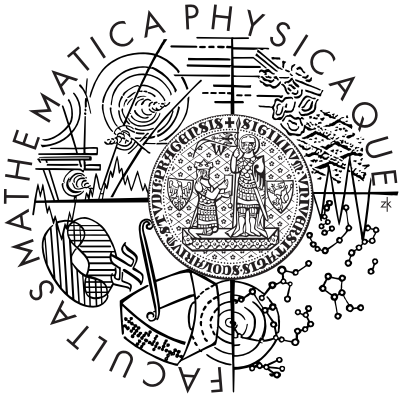
Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Jan Hladký

Strukturální vlastnosti grafů — pravděpodobnostní a deterministický pohled

(Bipartitní podgrafy v náhodném kubickém grafu)

Structural properties of graphs — probabilistic and deterministic point of view

(Bipartite subgraphs in a random cubic graph)

Katedra aplikované matematiky

Vedoucí bakalářské práce: RNDr. Robert Šámal, Ph.D.
Studijní program: Matematika

2006

Děkuji svému školiteli Robertovi Šámalovi za trpělivé a pozorné vedení této práce.

Část této práce vznikla během mého pobytu na DIMACS v létě 2006, kdy jsem byl podporovován programem DIMACS-DIMATIA International REU Program NSF EIA-01-38973. Jsem vděčný National Science Foundation za tuto štědrou podporu a Katedře aplikované matematiky, MFF UK, která mi pobyt umožnila.

Děkuji rodině za podporu během studia.

I acknowledge my advisor Robert Šámal for guiding me patiently and carefully with the thesis.

Part of the thesis was done while I was staying at DIMACS during summer 2006, supported under DIMACS-DIMATIA International REU Program NSF EIA-01-38973. I am grateful to National Science Foundation for the generous support as well as to Department of Applied Mathematics, Charles University in Prague, which made the stay possible.

I thank my family for support during the study.

Prohlašuji, že jsem bakalářskou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 8. srpna 2006

Jan Hladký

Contents

Název práce: Strukturální vlastnosti grafů—pravděpodobnostní a deterministický pohled
Autor: Jan Hladký
Pracoviště: Katedra aplikované matematiky, Univerzita Karlova v Praze
Vedoucí bakalářské práce: RNDr. Robert Šámal, Ph.D.
E-mail vedoucího: samal@kam.mff.cuni.cz
Abstrakt: V práci zkoumáme bipartitní podgrafy náhodného kubického grafu.

Ukážeme, že hranově maximální bipartitní podgraf náhodného kubického grafu na n vrcholech má asymptoticky skoro jistě méně než $\frac{3}{2} \cdot 0.9351n$ hran. Dále ukážeme, že počet vrcholů vrcholově maximálního indukovaného bipartitnho podgrafu náhodného kubického grafu asymptoticky skoro jistě leží v intervalu $[0.75n; 0.9082n]$. K získání dolního odhadu zkonstruujeme randomizovaný algoritmus na hledání velkého indukovaného biparitního podgrafu v náhodném kubickém grafu.

V závěru práce diskutujeme důsledky pro grafové homomorfismy, zejména pro Nešetřilovu Pětiúhelníkovou domněnku.
Klíčová slova: náhodný kubický graf, randomizovaný algoritmus, bipartitní podgraf

Title: Structural properties of graphs—probabilistic and deterministic point of view
Author: Jan Hladký
Department: Department of Applied Mathematics, Charles University in Prague
Supervisor: Robert Šámal, Ph.D.
Supervisor's e-mail address: samal@kam.mff.cuni.cz
Abstract: We study bipartite subgraphs of a random cubic graph in the thesis.

We show, that an edge-maximum bipartite subgraph of a random cubic graph on n vertices has asymptotically almost surely less then $\frac{3}{2} \cdot 0.9351n$ edges. We also show that the number of vertices of a vertex-maximum induced bipartite subgraph of a random cubic graph lies within interval $[0.75n; 0.9082n]$. To obtain the lower bound we design a randomized algorithm for finding a large induced bipartite subgraph of a random cubic graph.

We discuss consequences of the results for graph homomorphisms, namely for Pentagon Conjecture posed by Nešetřil.
Keywords: random cubic graph, randomized algorithm, bipartite subgraph

Chapter 1

Overview of the results

We give upper bounds for bipartite density (bipartite density is the ratio of the number of edges in the edge-maximum bipartite graph over the number of edges) and size of maximum induced bipartite subgraph of a random cubic graph in Chapter 3. We show that the number of vertices of a maximum induced bipartite subgraph of a random cubic graph is a.a.s. at most $0.9082n$ and bipartite density is a.a.s. at most 0.9351. First moment method is used to obtain the results. Chapter 4 is completely devoted to probabilistic analysis of randomized algorithm BIPGREEDY which finds large induced bipartite subgraphs in cubic graphs. The analysis gives us a lower bound on the size of maximum induced bipartite subgraph of a random cubic graph. Concretely, we show that the size of maximum induced bipartite subgraph of a random cubic graph on n vertices is a.a.s. at least $(0.75 - \varepsilon)n$ vertices. The analysis is not tight; we include data from computer simulation which show that the bipartite subgraph found by the computer has a.a.s. at least $0.8179n$ vertices. We sketch how a tight analysis could be done, with only (hopefully) little more effort. The estimates on the size of maximum induced bipartite subgraph are natural counterparts to estimates on the size of maximum independent set and are obtained by the same methods. McKay [McK87] showed that the size of a maximum independent set is a.a.s. at most $0.4554n$ for a random cubic graph. Frieze and Suen [FS94] obtained an lower bound for independence number; the size of maximum independent set is at most $0.4327n$.

