denotes the minimum degree of G, so $\delta(\mathcal{K}(n, \mathbf{P}))$ is a random variable which for some sets of parameters α, β, γ may quickly tend to infinity as $n \to \infty$.

Theorem 5.1. *For each set of parameters α, β, γ, such that either $\beta \neq 1$ or $\alpha + \gamma > 0$, the Kronecker graph $\mathcal{K}(n, \mathbf{P})$ is a.a.s. $\delta(\mathcal{K}(n, \mathbf{P}))$-edge connected.*

Proof. We need to prove that a.a.s. for every $S \subseteq V$, where $|S| \leq |V|/2$, we have

$$e(S, \overline{S}) \geq \delta = \delta(\mathcal{K}(n, \mathbf{P})). \tag{5.1}$$

Recall that by Theorem 2.2 $\mathcal{K}(n, \mathbf{P})$ is a.a.s. connected if and only if $\beta + \gamma > 1$ or $\beta = 1$, $\alpha > 0$ and $\gamma = 0$. Otherwise it is a.a.s. not connected. Assume first that $\beta = 1$, $\alpha > 0$, $\gamma = 0$. Then the vertex $(0, 0, \ldots, 0)$ is with probability one connected to a vertex $(1, 1, \ldots, 1)$, and with probability 0 connected to any other vertex. Therefore the minimum degree in this graph is 1 with probability 1. Since due to Theorem 2.2 we know that in this case the graph is connected, the assertion holds for this set of parameters.

Thus, from now on, we assume $\beta + \gamma > 1$.

Let S be a subset of vertices such that $|S| < \delta$. Then

$$e(S, \overline{S}) + 2e(S) = \sum_{v \in S} \deg(v) \geq \delta |S|.$$

However, clearly $e(S) \leq \binom{|S|}{2}$, and so

$$e(S, \overline{S}) \geq \delta |S| - 2e(S) \geq \delta |S| - |S|^2 + |S| \geq \delta,$$

where the last inequality holds since $1 \leq |S| \leq \delta$. Therefore every set S of size smaller than δ fulfills the condition (5.1).

Now let S be a vertex subset such that $\delta \leq |S| \leq |V|/2$. Let v be a vertex of $\mathcal{K}(n, \mathbf{P})$. Then by (3.1)

$$\mathbb{E}(\deg(v)) = (\alpha+\beta)^{\mathrm{w}(v)}(\beta+\gamma)^{n-\mathrm{w}(v)}.$$

As $\alpha \geq \gamma$, $\mathbb{E}\left(\deg(v)\right) \geq (\beta+\gamma)^n$. By Chernoff's inequality (1.2), $\deg(v) \leq \frac{1}{2}(\beta+\gamma)^n$ with probability at most $\exp\left(-(\beta+\gamma)^n/8\right)$. Thus the expected number of vertices in $\mathcal{K}(n, \mathbf{P})$ with degree smaller than $\frac{1}{2}(\beta+\gamma)^n$ is at most

$$2^n \exp\left(-\frac{(\beta+\gamma)^n}{8}\right) \leq \exp\left(-\frac{(\beta+\gamma)^n}{10}\right).$$

From the first moment method (Theorem 1.5) we infer that with probability at least $1-\exp\left(-(\beta+\gamma)^n/10\right)$,

$$\delta \geq \frac{1}{2}(\beta+\gamma)^n.$$

On the other hand, take vertex $v^0 = (0, 0, \ldots, 0)$. Its weight is 0, hence

$$\mathbb{E}\left(\deg(v^0)\right) = (\beta+\gamma)^n.$$

Again, by Chernoff's inequality (1.2)

$$\deg(v^0) \leq 2\mathbb{E}\left(\deg(v^0)\right)$$

with probability at least $1-\exp\left(-3(\beta+\gamma)^n/8\right)$. Consequently, with probability at least

$$1-\exp\left(\frac{(\beta+\gamma)^n}{10}\right) - \exp\left(-\frac{3(\beta+\gamma)^n}{8}\right) > 1-\exp\left(\frac{(\beta+\gamma)^n}{20}\right),$$

we have

$$\delta \in \left[\frac{1}{2}(\beta+\gamma)^n, 2(\beta+\gamma)^n\right]. \tag{5.2}$$

Now let t and H be defined as in the proof of Theorem 4.1. Then, by Lemma 4.2,

$$e_H(S, \overline{S}) \geq |S| c \binom{n}{t},$$

for some positive constant c. Since every edge of H is present in $\mathcal{K}(n, \mathbf{P})$ with probability at least $\beta^t\gamma^{n-t}$,

$$\mathbb{E}\left(e(S, \overline{S})\right) \geq |S| c \binom{n}{t} \beta^t \gamma^{n-t} \overset{(1.6)}{\geq} (1-o(1))\,|S| \frac{c(\beta+\gamma)^n}{n^2} \geq \frac{|S| c(\beta+\gamma)^n}{n^3}.$$

By Chernoff's inequality (1.2)

$$e(S, \overline{S}) \geq \frac{1}{2}\mathbb{E}\left(e(S, \overline{S})\right) \geq \frac{|S| c(\beta+\gamma)^n}{2n^3} \geq \frac{c\delta(\beta+\gamma)^n}{2n^3} \geq \delta$$

with probability at least

$$1 - \exp\left(-\frac{\mathbb{E}\left(e(S,\overline{S})\right)}{8}\right) \geq 1 - \exp\left(-\frac{|S|c(\beta+\gamma)^n}{8n^3}\right).$$

Therefore the probability that there exists a subset S, $|S| > \delta$, for which the condition (5.1) does not hold is at most

$$\sum_{|S|=\delta}^{2^n} \binom{2^n}{|S|} \exp\left(-\frac{|S|c(\beta+\gamma)^n}{8n^3}\right) \leq \sum_{|S|=\delta}^{2^n} 2^{n|S|} \exp\left(-\frac{|S|c(\beta+\gamma)^n}{8n^3}\right)$$
$$\leq \sum_{|S|=\delta}^{2^n} \left(\exp\left(n - \frac{c(\beta+\gamma)^n}{8n^3}\right)\right)^{|S|} \leq 2^n \left(\exp\left(n - \frac{c(\beta+\gamma)^n}{8n^3}\right)\right)^{\delta} = o(1)\,.$$

Thus a.a.s. the condition (5.1) holds for every vertex subset S of size $|S| \leq |V|/2$ and Theorem 5.1 follows.

■

Chapter 6

Diameter

The main goal of this chapter is to prove the following result of Banaszak and Łuczak [2] which states that the diameter of $\mathcal{K}(n, \mathbf{P})$ is a.a.s. bounded from above by a constant for any set of parameters for which $\mathcal{K}(n, \mathbf{P})$ is a.a.s. connected.

Theorem 6.1. *Let α, β, and γ be a set of parameters for which $\mathcal{K}(n, \mathbf{P})$ is a.a.s. connected. Then there exists a constant $a = a(\alpha, \beta, \gamma)$ such that a.a.s.* $\operatorname{diam}(\mathcal{K}(n, \mathbf{P})) \leq a$.

Recall that by Theorem 2.2, for every set of parameters $\mathcal{K}(n, \mathbf{P})$ is either a.a.s. connected or a.a.s. not connected. Thus, since the diameter of a disconnected graph is infinite, the result above gives a full characterization of the diameter of $\mathcal{K}(n, \mathbf{P})$.