In Chapter 5 we deduce some consequences for graph homomorphisms, namely to homomorphisms onto odd cycles. Our motivation comes from the Pentagon Conjecture of Nešetřil. We show that there is a cubic graph with arbitrary high girth which is not homomorphic to C_{11}. This is a better result than the one obtained analogously from McKay's bound on independence number. Still, the method gives result much weaker then approach of Hatami [Hat05].

Chapter 2

Preliminaries

2.1 Graph theory

The gentle reader is referred to [Die05] for the basic notation from graph theory (such as 'independent set' or 'girth'). The explanation of the term of 'graph' is left to chapter Random cubic graphs; we just emphasize at this point that the term will be used in a bit sloppy way throughout the thesis. We vary with the Diestel's book only in the notation of minimal degree of a graph; for graph G we write it as $\mathrm{mindeg}(G)$.

2.2 Probability

We assume some knowledge of probability theory. Book [Ros00] can serve as a good source. From the more advanced stuff, only moment generating functions will be used here without explanation.

All probability spaces we are dealing with in the thesis are finite or countable. This guarantees us we will not get into any measure theoretical problems when examining the probabilities only of the elementary events.

We usually do not write which space we are working in; it should be clear from the context. For a finite set A we denote by A_u the uniform probability space on with set of elementary events A. Next we explain symbol $\in_u$. By writing $\mathbf{Pr}[x \text{ has property} P; x \in_u A]$ (for a finite set A) we mean exactly the same as by $\mathbf{Pr}_{A_u}[x \text{ has property } P]$. For example $\mathbf{Pr}[x \text{ is odd}; x \in_u \{1,2,3\}] = 2/3$.

We write $X \sim D$ when random variable X has distribution D. We write $X \overset{\text{ind}}{\sim} D$ as a shortcut for introducing a new random variable X with distribution D independent on all variables and events used before. We describe what this means formally: let Ω be the probability space we worked in before introducing X. Let Ω_D be the probability space which has the structure that allows us to define distribution D on. Now, we take product space $\Omega \times \Omega_D$ and work (naturally) in it.

Distribution $Be(p)$ is Bernoulli distribution with success probability p, that is for

$X \sim Be(p)$ it holds that

$$\mathbf{Pr}[X = n] = \begin{cases} 1-p & \text{for } n = 0 \\ p & \text{for } n = 1 \end{cases}$$

We say that event E holds *asymptotically almost surely* (*a.a.s.*) in a sequence of probability spaces $\{\Omega_i\}_{i\in\mathbb{N}}$ if $\lim_{i\to\infty} \mathbf{Pr}_{\Omega_i}[E] = 1$.

We conclude the section with two notorious inequalities of probability theory: Markov Inequality and Chernoff Inequality.

Theorem 2.1 (Markov Inequality)**.** *Let X be a nonnegative random variable with finite expectation. Let $t > 0$ be arbitrary. Then*

$$\mathbf{Pr}[X \geq t] \leq \frac{\mathbf{E}[X]}{t}$$

The version of Chernoff Inequality presented here is not tight, see [Jan02] for a better tail estimate.

Theorem 2.2 (Chernoff Inequality, Remark 2.9 in [JŁR00])**.** *Let $X_1, X_2, \ldots, X_n$ be independent random variables attaining values in $[0, 1]$, denote $X = \sum_{i=1}^{n} X_i$, $\sigma^2 = \mathbf{Var}[X]$. Then for any $\varepsilon > 0$*

$$\mathbf{Pr}[X - \mathbf{E}[X] \geq \varepsilon] < \exp\left(\frac{-\varepsilon^2}{\sigma^2 + \varepsilon/3}\right)$$

and

$$\mathbf{Pr}[X - \mathbf{E}[X] \leq -\varepsilon] < \exp\left(\frac{-\varepsilon^2}{\sigma^2 + \varepsilon/3}\right)$$

2.3 Markov chains

We refer to [MU05] for a brief introduction to the theory of stochastic processes. We shortly recall here the basic notation of the theory of finite Markov chains, following [MU05]. Book [Ros96] deals with the subject in much more detail. This section contains two unproven statements: Fact 2.4 and Fact 2.9. Both are well known facts in the theory which can be found in Chapter 4 of [Ros96], for example. We give them tailored to what we need in the thesis.

Let $\{F_t\}_{t\in\mathbb{N}_0}$ be a discrete time stochastic process. We say that $\{F_t\}_{t\in\mathbb{N}_0}$ is a *Markov chain* if

$$\mathbf{Pr}[F_t = a_t | F_{t-1} = a_{t-1}, F_{t-2} = a_{t-2}, \ldots, F_0 = a_0] = \mathbf{Pr}[F_t = a_t | F_{t-1} = a_{t-1}]$$

for any time $t \in \mathbb{N}$ and $a_i \in \mathbb{R}, i = 0, \ldots, t$. We call F_t the *state* of the process at time t. The *state space* is the set of all possible states over all times. If the state space is finite, then we say that the Markov chain is *finite*. Without loss of generality we assume that the state of any finite Markov chain is $\{1, 2, \ldots, n\}$ (for some n). Markov chain $\{F_t\}_{t\in\mathbb{N}_0}$

is uniquely defined by F_0 and its *transition matrix* P. Entry $P_{i,j}$ is the probability that the process moves in one step from state i to j.

$$P_{i,j} = \mathbf{Pr}[F_t = j | F_{t-1} = i]$$

By $\pi(t)$ we denote the distribution of the states of the Markov chain at time t. We write its coordinates in brackets in the superscript, i.e. $\pi^{(i)}(t)$ denotes the probability that the Markov chain is in state i at time t. It follows immediately that

$$(\pi^{(1)}(t), \pi^{(2)}(t), \ldots, \pi^{(n)}(t)) = (\pi^{(1)}(t-1), \pi^{(2)}(t-1), \ldots, \pi^{(n)}(t-1))\ P$$

It is often useful to think of a Markov chain as of a walk on a directed graph. The graph is constructed as follows. Each vertex corresponds to a state. Edge (i, j) is contained in the graph if and only if $P_{i,j} > 0$; $w((i, j)) = P_{i,j}$ is the weight of the edge. At time $t = 0$ we pick a starting vertex of the walk at random with distribution of the probabilities $\pi(0)$. At each time $t > 0$ we move to vertex j with probability $w((i, j))$, where i is the vertex we are currently at.

Example 2.3. Figure 2.1 shows a graph corresponding to a Markov chain with transition matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ 0 & 0 & 1 \\ \frac{4}{5} & \frac{1}{5} & 0 \end{pmatrix}$$

Consider the starting distribution $\pi(0)$ is $\pi^{(1)}(0) = \frac{1}{2}, \pi^{(2)}(0) = \frac{1}{4}, \pi^{(3)}(0) = \frac{1}{4}$. The

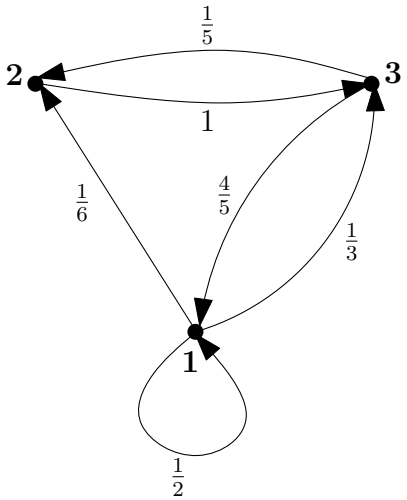


FIGURE 2.1: A random walk on a graph with three vertices

following table shows the distribution at times $t = 0, 1, 2, \ldots, 5$. The probabilities are rounded to four digits.

distribution	$\pi^{(1)}(t)$	$\pi^{(2)}(t)$	$\pi^{(3)}(t)$
$t=0$	.5000	.2500	.2500
$t=1$	.4500	.1333	.4167
$t=2$	.5583	.1583	.2833
$t=3$	.5058	.1497	.3444
$t=4$	.5285	.1532	.3183
$t=5$	.5189	.1517	.3294

One may guess that the distributions in the above example converge to a certain distribution $\overline{\pi}$. This is true and it does happen in a big family of Markov chains.

A finite Markov chain is *irreducible* if its graph representation is a strongly connected graph. A state i in a Markov chain $\{F_t\}_{t\in\mathbb{N}_0}$ is *periodic* if there exists an integer $\Delta > 1$ such that $\mathbf{Pr}[F_{t+s} = i|F_t = i] = 0$ whenever s is not divisible by Δ. Markov chain is *periodic* if any state is periodic. Markov chain is *aperiodic* if it is not periodic.

We say that $\overline{\pi}$ is a *stationary distribution* of a Markov chain with transition matrix P if $\overline{\pi} = \overline{\pi}\, P$.

Now we can give a statement about existence and uniqueness of a stationary distribution and about the convergence of the distribution to it.

Fact 2.4. *Any irreducible aperiodic finite Markov chain with transition matrix P has the following properties*

(a) The chain has a unique stationary distribution $\overline{\pi} = (\overline{\pi}^{(1)}, \overline{\pi}^{(2)}, \ldots, \overline{\pi}^{(n)})$.

(b) Given any starting distribution $\pi(0)$ the distribution converges to $\overline{\pi}$, $\lim_{t\to\infty} \pi(t) = \overline{\pi}$.

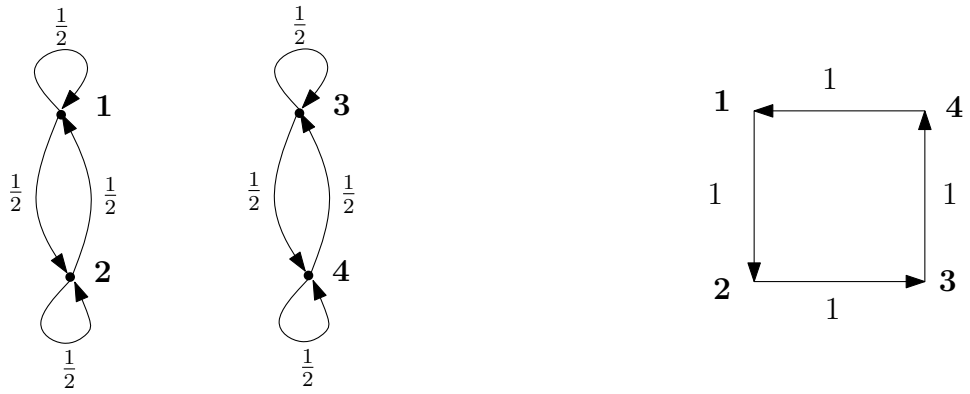
(c) Vector $\overline{\pi}$ is the only eigenvector (up to a scalar multiple) of P^T corresponding to its largest eigenvalue $\lambda = 1$.

The stationary distribution of the chain from Example 2.3 is $\overline{\pi} \doteq (0.5219, 0.1522, 0.3262)$.

Example 2.5. Figure 2.2(a) shows an example of a Markov chain which has infinitely many stationary distributions. Two of these are $\overline{\pi_1} = (\frac{1}{2}, \frac{1}{2}, 0, 0)$ and $\overline{\pi_2} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. The graph is not strongly connected.