The diameter of $\mathcal{K}(n, \mathbf{P})$ has been studied by Mahdian and Xu [25] under the condition $\gamma \leq \beta \leq \alpha$ using the standard approach for binomial random graphs. This approach was possible in that case, because if u, v are two vertices of weights $\mathrm{w}(u) > \mathrm{w}(v)$, the probability that there exists an edge uv is at least

$$\beta^{\mathrm{w}(u)} \gamma^{n - \mathrm{w}(u)} .$$

Taking the set S of vertices with weight at least $\frac{\beta}{\beta+\gamma} n$ the authors have shown that each vertex outside S has a neighbour in S. Then the fact that S has the diameter bounded by a constant was deduced from corresponding results of Klee and Larman [23] for the graph $G\left(n, p\right)$ with $n = |S|$ and $p = \beta^{\frac{\beta}{\beta+\gamma} n} \gamma^{\frac{\gamma}{\beta+\gamma} n}$. This approach fails in our case, as we cannot satisfactorily bound from below the probability of an edge between two vertices, knowing their weights.

Let us also recall how the diameter of random graphs was bounded for the binomial model of a random graph $G(n, p)$ in [4], [6], and [7]. One can show that for some small k the k-th neighbourhood of every vertex in $G(n, p)$ is a.a.s. much larger than $\sqrt{n}$ and as for every two vertices u, v their k-th neighbourhoods are either independent or intersecting, a.a.s. they are not disjoint. For our model this procedure is impossible due to the fact that we do not understand well the expanding properties of $\mathcal{K}(n, \mathbf{P})$ which are easy to investigate in most of the other random graph models. Moreover, for a vertex v in $\mathcal{K}(n, \mathbf{P})$ and two vertices u, u' in $N(v)$, the neighbourhoods of u, u' are strongly correlated. Thus, we cannot bound from below the size of the k-th neighbourhood of v and, furthermore, the k-th neighbourhood of v is far from being the random subset of the n-cube.

Thus, we apply a different approach. For the set of vertices with weight other than $n/2$ we use the results from Chapter 3 to show that every such vertex is connected by a short path to the middle layer, i.e. to the set of vertices with weight $n/2$. Next, inside the middle layer, we pick two vertices which lie within small Hamming distance and generate their neighbourhoods at the same time until, for some k, we observe that the k-th neighbourhood of v does not expand satisfactorily. This can happen if most of the neighbours of vertices in k-th neighbourhood of v are already in the k-th neighbourhood of v. The probability of such event is however roughly the same as the probability that they are in k-th neighbourhood of u, since u, v are close to each other and so their k-neighbourhoods are similar. Thus, a.a.s. there exists a path of length at most $2k$ between u and v. Although the main idea of our argument seems to be simple, its rigorous implementation, presented in Section 6.2.2, is rather technical and complicated.

6.1 At the connectivity threshold

First, let us consider the case on the connectivity threshold, when $\mathcal{K}(n, \mathbf{P})$ is a.a.s. connected, i.e. when $\alpha > 0$, $\beta = 1$, $\gamma = 0$. For a vertex $v = (v_1, v_2, \ldots, v_n)$, denote $\bar{v} = (1 - v_1, 1 - v_2, \ldots, 1 - v_n)$.

Notice that

$$p_{v,\bar{v}} = \beta^n = 1\,,$$

and that $\mathrm{w}(\bar{v}) = n - \mathrm{w}(v)$, so either v or $\bar{v}$ has weight at least $n/2$. Denote by $\mathcal{K}_{upper}$ the subgraph of $\mathcal{K}(n, \mathbf{P})$ induced on the vertex set

$$\{v \in V :\ \mathrm{w}(v) \geq n/2\}\,.$$

Let v be a vertex in $\mathcal{K}_{upper}$ other than $(1, 1, \ldots, 1)$. We show that there exists a short path joining v and $(1, 1, \ldots, 1)$ in $\mathcal{K}(n, \mathbf{P})$ and, consequently, there is a short path between any pair of vertices from $\mathcal{K}(n, \mathbf{P})$.

Let $\eta \in (0, 1)$ denote the largest solution of the equality

$$\alpha^{\eta}(\alpha^2 + 1)^{1/2} = 1 + \eta\,.$$

Such a solution exists, as both sides of the equation are continuous functions of η, for $\eta = 0$ the left hand side is greater than the right hand side, and for $\eta = 1$ the right hand side is greater than the left hand side of the equation. Let $t < \eta n$, $t < n - \mathrm{w}(v)$ be fixed and let u be a vertex with ones on all the coordinates where v has ones, and with t ones on the positions where v has zeros. Then u is a vertex of $\mathcal{K}_{upper}$. We show that v and u have a.a.s. a common neighbour in $\mathcal{K}(n, \mathbf{P})$. Let $x \in V$ have ones on all the coordinates, where v has zeros, and $j < \mathrm{w}(v)$ more ones on the positions where v has ones (see Figure 6.1).

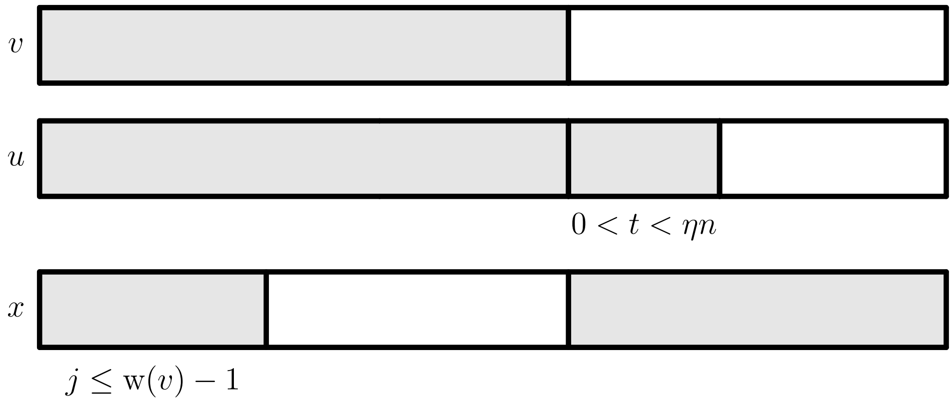

Figure 6.1: Vertex x in $\mathcal{K}(n, \mathbf{P})$

Then

$$\mathbb{P}(x \sim v, x \sim u) = p_{x,v} p_{x,u} = \alpha^j \beta^{n-j} \alpha^{j+t} \beta^{n-j-t} = \alpha^{2j+t} .$$

Hence,

$$\begin{aligned}
\mathbb{P}(v, u \text{ have no common neighbours in } \mathcal{K}(n, \mathbf{P})) &\leq \prod_{j=1}^{\mathrm{w}(v)-1} \left(1 - \alpha^{2j+t}\right)^{\binom{\mathrm{w}(v)}{j}} \\
&\overset{(1.3)}{\leq} \exp\left(- \sum_{j=1}^{\mathrm{w}(v)-1} \binom{\mathrm{w}(v)}{j} \alpha^{2j+t}\right) \leq \exp\left(-\alpha^t \left((\alpha^2+1)^{\mathrm{w}(v)} - \alpha^{2\,\mathrm{w}(v)} - 1\right)\right) \\
&\leq \exp\left(-\frac{\left(\alpha^\eta (\alpha^2+1)^{1/2}\right)^n}{2}\right) = \exp\left(-\frac{1}{2}(1+\eta)^n\right) = o(2^{-2n}) ,
\end{aligned}$$

Thus the expected number of such pairs (v, u) in $\mathcal{K}_{upper}$ that do not have a common neighbour is $o(1)$. By Markov's inequality (Theorem 1.4), a.a.s. each such pair $u, v \in \mathcal{K}_{upper}$ has a common neighbour. One can construct a path $vv^1 \ldots v^{2r}$ such that $r < \frac{n-\mathrm{w}(v)}{\eta n} < 1/(2\eta)$ and $v^{2r} = (1, 1, \ldots, 1)$. Hence every vertex in $\mathcal{K}_{upper}$ is a.a.s. connected to the vertex $(1, 1, \ldots, 1)$ by a path of length at most $1/\eta$. Consequently, a.a.s. each pair of vertices of $\mathcal{K}(n, \mathbf{P})$ is joined by a path of length at most $2/\eta + 2$, where η is a constant.