Graph on Figure 2.2(b) does not correspond to an aperiodic Markov chain. It is easy to check that there is only one starting distribution, which is convergent, $\pi(0) = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

What is often of interest is how fast a Markov chain converges to the stationary distribution, irrespective of the starting distribution $\pi(0)$. To measure this we introduce *variation distance* $\|\cdot\|$ of two distributions π_1 and π_2. It is given by $\|\pi_1 - \pi_2\| = \frac{1}{2}\sum_i |\pi_1^{(i)} - \pi_2^{(i)}|$. Mixing time of a Markov chain is a time when, irrespective to $\pi(0)$, the variation distance of $\pi(t)$ and $\overline{\pi}$ starts to be small enough. The following theorem is useful for bounding mixing time of a Markov chain.



(a) (b)

FIGURE 2.2: An example of a Markov chain which is not (a) irreducible, (b) aperiodic

Theorem 2.6 (Theorem 11.5 in [MU05])**.** *Let P be the transition matrix for a finite, irreducible, aperiodic Markov chain with stationary distribution $\overline{\pi}$. Let m_j be the smallest entry in the j-th column of matrix P and let $m = \sum_j m_j$. Then, for any starting distribution $\pi(0)$ it holds for all t that*

$$||\pi(t) - \overline{\pi}|| \leq (1-m)^t$$

The next two lemmas deal with a special kind of random walk $\{X_t\}_{t\in\mathbb{N}_0}$ on $\mathbb{N}_0$. Let Z be a random variable with support on $\mathbb{N}_0$, let $G_Z(s)$ be its probability generating function,

$$G_Z(s) = \sum_{i=0}^{\infty} s^i \, \mathbf{Pr}[Z = i]$$

Suppose that the radius of convergence of G_Z is a constant greater than 1. Let p and q be constants such that $0 \leq p < q \leq 1$. The Markov chain $\{X_t\}_t$ is defined by the transition probabilities

$$\begin{aligned} \mathbf{Pr}[X_{t+1} = i \mid X_t = 0] &= \mathbf{Pr}[Z = i] \\ \mathbf{Pr}[X_{t+1} - X_t = i \mid X_t \neq 0] &= \begin{cases} p & \text{if } i = 1 \\ 1-p-q & \text{if } i = 0 \\ q & \text{if } i = -1 \end{cases} \end{aligned}$$

and initial state $X_0 = 0$. Random variable R_k is the time elapsed when X_t returns for the k-th time to state 0. Random variable $W(t)$ counts how many times state 0 was visited during time interval $[0, t]$.

The proof of the following lemma goes along the same lines as proof of Lemma 3 in [FS94] (the lemma itself is contained in [FS94] as Lemma 4 with the proof omitted).

Lemma 2.7. *For any $A > 0$ we have as $k \to \infty$,*

$$\mathbf{Pr}\left[\left|R_k - \frac{k(q-p+\mathbf{E}[Z])}{q-p}\right| \geq A\sqrt{k}\right] \leq O(\exp(-A))$$

Proof. First consider Markov chain $\{X'_t\}_t$ with the same states and transition probabilities as $\{X_t\}_t$ and initial state $X'_0 = 1$. Let τ be the first time t when $X'_t = 0$, let τ_1 and τ_2 be two independent copies of τ. Then τ equals in distribution to constant 1 with probability q, $1+\tau_1$ with probability, $1 + \tau_1 + \tau_2$ with probability p. Let $\varphi(s)$ be the moment generating function of τ, $\varphi(s) = \mathbf{E}[\exp(s\tau)]$. Then we have

$$\varphi(s) = \exp(s)\left(q \cdot 1 + (1-p-q)\cdot\varphi(s) + p\cdot\varphi^2(s)\right)$$

Solving it (algebraically), we get

$$\varphi(s) = \frac{1 - \exp(s)(1-p-q) - \sqrt{(\exp(s)(1-p-q)-1)^2 - 4pq\exp(2s)}}{2p\exp(s)}$$

From this, one can easily calculate the expectation of τ,

$$\mathbf{E}[\tau] = \frac{\mathbf{d}}{\mathbf{d}s}\varphi(s)_{|s=0} = \frac{1}{q-p}$$

Let $\psi(s)$ be the moment generating function for R_1, $\psi(s) = \mathbf{E}[\exp(sR_1)]$. From how the transition probabilities look like we see, that

$$\psi(s) = \exp(s)G_Z(\varphi(s)) \tag{2.1}$$

and direct computation gives us

$$\mathbf{E}[R_1] = \frac{\mathbf{d}}{\mathbf{d}s}\psi(s)_{|s=0} = 1 + \frac{\mathbf{E}[Z]}{q-p}$$

From (2.1) it is easy to extract terms $[s^0]$ and $[s^1]$ of $\psi(s)$; as $s \to 0$ we have

$$\psi(s) = 1 + \frac{q-p+\mathbf{E}[Z]}{q-p}s + O(s^2)$$

Variable R_k equals in distribution to the sum of k independent copies of R_1, which gives us $\psi(s)^k = \mathbf{E}[\exp(sR_k)]$.

$$\begin{aligned}
&\mathbf{Pr}\left[R_k - \frac{k(q-p+\mathbf{E}[Z])}{q-p} \geq A\sqrt{k}\right] \\
&= \mathbf{Pr}\left[\exp(sR_k) \geq \exp\left(Ask^{1/2} + \frac{sk(q-p+\mathbf{E}[Z])}{q-p}\right)\right] \\
&\leq \frac{\mathbf{E}\left[\exp(sR_k)\right]}{\exp\left(Ask^{1/2} + \frac{sk(q-p+\mathbf{E}[Z])}{q-p}\right)} \\
&= \psi(s)^k \exp\left(-Ask^{1/2} - \frac{sk(q-p+\mathbf{E}[Z])}{q-p}\right) \\
&= \left(1 + \frac{q-p+\mathbf{E}[Z]}{q-p}s + O(s^2)\right)^k \exp\left(-Ask^{1/2} - \frac{sk(q-p+\mathbf{E}[Z])}{q-p}\right) \\
&= \exp\left(O(ks^2) - Ask^{1/2}\right)
\end{aligned}$$

where we used Markov Inequality between lines 2 and 3. Substituting $s = k^{-1/2}$ we get

$$\mathbf{Pr}\left[R_k - \frac{k(q-p+\mathbf{E}[Z])}{q-p} \geq A\sqrt{k}\right] = O(\exp(-A))$$