6.2 Above the connectivity threshold

Let $\beta + \gamma > 1$. As if we increase α, the diameter of $\mathcal{K}(n, \mathbf{P})$ can only decrease, from now on we assume $\alpha = \gamma$.

Let $\mathcal{K}_{mid}$ denote the subgraph of $\mathcal{K}(n, \mathbf{P})$ induced by the middle layer, i.e. the set of vertices with weight $n/2$. First we argue that every vertex of $\mathcal{K}(n, \mathbf{P})$ outside $\mathcal{K}_{mid}$ is connected to $\mathcal{K}_{mid}$ by a short path. In the second part of the proof we show that $\mathcal{K}_{mid}$ has a constant diameter.

6.2.1 Outside the middle layer

The main goal of this section is to prove the following result.

Theorem 6.2. *Let $\alpha = \gamma$ and $\beta + \gamma > 1$. There exists a constant $c = c(\alpha, \beta)$ such that a.a.s. every vertex v of $\mathcal{K}(n, \mathbf{P})$ is connected by a path of length at most c to the graph $\mathcal{K}_{mid}$.*

PROOF. Our method is somewhat similar to that used by Horn and Radcliffe [20] to show that, under some conditions, $\mathcal{K}(n, \mathbf{P})$ contains a giant component. In this paper the authors first proved that a layer of vertices with weight $\frac{\alpha+\beta}{\alpha+2\beta+\gamma}n$ is connected and then argued that many vertices are connected to this layer by a path in which subsequent vertices have weights closer to $\frac{\alpha+\beta}{\alpha+2\beta+\gamma}n$ than the previous ones.

Our argument is based on the following three observations.

Lemma 6.3. *Let $\alpha = \gamma$ and $\beta + \gamma > 1$. Then a.a.s. each vertex v of $\mathcal{K}(n, \mathbf{P})$ of weight $\mathrm{w}(v) \neq n/2$ has a good neighbour, i.e. a neighbour of weight*

$$\widetilde{\mathrm{w}}(v) = \frac{\alpha}{\alpha+\beta}\,\mathrm{w}(v) + \frac{\beta}{\alpha+\beta}\,(n - \mathrm{w}(v)) = \frac{n}{2} + \frac{\alpha-\beta}{\alpha+\beta}\left(\mathrm{w}(v) - \frac{n}{2}\right).$$

PROOF. By Lemma 3.1 the expected number of neighbours of v of weight $\widetilde{\mathrm{w}}(v)$ is at least

$$(1 - o(1))\frac{(\alpha+\beta)^n}{n^2} \geq \frac{(\alpha+\beta)^n}{n^3}.$$

By Chernoff's inequality (1.2) with probability at most $\exp\left(-\frac{(\alpha+\beta)^n}{8n^3}\right)$, the vertex v has fewer than $\frac{(\alpha+\beta)^n}{2n^3}$ neighbours of weight $\widetilde{\mathrm{w}}(v)$. Thus, the expected number of vertices v outside $\mathcal{K}_{mid}$ with no neighbours of weight $\widetilde{\mathrm{w}}(v)$ is at most

$$|V(\mathcal{K}(n, \mathbf{P}))| \exp\left(-\frac{(\alpha+\beta)^n}{8n^3}\right) < 2^n \exp(-n^2) = o(1).$$

Consequently, by the first moment method (Theorem 1.5) a.a.s. every vertex v with weight $\mathrm{w}(v) \neq n/2$ has a neighbour of weight

$$\widetilde{\mathrm{w}}(v) = \frac{n}{2} + \frac{\alpha-\beta}{\alpha+\beta}\left(\mathrm{w}(v) - \frac{n}{2}\right).$$

■

Lemma 6.4. *Let $\alpha = \gamma$ and $\beta + \gamma > 1$. Let $\zeta > 0$ be a constant and denote*

$$b = \begin{cases} 1 & \text{if } \alpha = \beta \\ \log_{\left|\frac{\alpha-\beta}{\alpha+\beta}\right|}(\zeta) & \text{if } \alpha \neq \beta. \end{cases}$$

Then a.a.s. every vertex v of $\mathcal{K}(n, \mathbf{P})$ such that $|\,\mathrm{w}(v) - n/2| > \zeta n/2$ is connected by a path of length at most b to a vertex u such that

$$\left|\mathrm{w}(u) - \frac{n}{2}\right| \leq \zeta \frac{n}{2}\,.$$

Proof. By Lemma 6.3, a.a.s. for each vertex v there exists a path

$$v = v^0 v^1 v^2 \ldots v^b\,,$$

such that

$$\mathrm{w}(v^i) = \widetilde{\mathrm{w}}(v^{i-1}) = \frac{n}{2} + \frac{\alpha-\beta}{\alpha+\beta}\left(\mathrm{w}(v^{i-1}) - \frac{n}{2}\right)\,,$$

for $i \in [b]$. Solving the above recurrence we get

$$\mathrm{w}(v^b) = \frac{n}{2} + \left(\frac{\alpha-\beta}{\alpha+\beta}\right)^b \left(\mathrm{w}(v) - \frac{n}{2}\right)\,.$$

Thus

$$\left|\mathrm{w}(v^b) - \frac{n}{2}\right| \leq \zeta \left|\mathrm{w}(v) - \frac{n}{2}\right| \leq \zeta \frac{n}{2}\,.$$

■

Lemma 6.5. *Let $\alpha = \gamma$ and $\beta + \gamma > 1$. Moreover, let $\zeta > 0$ be such a constant that*

$$(\alpha+\beta)^{1-\zeta} \left(\min\{\alpha, \beta\}\right)^{\zeta} > 1\,. \tag{6.1}$$

Then a.a.s. every vertex v of $\mathcal{K}(n, \mathbf{P})$ of weight $\mathrm{w}(v)$ such that $|\,\mathrm{w}(v) - n/2| \leq \zeta n/2$ has a neighbour of weight $n/2$.

Proof. Let v be such a vertex that $|\,\mathrm{w}(v) - n/2| \leq \zeta n/2$.

Let us choose $(1-\zeta)n/2$ ones and $(1-\zeta)n/2$ zeros in the label of v. Let $A(v)$ denote the set of those vertices which have precisely $\frac{\alpha}{\alpha+\beta}(1-\zeta)n/2$ ones among the chosen one positions in the label of v, $\frac{\alpha}{\alpha+\beta}(1-\zeta)n/2$ zeros among the chosen zero-positions in the label of v, ones on half of the remaining one positions in the label of v and zeros on half of the remaining zero positions in the label of v (see Figure 6.2).

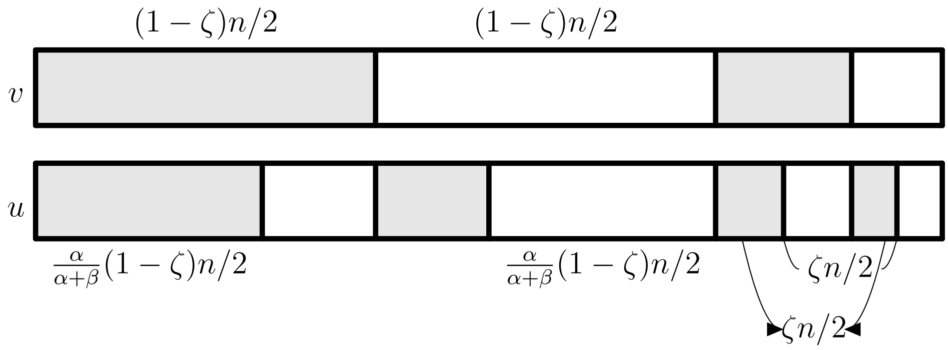

Figure 6.2: Vertex u from $A(v)$

Then for every $u \in A(v)$,

$$\mathrm{w}(u) = \frac{\alpha}{\alpha+\beta}(1-\zeta)\frac{n}{2} + \frac{\beta}{\alpha+\beta}(1-\zeta)\frac{n}{2} + \zeta\frac{n}{2} = \frac{n}{2}\,.$$

Thus all the vertices of $A(v)$ lie in the middle layer. Moreover

$$|A(v)| > \binom{(1-\zeta)n/2}{\frac{\alpha}{\alpha+\beta}(1-\zeta)n/2}^2,$$

and for every $u \in A(v)$ the probability that $u \sim v$ is

$$p_{u,v} \geq \alpha^{\frac{\alpha}{\alpha+\beta}(1-\zeta)n}\beta^{\frac{\beta}{\alpha+\beta}(1-\zeta)n}\left(\min\{\alpha,\beta\}\right)^{\zeta n}\,.$$

Now let $N_{mid}(v)$ denote the neighbourhood of v in the middle layer. Then we have

$$\mathbb{E}|N_{mid}(v)| \geq \mathbb{E}|A(v) \cap N(v)| > \binom{(1-\zeta)n/2}{\frac{\alpha}{\alpha+\beta}(1-\zeta)n/2}^2 \alpha^{\frac{\alpha}{\alpha+\beta}(1-\zeta)n}\beta^{\frac{\beta}{\alpha+\beta}(1-\zeta)n}\left(\min\{\alpha,\beta\}\right)^{\zeta n}$$

$$\overset{(1.5)}{\geq} \frac{\left((\alpha+\beta)^{1-\zeta}\left(\min\{\alpha,\beta\}\right)^{\zeta}\right)^n}{n^5}\,.$$

Since by (6.1) $(\alpha+\beta)^{1-\zeta}\left(\min\{\alpha,\beta\}\right)^{\zeta} > 1$,

$$\mathbb{E}|N_{mid}(v)| \geq \tau^n,$$

for some constant $\tau > 1$. By Chernoff's inequality (1.2) with probability at least $1 - \exp\left(-\mathbb{E}|N_{mid}(v)|/8\right)$, v has a neighbour with weight $n/2$. Thus, the expected number of vertices v of weight $\mathrm{w}(v)$ such that

$$\left|\mathrm{w}(v) - \frac{n}{2}\right| \leq \zeta\frac{n}{2}\,,$$

and v has no neighbour in $\mathcal{K}_{mid}$ is at most

$$2^n \exp\left(-\frac{\mathbb{E}|N_{mid}(v)|}{8}\right) = o(1).$$

Hence by the first moment method, the probability that there exists such a vertex v is $o(1)$.

■

Note that there always exists a constant ζ which fulfills (6.1), as the function on the left hand side of the inequality is a continuous function of ζ greater than 1 for $\zeta = 0$. Let ζ be such a constant. By Lemma 6.4 a.a.s. every vertex in $\mathcal{K}(n, \mathbf{P})$ outside the middle layer is joint by a path of constant length to a vertex u with $|\,\mathrm{w}(u) - n/2| < \zeta n/2$, which is by Lemma 6.5 a.a.s. connected by an edge to the middle layer. That completes the proof of Theorem 6.2.

■

6.2.2 Middle layer

The main ingredient of the proof of Theorem 6.1 is the following result.

Theorem 6.6. *For every $\alpha = \gamma$ and $\beta + \gamma > 1$ there exists a constant $c' = c'(\alpha, \beta)$ such that a.a.s.*

$$\mathrm{diam}(\mathcal{K}_{mid}) < c' .$$

Let us first notice the following fact.

Lemma 6.7. *If $\mathcal{K}(2\bar{n}, \mathbf{P})$ has a.a.s. diameter bounded by a, then a.a.s.*

$$\mathrm{diam}\left(\mathcal{K}(2\bar{n} + 1, \mathbf{P})\right) < 2a + 1.$$

Proof. Assume $\mathcal{K}(2\bar{n}, \mathbf{P})$ has a.a.s. diameter bounded by a. We can split $\mathcal{K}(2\bar{n} + 1, \mathbf{P})$ into two disjoint subgraphs $\mathcal{K}_0(2\bar{n} + 1, \mathbf{P})$ and $\mathcal{K}_1(2\bar{n} + 1, \mathbf{P})$, induced respectively by vertex sets $V_0 = \{v \in V : \ v_{2n+1} = 0\}$ and $V_1 = \{v \in V : \ v_{2n+1} = 1\}$. Since a.a.s. $\mathrm{diam}(\mathcal{K}(2\bar{n}, \mathbf{P})) < a$, the diameters of $\mathcal{K}_0(2\bar{n} + 1, \mathbf{P})$ and $\mathcal{K}_1(2\bar{n} + 1, \mathbf{P})$ are also a.a.s. bounded by a. From the connectivity of $\mathcal{K}(2\bar{n} + 1, \mathbf{P})$ we infer that a.a.s. there always exists an edge from $\mathcal{K}_0(2\bar{n} + 1, \mathbf{P})$ to $\mathcal{K}_1(2\bar{n} + 1, \mathbf{P})$ and consequently the diameter of $\mathcal{K}(2\bar{n} + 1, \mathbf{P})$ is bounded by $2a + 1$.

■

Due to Lemma 6.7, it is enough to consider the case when n is even, and so we do to the end of this proof.

In order to show that the theorem holds, we show that a.a.s. for each pair of vertices v and u with a small Hamming distance between them, the expected number of edge-disjoint paths of bounded length is at least $n^2/4$ and we use Talagrand's inequality to infer that there a.a.s. exists such a path between them. Note, however, that although it is easy to bound the expected number of short paths between v and u, it is hard to bound the number of such edge-disjoint paths (which is crucial for Talangrand's theorem), as they are correlated. Thus, we label each edge of $\mathcal{K}_{mid}$ independently at random with one of n^2 labels - we split graph $\mathcal{K}_{mid}$ randomly into n^2 edge-disjoint graphs on vertex set $V(\mathcal{K}_{mid})$. Of course each of these n^2 graphs is the same random object, which can be obtained by deleting edges of $\mathcal{K}_{mid}$ with probability $1 - n^{-2}$. We denote it by $\widehat{\mathcal{K}_{mid}}$. We show that in each of them with some probability there exists a path between v and u. The following lemma is a crucial part of the proof.

Lemma 6.8. *For every $\alpha = \gamma$ and $\beta + \gamma > 1$, there exist constants $\epsilon > 0$ and $\hat{c} > 0$ such that for each pair v, u of vertices of $\widehat{\mathcal{K}_{mid}}$ such that $\mathrm{d}(v, u) < \epsilon n$ the probability that u and v are connected in $\widehat{\mathcal{K}_{mid}}$ by a path of length at most $2\hat{c}$ is at least $1/4$.*

The proof of the above lemma is technical and complicated and thus moved to the next subsection of this chapter.