The tail estimate of $\mathbf{Pr}\left[R_k - \frac{k(q-p+\mathbf{E}[Z])}{q-p} \leq -A\sqrt{k}\right]$ is analogous. □

Lemma 2.8.

$$\mathbf{Pr}\left[W(t) < \frac{t(q-p)}{q-p-\mathbf{E}[Z]} - A\sqrt{t}\right] = O\left(\exp(-A)\right)$$

Proof.

$$\begin{aligned}
&\mathbf{Pr}\left[W(t) < \frac{t(q-p)}{q-p+\mathbf{E}[Z]} - A\sqrt{t}\right] \\
&= \mathbf{Pr}\left[R_{\frac{t(q-p)}{q-p+\mathbf{E}[Z]}-A\sqrt{t}} > t\right] \\
&\leq \mathbf{Pr}\left[\left|R_{\frac{t(q-p)}{q-p+\mathbf{E}[Z]}-A\sqrt{t}} - t + \frac{A\sqrt{t}(q-p+\mathbf{E}[Z])}{q-p}\right| > \frac{A\sqrt{t}(q-p+\mathbf{E}[Z])}{q-p}\right] \\
&= O\left(\exp\left(-\frac{A(q-p+\mathbf{E}[Z])}{q-p}\right)\right) \\
&= O\left(\exp(-A)\right)
\end{aligned}$$

where we used Lemma 2.7 between lines 3 and 4. □

A state S of Markov chain $\{F_t\}_{t\in\mathbb{N}_0}$ is called *absorbing* if it is impossible to leave it, i.e. $\mathbf{Pr}[F_{t+1} \neq S \mid F_t = S] = 0$ for every time $t \in \mathbb{N}_0$. A Markov chain is *absorbing* if it is possible to get from every state (in one or more steps) to an absorbing state. For an absorbing Markov chain, a state which is not absorbing is called *transient* (in literature, transient states are defined for any Markov chain in a more general way, however for absorbing Markov chains the two definitions coincide).

By renumbering the states in the transition matrix P of an absorbing Markov chain we can get in to the *canonical form*

$$P = \left(\begin{array}{c|c} Q & R \\ \hline 0 & \mathbb{1} \end{array}\right)$$

where matrix Q is a transition matrix between the transient states, matrix R records transition probabilities from the transient states to the absorbing states and $\mathbb{1}$ is an identity matrix representing equability of the absorbing states. The following fact is a basic fact about absorbing Markov chains.

Fact 2.9. *For an absorbing Markov chain the matrix $\mathbb{1} - Q$ has an inverse N, $N = \mathbb{1} + Q + Q^2 + \cdots + Q^i + \cdots$. Entry $N_{i,j}$ is the expected number of times the chain is in state j, given that it starts in state i.*

Matrix N from the Fact is called the *fundamental matrix.*

In the thesis we will use absorbing Markov chains with only one absorbing state.

Proposition 2.10. *Let $\{F_t\}_{t\in\mathbb{N}_0}$ be a Markov chain with only one absorbing state S, such that $F_0 = j$ and P is its transition matrix,*

$$P = \left(\begin{array}{c|c} & p_1 \\ Q & \vdots \\ & p_{n-1} \\ \hline 0 & 1 \end{array}\right)$$

Then the probability that the last visited transient state was state i is $p_i N_{j,i}$.

Proof.

$$\begin{aligned} &\mathbf{Pr}[\text{the last visited transient state was state } i] \\ &= \sum_{t=0}^{\infty} \mathbf{Pr}[F_{t+1} = S \mid F_t = i] \\ &= p_i \sum_{t=0}^{\infty} \mathbf{Pr}[F_t = i] \\ &= p_i\, \mathbf{E}[\text{number of times the chain is in state } i] \\ &= p_i N_{j,i} \end{aligned}$$

□

Chapter 3

Random cubic graphs

The main results in the thesis deal with random graphs. Book [AS00] is a comprehensive introduction to the Probabilistic method (which a tool used not only to obtain results in graph theory but also in other parts of mathematics). All the important concepts needed here are introduced there. For further results on random graphs we refer to [JŁR00] and [Bol01].

By $\mathcal{G}_{3,n}$ we denote the uniform probability space of all cubic graphs on vertices $v_1, v_2, \ldots, v_n$. This model is what one would imagine saying 'a random cubic graph'. Alas, any direct computation in the model is practically unmanageable.

We will use the *pairing model* of cubic graphs to deal with random cubic graphs. The model was first introduced by Bender and Canfield [BC78]. The elements of the probability space $\mathcal{G}^*_{3,n}$ are all perfect matching on $\{1,\ldots,n\}\times\{1,2,3\}$ (for n even; we will not emphasize the condition henceforth); every element of $\mathcal{G}^*_{3,n}$ has the same probability. To obtain from a matching $G \in \mathcal{G}^*_{3,n}$ a cubic multigraph (with possible loops) one projects the elements $\{1,2,\ldots,n\}\times\{1,2,3\}$ in a natural way onto the set of *vertices* of G which we denote by $V(G) = \{v_1, v_2, \ldots, v_n\}$. The matchings in $\mathcal{G}^*_{3,n}$ will be called inaccurately *graphs*; and we also translate graph-theoretical terms to $\mathcal{G}^*_{3,n}$. For a vertex v_i of G we write $v_i = (v_i^1, v_i^2, v_i^3)$ where $v_i^l = (i,l) \in \{1,\ldots,n\}\times\{1,2,3\}, l = 1,2,3$.