PROOF OF THEOREM 6.6. Let $\epsilon > 0$ and $\hat{c} > 0$ be such constants for which the assertion of Lemma 6.8 holds. Let us split at random all edges of the middle layer of the n-cube into n^2 sets and on each of them we generate a random graph $\widehat{\mathcal{K}_{mid}}$. Equivalently, we may say of n^2 random graphs as of a random partition of edges of $\mathcal{K}_{mid}$ into n^2 random subgraphs. For any two vertices v, u of $\mathcal{K}_{mid}$ such that $\mathrm{d}(v, u) < \epsilon n$, denote by $X_{v,u}$ maximum number of edge-disjoint paths of length at most $2\hat{c}$ between v and u which lie entirely in one of n^2 graphs $\widehat{\mathcal{K}_{mid}}$. By Lemma 6.8 used for each of n^2 graphs $\widehat{\mathcal{K}_{mid}}$ we know that $\mathbb{E}(X_{v,u}) \geq n^2/4$. Now, let us consider $X_{v,u}$ as a function of $\binom{|V(\mathcal{K}_{mid})|}{2}$ indicator random variables, each of them representing a possible edge in $\mathcal{K}_{mid}$ and thus in one of $\widehat{\mathcal{K}_{mid}}$. Note that:

1) Adding or removing a single edge cannot change the value of $X_{v,u}$ by more than one.

2) If $X_{v,u} \geq m$, we can verify this fact by checking only $2\hat{c}m$ edges.

Thus, by Talagrand's inequality (Theorem 1.8) we get

$$\mathbb{P}\left(X_{v,u} < \frac{n^2}{8}\right) \leq 2\exp\left(-\frac{\mathbb{E}X_{v,u}}{32\hat{c}}\right) \leq 2\exp(-\mu n^2)\,,$$

for some constant $\mu = \mu(\epsilon, \hat{c}) > 0$, so the expected number of such pairs v, u for which $X_{v,u} < n^2/8$ is at most

$$2^{2n+1}\exp(-\mu n^2) = o(1)\,.$$

By the first moment method (Theorem 1.5) a.a.s. each pair of vertices v, u, of $\mathcal{K}_{mid}$ such that $\mathrm{d}(v, u) \leq \epsilon n$ is connected by a path of length at most $2\hat{c}$. To complete the proof it is enough to observe that for every pair of vertices v, v' of $\mathcal{K}_{mid}$ one can find a sequence of vertices

$$v = v^0, v^1, \ldots, v^r = v'$$

such that for $i \in [r]$, $\mathrm{d}(v^i, v^{i-1}) < \epsilon n$, and $r < 1/\epsilon$. Taking into account Lemma 6.7, we get

$$\mathrm{diam}(\mathcal{K}_{mid}) \leq \frac{4\hat{c}}{\epsilon} + 1\,.$$

■

6.2.3 Proof of Lemma 6.8

PROOF. Let u, v be two vertices from $\widehat{\mathcal{K}_{mid}}$ such that $\mathrm{d}(u, v) < \epsilon n$, where $\epsilon > 0$ satisfies the condition

$$(\alpha + \beta)^{1-\epsilon}(\min\{\alpha, \beta\})^{\epsilon} > 1\,.$$

Notice that such an ϵ exists, as the left hand side is a continuous function of ϵ which is greater than 1 for $\epsilon = 0$.

Our argument is based on the fact that the neighbourhoods of u and v are very similar. Thus, we first modify $\widehat{\mathcal{K}_{mid}}$ slightly to introduce a new random graph $\widetilde{\mathcal{K}_{mid}}$, which can be viewed as a subgraph of $\widehat{\mathcal{K}_{mid}}$, in which these neighbourhoods are basically indistinguishable.

Let $I \subseteq [n]$ denote the set of those positions on which v, u differ. Clearly $|I| = \mathrm{d}(v, u)$. The set of vertices of $\widetilde{\mathcal{K}_{mid}}$ is the set of those vertices which have $|I|/2$ ones inside I and $(n - |I|)/2$ ones outside of I (see Figure 6.3), i.e.

$$V(\widetilde{\mathcal{K}_{mid}}) = \left\{ x \in V(\mathcal{K}(n, \mathbf{P})) : \ |\{i \in I : x_i = 1\}| = \frac{|I|}{2},\ |\{i \in [n] \setminus I : x_i = 1\}| = \frac{n - |I|}{2} \right\}.$$

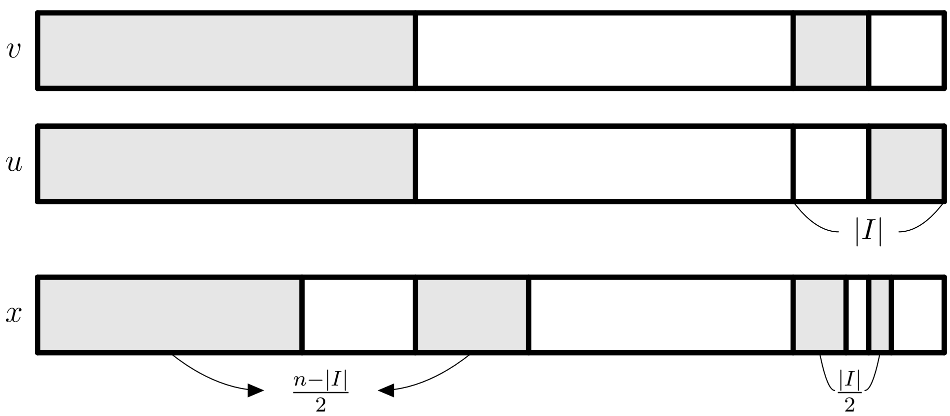

Figure 6.3: Vertex x from $V(\widetilde{\mathcal{K}_{mid}})$

Notice that each such vertex has weight $n/2$, so $V(\widetilde{\mathcal{K}_{mid}}) \subseteq V(\widehat{\mathcal{K}_{mid}})$. Furthermore let the edge set of $\widetilde{\mathcal{K}_{mid}}$ consist only of pairs xy which satisfy the condition

$$|\{i : x_i = y_i\} \setminus I| = \frac{\alpha}{\alpha + \beta}(n - |I|), \tag{6.2}$$

and we put each of them in $\widetilde{\mathcal{K}_{mid}}$ independently with probability

$$\rho = \rho(\alpha, \beta) = \alpha^{\frac{\alpha}{\alpha+\beta}(n-|I|)} \beta^{\frac{\beta}{\alpha+\beta}(n-|I|)} (\min\{\alpha, \beta\})^{|I|} n^{-2}. \tag{6.3}$$

Notice that each such edge is present in $\widehat{\mathcal{K}_{mid}}$ with probability at least ρ. Thus $\widetilde{\mathcal{K}_{mid}}$ can be viewed as a subgraph of $\widehat{\mathcal{K}_{mid}}$ on the vertex set defined above which we obtain from $\widehat{\mathcal{K}_{mid}}$ by keeping only the edges e which fulfill the condition (6.2), with probability ρ/ρ', where $\rho' > \rho$ is the probability of existence of e in $\widehat{\mathcal{K}_{mid}}$.

Now notice that for every pair of vertices x, y of $\widetilde{\mathcal{K}_{mid}}$ which differ only on I (such as v, u) and for every other vertex z of $\widetilde{\mathcal{K}_{mid}}$, the probability that $z \sim x$ is the same as the probability that $z \sim y$.

For a vertex x of $\widetilde{\mathcal{K}_{mid}}$ denote by $\widetilde{N}(x)$ the neighbourhood of x in $\widetilde{\mathcal{K}_{mid}}$ and, more generally, denote by $\widetilde{N}^i(x)$ the i-th neighbourhood of x in $\widetilde{\mathcal{K}_{mid}}$. Recall that by (3.1) the expected degree of every vertex in $\mathcal{K}(n, \mathbf{P})$ is exponential. We show that it is still exponential in $\widetilde{\mathcal{K}_{mid}}$.

Fact 6.9. *Let $\epsilon > 0$ be such a constant that*

$$(\alpha + \beta)^{1-\epsilon} (\min\{\alpha, \beta\})^{\epsilon} > 1\,,$$

and let $I \subseteq [n]$ be of size $|I| \le \epsilon n$. Then, there exists a constant $\xi > 1$ such that a.a.s. for every vertex x of $\widetilde{\mathcal{K}_{mid}}$

$$|\widetilde{N}(x)| \ge \xi^n.$$

Proof. Let x be a vertex of $\widetilde{\mathcal{K}_{mid}}$. Recall that y is a possible neighbour of x in $\widetilde{\mathcal{K}_{mid}}$ if it fulfills the condition (6.2) - see Figure 6.4 below.

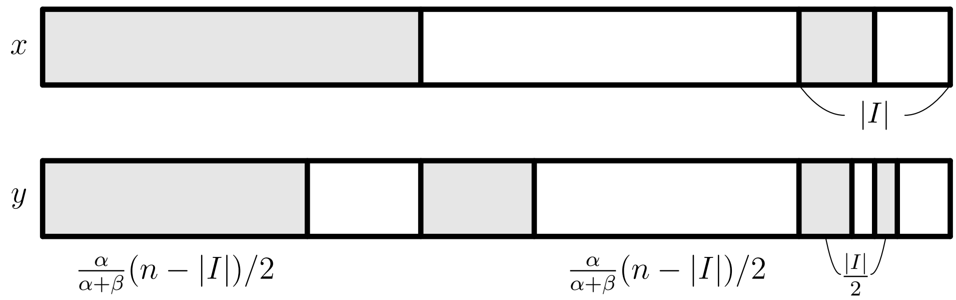

Figure 6.4: Possible neighbour y of x in $\widetilde{\mathcal{K}_{mid}}$

Thus the random variable $|\widetilde{N}(x)|$ has the binomial distribution $\mathrm{Bi}(m, \rho)$, where

$$m \ge \binom{(n-|I|)/2}{\frac{\alpha}{\alpha+\beta}(n-|I|)/2}^2 \binom{|I|}{|I|/2} \ge \binom{(n-|I|)/2}{\frac{\alpha}{\alpha+\beta}(n-|I|)/2}^2$$

and ρ is defined in (6.3). Hence

$$\mathbb{E}|\widetilde{N}(x)| \geq \binom{(n-|I|)/2}{\frac{\alpha}{\alpha+\beta}(n-|I|)/2}^2 \alpha^{\frac{\alpha}{\alpha+\beta}(n-|I|)} \beta^{\frac{\beta}{\alpha+\beta}(n-|I|)} (\min\{\alpha,\beta\})^{|I|} n^{-2}$$

$$\stackrel{(1.5)}{\geq} \frac{(\min\{\alpha,\beta\})^{|I|}(\alpha+\beta)^{(n-|I|)}}{n^5} \geq \frac{\left((\min\{\alpha,\beta\})^{\epsilon}(\alpha+\beta)^{(1-\epsilon)}\right)^n}{n^5}.$$

Thus for some constant $\xi > 1$ we have $\mathbb{E}|\widetilde{N}(x)| \geq 2\xi^n$. By Chernoff's inequality (1.2),

$$\mathbb{P}(|\widetilde{N}(x)| < \xi^n) \leq \exp\left(-\frac{\xi^n}{4}\right).$$

Consequently, the expected number of vertices in $\widetilde{\mathcal{K}_{mid}}$ with the neighbourhood of size smaller than ξ^n is at most

$$2^n \exp\left(-\frac{\xi^n}{4}\right) = o(1),$$

and by the first moment method (Theorem 1.5) the assertion holds.

■

Now, let $\hat{c} = 2\lceil \log_\xi 2 \rceil$, where ξ is a constant for which the above fact holds. We prove that

$$\mathbb{P}(\widetilde{N}^{\hat{c}}(v) \cap \widetilde{N}^{\hat{c}}(u) \neq \emptyset) \geq \frac{1}{4}.$$

Since the further part of the proof is quite technical, let us explain first its idea. As we know from Fact 6.9, the expected degree of a vertex in $\widetilde{\mathcal{K}_{mid}}$ is exponential, bounded from below by ξ^n, where $\xi > 1$. Since $\widetilde{\mathcal{K}_{mid}}$ has fewer than 2^n vertices, for some $k \leq 2\lceil \log_\xi 2 \rceil$, the k-th neighbourhood of v in $\widetilde{\mathcal{K}_{mid}}$ is smaller than $\xi^{kn/2}$. But, as we have shown, the degree of every vertex in $\widetilde{\mathcal{K}_{mid}}$ is at least ξ^n, which means there are pairs of vertices in the $(k-1)$-th neighbourhood of v which are connected. As in $\widetilde{\mathcal{K}_{mid}}$ the neighbourhoods of v and u are indistinguishable, there must exist edges between $(k-1)$-th neighbourhoods of u and v in $\widetilde{\mathcal{K}_{mid}}$.

Let us make the above heuristic argument rigorous. Let us split randomly the set $\widetilde{N}(v)$ into two sets: head neighbourhood $\widetilde{N}_{(H)}(v)$ and tail neighbourhood $\widetilde{N}_{(T)}(v)$ by tossing for each vertex in $\widetilde{N}(v)$ a symmetric coin. In the same way let us split randomly $\widetilde{N}(u)$ into $\widetilde{N}_{(H)}(u)$ and $\widetilde{N}_{(T)}(u)$. Denote by $\widetilde{N}^k_{-y}(x)$ the k-th neighbourhood of x in $\widetilde{\mathcal{K}_{mid}}$ without an edge xy if such edge exists. Observe that for every $x \in \widetilde{N}(v)$

$$\widetilde{N}^{\hat{c}-1}_{-v}(x) \subseteq \widetilde{N}^{\hat{c}}(v),$$

and the same holds for u, so

$$\mathbb{P}(\widetilde{N}^{\hat{c}}(v) \cap \widetilde{N}^{\hat{c}}(u) \neq \emptyset) \geq \mathbb{P}\left(\bigcup_{x \in \widetilde{N}_{(H)}(v)} \widetilde{N}^{\hat{c}-1}_{-v}(x) \cap \bigcup_{y \in \widetilde{N}_{(T)}(u)} \widetilde{N}^{\hat{c}-1}_{-u}(y) \neq \emptyset\right).$$

Assume

$$\mathbb{P}\left(\bigcup_{x\in\widetilde{N}_{(H)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(x) \cap \bigcup_{y\in\widetilde{N}_{(T)}(u)} \widetilde{N}_{-u}^{\hat{c}-1}(y) \neq \emptyset\right) < \frac{1}{4}\,, \tag{6.4}$$

to show that it leads to a contradiction.