We let $\mathrm{f}(n)$ denote the number of matchings on n elements. For n odd, $\mathrm{f}(n) = 0$, for n even, $\mathrm{f}(n) = \frac{\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2}\cdots\binom{2}{2}}{(n/2)!} = \frac{n!}{(n/2)!2^{n/2}}$. It follows from the definition that $|\mathcal{G}^*_{3,n}| = \mathrm{f}(3n)$.

We draw attention to the fact that after projecting, not all the graphs (in the proper sense) have the same probability in $\mathcal{G}^*_{3,n}$. Figure 3.1 shows two graphs G_1 and G_2, G_1 corresponding to six matchings $P_1^1, P_1^2 \ldots, P_1^6$ and G_2 corresponding to nine matchings $P_2^1, P_2^2 \ldots, P_2^9$. However, for simple graphs without loops the situation gets easier as the following facts ([Wor99b]) show.

Fact 3.1. *Conditioning that $G \in \mathcal{G}^*_{3,n}$ is simple and without loops we get a uniform sample from $\mathcal{G}_{3,n}$.*

Fact 3.2. *The probability that $G \in \mathcal{G}^*_{3,n}$ is simple and without loops tends to $e^{-9/4}$.*

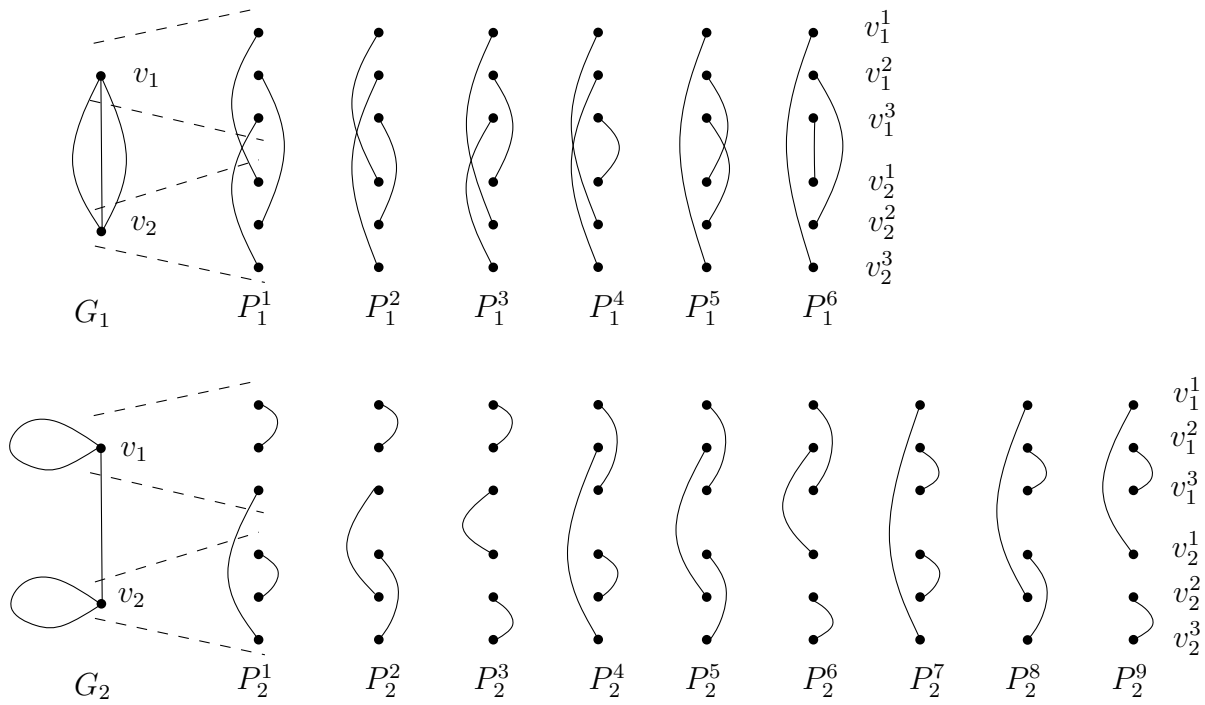


FIGURE 3.1: G_1, G_2 corresponding to graphs of different probabilities in $\mathcal{G}^*_{3,n}$

Generalization of Fact 3.2 is the following

Fact 3.3. *The probability that $G \in \mathcal{G}^*_{3,n}$ has girth at least k tends to $\prod_{i=2}^{k-1} \exp\left(-\frac{2^i}{i+1}\right)$.*

The following theorem can be used to show strong concentration of a random variable in $\mathcal{G}^*_{3,n}$ (such as the size of maximum independent set) around its expectation. The theorem uses notion of *simple switching*. Denote the set of elements $\{1,\ldots,n\} \times \{1,2,3\}$ on which we perform the matching by $\{e_1, e_2, \ldots, e_{3n}\}$. We say that two graphs $G_1, G_2 \in \mathcal{G}^*_{3,n}$ differ by a simple switching if there are two edges $f_1 = \{e_i, e_j\}, f_2 = \{e_k, e_l\}$ of G_1 such that $\{e_i, e_k\}, \{e_j, e_l\}$ are edges of G_2 and otherwise G_1 and G_2 have the same edges. Figure 3.2 shows two such graphs.