Since the distribution of $\bigcup_{y\in\widetilde{N}_{(T)}(u)} \widetilde{N}_{-u}^{\hat{c}-1}(y)$ is identical with the distribution of $\bigcup_{y\in\widetilde{N}_{(T)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(y)$, we get

$$\begin{aligned}
&\mathbb{P}\left(\bigcup_{x\in\widetilde{N}_{(H)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(x) \cap \bigcup_{y\in\widetilde{N}_{(T)}(u)} \widetilde{N}_{-u}^{\hat{c}-1}(y) \neq \emptyset\right)\\
&\geq \mathbb{P}\left(\bigcup_{x\in\widetilde{N}_{(H)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(x) \cap \bigcup_{y\in\widetilde{N}_{(T)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(y) \neq \emptyset\right)\,,
\end{aligned}$$

where there is inequality instead of equality because it is possible that $\widetilde{N}_{(H)}(v)$ and $\widetilde{N}_{(T)}(u)$ intersect. Thus, by (6.4),

$$\mathbb{P}\left(\bigcup_{x\in\widetilde{N}_{(H)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(x) \cap \bigcup_{y\in\widetilde{N}_{(T)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(y) \neq \emptyset\right) < \frac{1}{4}\,.$$

Furthermore

$$\begin{aligned}
&\mathbb{P}\left(\bigcup_{x\in\widetilde{N}_{(H)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(x) \cap \bigcup_{y\in\widetilde{N}_{(T)}(v)} \widetilde{N}_{-v}^{\hat{c}-1}(y) \neq \emptyset\right)\\
&= \mathbb{P}\left(\exists_{x\neq y\in\widetilde{N}(v)} :\, x\in\widetilde{N}_{(H)}(v)\ \&\ y\in\widetilde{N}_{(T)}(v)\ \&\ \widetilde{N}_{-v}^{\hat{c}-1}(x)\cap\widetilde{N}_{-v}^{\hat{c}-1}(y)\neq\emptyset\right)\\
&= \frac{1}{2}\mathbb{P}\left(\exists_{x\neq y\in\widetilde{N}(v)} :\, \widetilde{N}_{-v}^{\hat{c}-1}(x)\cap\widetilde{N}_{-v}^{\hat{c}-1}(y)\neq\emptyset\right)\,,
\end{aligned}$$

so again by (6.4)

$$\mathbb{P}\left(\exists_{x\neq y\in\widetilde{N}(v)} :\, \widetilde{N}_{-v}^{\hat{c}-1}(x)\cap\widetilde{N}_{-v}^{\hat{c}-1}(y)\neq\emptyset\right) < \frac{1}{2}\,. \tag{6.5}$$

Denote by J a random variable which indicates a minimum i, for which i-th neighbourhood of v in $\widetilde{\mathcal{K}_{mid}}$ is smaller than expected, i.e.

$$J = \min\left\{i : |\widetilde{N}^i(v)| \leq \xi^{ni/2}\right\}\,.$$

Since $\xi^{\hat{c}n/2} \geq 2^n$, we have $J \in [\hat{c}]$. Let $j \in [\hat{c}]$ be such that $\mathbb{P}(J=j)$ is maximal. Clearly $\mathbb{P}(J=j) \geq 1/\hat{c}$. Hence

$$\mathbb{P}\left(|\widetilde{N}^{j-1}(v)| > \xi^{n(j-1)/2}\ \&\ |\widetilde{N}^{j}(v)| \leq \xi^{nj/2}\right) \geq \frac{1}{\hat{c}}\,. \tag{6.6}$$

Moreover, for every $x \neq y \in \widetilde{N}(v)$

$$\widetilde{N}^{j-1}(v) \subseteq \widetilde{N}^{j-1}_{-x}(v) \cup \widetilde{N}^{j-1}_{-y}(v)\,,$$

and $|\widetilde{N}^{j-1}_{-x}(v)|$, $|\widetilde{N}^{j-1}_{-y}(v)|$ have identical distribution. Thus

$$\begin{aligned}\mathbb{P}\left(|\widetilde{N}^{j-1}(v)| \geq \xi^{n(j-1)/2}\right) &\leq \mathbb{P}\left(|\widetilde{N}^{j-1}_{-x}(v)| \geq \frac{\xi^{n(j-1)/2}}{2}\right) + \mathbb{P}\left(|\widetilde{N}^{j-1}_{-y}(v)| \geq \frac{\xi^{n(j-1)/2}}{2}\right)\\ &= 2\mathbb{P}\left(|\widetilde{N}^{j-1}_{-x}(v)| \geq \frac{\xi^{n(j-1)/2}}{2}\right).\end{aligned}$$

Furthermore $|\widetilde{N}^{j-1}_{-x}(v)|$ and $|\widetilde{N}^{j-1}_{-v}(x)|$ have also identical distribution, so by (6.6) for every $x \in \widetilde{N}(v)$

$$\mathbb{P}\left(|\widetilde{N}^{j-1}_{-v}(x)| \geq \frac{\xi^{n(j-1)/2}}{2}\right) \geq \frac{1}{2\hat{c}}\,. \tag{6.7}$$

Now let us order all neighbours of v in a sequence $x^1, x^2, \ldots, x^r$, where by Fact 6.9 a.a.s. $r \geq \xi^n$. For every $i \in [r]$ we define a set W_i recursively in the following way

$$W_1 = \widetilde{N}^{j-1}_{-v}(x^1).$$

Once the sets $W_1, W_2, \ldots W_{i-1}$ are found, we construct W_i by putting there the vertices of $\widetilde{N}^{j-1}_{-v}(x^i)$ one by one, and we stop when we either generate the whole set $\widetilde{N}^{j-1}_{-v}(x^i)$ or when we first find a vertex which is already in $\bigcup_{k=1}^{i-1} W_k$. In this latter case we stop generating elements of W_i, we move to the vertex x^{i+1} and we say that vertex x^i is bad.

Fact 6.10. *A.a.s. none of the vertices $x^1, x^2, \ldots, x^t$, where $t = \xi^{2n/3}$ is bad.*

PROOF. Denote by $\mathcal{B}$ the probability that one of the vertices $x^1, x^2, \ldots, x^r$ is bad, where, let us recall, $r \geq \xi^n$. For $i \in [r]$ denote by ψ_i the probability that x^i is bad. Clearly, for each $i \in [r-1]$, we have $\psi_i \leq \psi_{i+1}$. Hence

$$\mathbb{P}(\mathcal{B}) = 1 - \prod_{i=1}^{r}(1-\psi_i) \geq 1 - \prod_{i=t+1}^{r}(1-\psi_i) \geq 1 - (1-\psi_{t+1})^{r-t}\,.$$

Since due to (6.5) we have

$$\mathbb{P}(\mathcal{B}) < \frac{1}{2}\,,$$

we get

$$(1-\psi_{t+1})^{r-t} > \frac{1}{2}\,,$$

and so, estimating very crudely, we infer that $\psi_{t+1} \leq \xi^{-3n/4}$. Thus the probability that there is a bad vertex among $x^1, x^2, \ldots, x^t$ is at most

$$\sum_{i=1}^{t} \phi_i < t\phi_{t+1} < \xi^{2n/3}\xi^{-3n/4} = o(1)\,,$$

and hence the assertion of the fact holds.