Theorem 3.4 (Theorem 2.19 in [Wor99b])**.** *Let X_n be a random variable defined on $\mathcal{G}^*_{3,n}$ such that $|X_n(G_1) - X_n(G_2)| \leq c$ whenever G_1 and G_2 differ by a simple switching. Then*

$$\mathbf{Pr}[|X_n - \mathbf{E}[X_n]| \geq t] \leq 2\exp\left(\frac{-t^2}{3nc^2}\right)$$

for all $t > 0$.

We conclude this preparatory part by defining two graph parameters. For graph G, number $\mathrm{bi}(G)$ denotes the number of vertices of vertex-maximum induced bipartite subgraph of G. *Bipartite density* $\mathrm{bd}(G)$ of a graph G is the fraction of edges in the edge-maximum bipartite subgraph of G. Obviously,

$$\mathrm{bi}(G) = \frac{MAXCUT(G)}{|E(G)|}$$

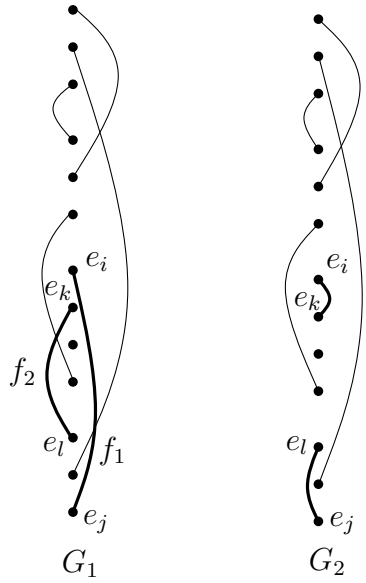


FIGURE 3.2: Two graphs differing by a simple switching.

Since determining $MAXCUT(G)$ for a cubic graph is well-known to be an NP-hard problem, so is the problem of determining the value of $\mathrm{bi}(G)$.

3.1 Maximum induced bipartite subgraph — the upper bound

Theorem 3.5. $\frac{\mathrm{bi}(G)}{n} \leq 0.9301$ *a.a.s. for* $G \in_u \mathcal{G}^*_{3,n}$.

Proof. We bound the probability of the event A that there is an induced bipartite subgraph of $G \in_u \mathcal{G}^*_{3,n}$ of size at least k.

$$\begin{aligned}\mathbf{Pr}[A] &\leq \sum_{\substack{B\subseteq V\\|B|=k}} \sum_{l=0}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{B_1\subseteq B\\|B_1|=l}} \mathbf{Pr}[B_1 \text{ and } B\setminus B_1 \text{ are independent}] \\ &= \binom{n}{k} \sum_{l=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{l} \sum_{h=0}^{3l} \mathbf{Pr}[A(k,l,h)]\end{aligned}$$

where denotes $A(k,l,h)$ the event that fixed sets of vertices B_1 and B_2 (of sizes k and $k-l$) induce a bipartite graph with exactly h edges.

$$\mathbf{Pr}[A(k,l,h)] = \frac{\binom{3l}{h}\binom{3(k-l)}{h} h! \binom{3(n-k)}{3k-2h} (3k-2h)!\, \mathrm{f}(3n-6k+2h)}{\mathrm{f}(3n)}$$

At the end of the estimates we will show that $\sqrt[n]{\mathbf{Pr}[A(0.9301n,l,h)]} \leq c$ for some constant $c<1$ and for all l,h. To deal with the inequalities easier we introduce notion

of asymptotically similar behavior in the n-th square root: we write $f(n) \approx g(n)$ if $\lim_{n\to\infty} \sqrt[n]{f(n)/g(n)} = 1$ and $f(n) \lessapprox g(n)$ if $\limsup_{n\to\infty} \sqrt[n]{f(n)/g(n)} \leq 1$. We use two approximations: $n! \approx (n/e)^n$ and $\sum_{i=1}^{n} F(i) \approx \max\{F(i); i = 1, 2, \ldots, n\}$ for a nonnegative function F.

$$\begin{aligned}
&\mathbf{Pr}[A(k,l,h)] \\
&= \frac{\binom{3l}{h}\binom{3(k-l)}{h} h! \binom{3(n-k)}{3k-2h} (3k-2h)!\, \mathrm{f}(3n-6k+2h)}{\mathrm{f}(3n)} \\
&\approx \frac{3^{1.5n}(n-k)^{3n-3k}\, l^{3l}\, (k-l)^{3k-3l}}{(3n-6k+2h)^{1.5n-3k+2h}\, (3k-3l-h)^{3k-3l-h}\, n^{1.5n}\, h^h\, (3l-h)^{3l-h}} \\
&= t(k,l,h)
\end{aligned}$$

For fixed k, l we find the maximum of $t(k,l,h)$ over $h \in [0, 3l]$. Since the derivative $\frac{\mathbf{d}}{\mathbf{d}h} t(k,l,h)$ tends from infinity at $h = 0$ to minus infinity at $h = 3l$, the function is maximized at the only point of zero derivative $h = \frac{3}{2}(k - n + \sqrt{n^2 - 2nk + k^2 - 4l^2 + 4kl})$.

We have the following bound:

$$\mathbf{Pr}[A] \lessapprox \binom{n}{k} \sum_{l=0}^{\lfloor \frac{k}{2} \rfloor} \frac{k^k}{l^l (k-l)^{k-l}} t(k, l, 1.5(k - n + \sqrt{n^2 - 2nk + k^2 - 4l^2 + 4kl}))$$

Evaluating the derivative $\frac{\mathbf{d}}{\mathbf{d}l}$ of the inner term of the sum we get that the term is maximized at $l = 0.5k$.

$$\mathbf{Pr}[A] \lessapprox \frac{n^n}{k^k (n-k)^{n-k}}\, 2^k \cdot t(k, 0.5k, 1.5(k - n + \sqrt{n^2 - 2nk + 2k^2}))$$

Substituting $k = 0.9301n$ we get

$$\mathbf{Pr}[A] \lessapprox 0.99999^n$$

which proves the theorem. □

We shall use the method of switching to sharpen the bound. The idea of the method is simple: we shall show that if there is one maximum induced bipartite graph, then there are at least M of them. This will allow us to improve previous estimate

$$\mathbf{Pr}[\text{exists a large induced bipartite graph}] \leq \mathbf{E}[\text{number of large induced bipartite subgraphs}]$$

to

$$\mathbf{Pr}[\text{exists a large induced bipartite graph}] \leq \frac{\mathbf{E}[\text{number of large induced bipartite subgraphs}]}{M}$$

First we make a simple observation about the structure of maximum induced bipartite subgraphs.