■

By Fact 6.10, for $i \in [t]$, $t = \xi^{2n/3}$,

$$W_i = \widetilde{N}_{-v}^{j-1}(x^i)\,,$$

and moreover, these sets are disjoint. Thus

$$\mathbb{P}\left(|\widetilde{N}^j(v)| \geq \left|\bigcup_{i=1}^{t} \widetilde{N}_{-v}^{j-1}(x_i)\right|\right) = 1 - o(1)\,.$$

Since by (6.7), for each $i \in [t]$,

$$\mathbb{P}\left(|W_i| \geq \frac{\xi^{n(j-1)/2}}{2}\right) \geq \frac{1}{2\hat{c}}\,,$$

$$\mathbb{E}\left|\bigcup_{i=1}^{t} \widetilde{N}_{-v}^{j-1}(x_i)\right| \geq \frac{t\xi^{n(j-1)/2}}{4\hat{c}}\,.$$

By Chernoff's inequality (1.2), with probability $1 - o(1)$,

$$\left|\bigcup_{i=1}^{t} \widetilde{N}_{-v}^{j-1}(x_i)\right| \geq \frac{\mathbb{E}\left|\bigcup_{i=1}^{t} \widetilde{N}_{-v}^{j-1}(x_i)\right|}{2} \geq \frac{t\xi^{n(j-1)/2}}{8\hat{c}}\,.$$

Hence with probability $1 - o(1)$,

$$|\widetilde{N}^j(v)| \geq \frac{t\xi^{n(j-1)/2}}{8\hat{c}} > \xi^{jn/2},$$

contradicting (6.6). Thus (6.4) leads to a contradiction and Lemma 6.8 holds.

■

6.3 Proof of Theorem 6.1

Proof. Let $\alpha = \gamma$ and $\beta + \gamma > 1$. By Theorem 6.2 a.a.s. for each vertex of $\mathcal{K}(n, \mathbf{P})$ there exists a path of length at most c joining it to $\mathcal{K}_{mid}$. By Theorem 6.6 the diameter of $\mathcal{K}_{mid}$ is bounded from above by c'. Consequently

$$\text{diam}(\mathcal{K}(n, \mathbf{P})) < 2c + c'\,,$$

where c and c' are constants. This completes the proof of Theorem 6.1.

■

Notation

$V(G)$	vertex set of G
$E(G)$	edge set of G
$u \sim v$	u, v are adjacent
$\deg_G(v)$	degree of $v \in V(G)$
$\delta(G), \Delta(G)$	minimum and maximum degrees of G
$\bar{S}$	complement of $S \subseteq V(G)$
$N_G(S)$	neighbourhood of S
$E_G(S, T)$	set of edges between $S, T \subseteq V(G)$
$e_G(S, T) = \lvert E_G(S, T) \rvert$	
$\mathrm{Vol}(S) = \sum_{v \in S} \deg_G(v)$	volume of $S \subseteq V(G)$
$\mathrm{dist}_G(u, v)$	distance between vertices v and u in G
$N_G^k(v)$	k-th neighbourhood of $v \in V(G)$
$\mathrm{diam}(G)$	diameter of G
$\mathbb{P}(\mathcal{A})$	probability of $\mathcal{A}$
$\mathbb{E}(X)$	expected value of X
$\mathrm{Bi}(n, p)$	random variable with binomial distribution with parameters n and p
$G(n, p)$	random graph with n vertices and probability p of every edge
$\mathcal{K}(n, \mathbf{P})$	Kronecker graph
$p_{u,v}$	probability that $u \sim v$ in $\mathcal{K}(n, \mathbf{P})$
$N = 2^n$	size of vertex set of $\mathcal{K}(n, \mathbf{P})$
$\mathrm{w}(v)$	weight of v
$\widetilde{\mathrm{w}}(v)$	weight of good neighbours of v
$\mathrm{d}(v, u)$	Hamming distance between vertices u, v of $\mathcal{K}(n, \mathbf{P})$

Bibliography

[1] Justyna Banaszak, On matchings in stochastic Kronecker graphs. *The Electronic Journal of Combinatorics*, **23** (2016).

[2] Justyna Banaszak and Tomasz Łuczak, On the diameter of Kronecker graphs. *Discrete Mathematics*, **341** (2018), 3165–3173.

[3] Elizabeth Bodine, Babak Hassibi and Adam Wierman, Generalizing Kronecker graphs in order to model searchable networks. *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing, Allerton'09*, 2009, 194–201.

[4] Béla Bollobás, The diameter of random graphs. *Transactions of the American Mathematical Society*, **267** (1981), 41–52.

[5] Béla Bollobás, *Random Graphs.* Cambridge studies in advanced mathematics, Cambridge University Press, 2001.

[6] Yu. D. Burtin, On extreme metric characteristics of a random graph, I. *Teor. Veroyatnost. i Primenen.*, **19** (1974), 740–754.

[7] Yu. D. Burtin, On extreme metric characteristics of a random graph, II. *Teor. Veroyatnost. i Primenen.*, **20** (1975), 82–99.

[8] Deepayan Chakrabarti, Christos Faloutsos, Zoubin Ghahramani, Jon Kleinberg and Jure Leskovec, Kronecker graphs: an approach to modeling networks. *Journal of Machine Learning Research*, **11** (2010), 985–1042.

[9] Deepayan Chakrabarti, Christos Faloutsos, Jon Kleinberg and Jure Leskovec, Realistic, mathematically tractable graph generation and evolution, using Kronecker multiplication. *Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases*, 2005, 133–145.

[10] Fan R. K. Chung, *Spectral Graph Theory.* Regional conference series in mathematics **92** Conference Board of the mathematical sciences by the American Mathematical Society, 1997.

[11] Reinhard Diestel, *Graph Theory.* Graduate Texts in Mathematics. Springer, 3rd edition, 2006.

[12] Paul Erdős and Alfréd Rényi, On random graphs I. *Publicationes Mathematicae (Debrecen)*, **6** (1959), 290–297.

[13] Paul Erdős and Alfréd Rényi, On the evolution of random graphs. *Publ. Math. Inst. Hungary. Acad. Sci.*, **5** (1960), 17–61.

[14] Paul Erdős and Alfréd Rényi, On the strength of connectedness of random graphs. *Acta Math. Acad. Sci. Hungar*, **12** (1961), 261–267.

[15] Paul Erdős and Alfréd Rényi, On random matrices. *Publ. Math. Inst. Hungary. Acad. Sci.*, **8** (1964), 455–461.

[16] Paul Erdős and Alfréd Rényi, On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungar*, **17** (1966), 359–368.

[17] Christos Faloutsos and Jure Leskovec, Scalable modeling of real graphs using Kronecker multiplication. *Proceedings of the 24th international conference on Machine learning*, 2007, 497–504.

[18] Alan Frieze and Michał Karoński, *Introduction to Random Graphs.* Cambridge University Press, 2015.

[19] E. N. Gilbert, Random graphs. *The Annals of Mathematical Statistics*, **30** (1959), 1141–1144.

[20] Paul Horn and Mary Radcliffe, Giant components in Kronecker graphs. *Random Structures & Algorithms*, **40** (2012), 385–397.

[21] Svante Janson, Tomasz Łuczak and Andrzej Ruciński, *Random Graphs.* Wiley-Interscience series in discrete mathematics and optimization. John Wiley, 2000.

[22] Mihyun Kang, Michał Karoński, Christoph Koch and Tamás Makai, Properties of stochastic Kronecker graphs. *Journal of Combinatorics*, **6** (2015), 395–432.

[23] Victor Klee and David Larman, Diameters of random graphs. *Canad. J. Math*, **33** (1981), 618–640.

[24] Tamara Kolda, Ali Pinar and Comandur Seshadhri, An in-depth study of stochastic Kronecker graphs. *IEEE 11th International Conference on Data Mining*, 2011, 587–596.

[25] Mohammad Mahdian and Ying Xu, Stochastic Kronecker graphs. *Random Structures & Algorithms*, **38** (2011), 453–466.

[26] Mary Radcliffe and Stephen Young, Connectivity and giant components of stochastic Kronecker graphs. *Journal of Combinatorics*, **6** (2015), 457–482.