Proposition 3.6. *Let G be a cubic graph, H be its maximum induced bipartite subgraph with parts $B_1, B_2 \subseteq V(G)$. Let $C = V \setminus (B_1 \cup B_2)$.*

(a) Any vertex $v \in C$ neighbors with at least one vertex from B_1 and from B_2.

(b) For any vertex $v \in B_1$ such that at least two of its neighbors v_1, v_2 (or v_1, v_2, v_3) lie in C, there are at most one of these are not adjacent to any vertex in B_1 other then v. (and symmetrically for $v \in B_2$)

Proof. (a) This is obvious.

(b) If there were two neighbors $v_1, v_2 \in C$ of v which are not adjacent to any other vertex of B_1, we obtain a larger induced bipartite subgraph by taking $\tilde{B_1} = B_1 \setminus \{v\} \cup \{v_1, v_2\}, \tilde{B_2} = B_2$, a contradiction.

□

Lemma 3.7. *Let G be a cubic graph on n vertices with maximum induced bipartite subgraph with parts B_1, B_2 of sizes $l, k-l$. Then G has at least $3^{0.5(n-k)}$ maximum induced bipartite subgraphs with parts of the same sizes.*

Proof. For any vertex $x \in C = V \setminus (B_1 \cup B_2)$ let w_x be its arbitrary (but fixed) neighbor in B_1 or B_2, such that x is adjacent to no other neighbor in that part (since $\deg x = 3$, there is always at least one). For any $v \in B_1 \cup B_2$ we denote $V_v = \{x \in C; w_x = v\} \cup \{v\}$. By Proposition 3.6 (b), if $x, y \in C, x \neq y$ then $w_x \neq w_y$, i.e. $|V_v| \leq 2$. Moreover, $|\{v \in B_1 \cup B_2; |V_v| = 2|\}| = n - k$.

Switching a vertex $v \in B_1 \cup B_2$, such that $V_v = \{v, x\}$ is an operation which replaces induced bipartite subgraph with parts B_1, B_2 by one with parts either $B_1 \setminus \{v\} \cup \{x\}, B_2$ (if $v \in B_1$) or $B_1, B_2 \setminus \{v\} \cup \{x\}$ (if $v \in B_2$).

If we switch any vertex u we obtain an induced bipartite graph. Now, it may happen that switching another vertex u' (according to the original list $\{V_v\}_{v \in B_1 \cup B_2}$) we do not get an induced bipartite graph; this is the case $V_u = \{u, c\}$, neighbors of c are u, x, c', $x \in B_1 \cup B_2$, $w_c = u'$ and u, u' lie in the same part of the bipartite graph — but it is the only case (and we observe that symmetrically switching u' first, we cannot switch u). In this case we say that u *blocks* u'.

The set $B_1 \cup B_2$ is split into a set A of d vertices which do not block any other vertex and $0.5(n-k-d)$ pairs $\{u, u'\}$ of vertices such that u blocks u'.

We can perform switching on the whole bipartite subgraph in the following way: we switch vertices of an arbitrary subset of A and for any pair $\{u, u'\}$ of blocking vertices we choose one of the three possibilities: (1) either we do not switch u nor u' or (2) we switch u only or (3) we switch u' only. The resulting bipartite subgraph will stay induced with parts of the same size; different switchings give different bipartite graphs. So the number of bipartite graphs obtained by switching the original one is $2^d \cdot 3^{0.5(n-k-d)} \geq 3^{0.5(n-k)}$. □

Theorem 3.8. $\frac{\mathrm{bi}(G)}{n} \leq 0.9082$ *a.a.s. for* $G \in_u \mathcal{G}^*_{3,n}$.

Proof. We calculate the number of cubic graphs (in the pairing model) which contain an induced bipartite subgraph on fixed parts B_1, B_2, $|B_1| = l, |B_2| = k-l$, span over h edges and satisfy condition given by Proposition 3.6 (a).

For any vertex $v = (v^1, v^2, v^3) \in C$, $C = V \setminus (B_1 \cup B_2)$ we distinguish to which vertex set (either B_1, B_2 or C) each v^i is adjacent to. There are twelve admissible cases.

In the following list of possibilities $v^i \to W$ denotes that v^i is adjacent to some vertex in $W \subseteq V$):

- *Type 0*: one of v^1, v^2, v^3 is adjacent to a vertex in C (6 cases): (1) $v^1 \to B_1, v^2 \to B_2, v^3 \to C$; (2) $v^1 \to B_1, v^2 \to C, v^3 \to B_2$; (3) $v^1 \to C, v^2 \to B_1, v^3 \to B_2$; (4) $v^1 \to B_2, v^2 \to B_1, v^3 \to C$; (5) $v^1 \to B_2, v^2 \to C, v^3 \to B_1$; (6) $v^1 \to C, v^2 \to B_1, v^3 \to B_2$.

- *Type 1*: two of v^1, v^2, v^3 are adjacent to vertices in B_1 (3 cases): (1) $v^1 \to B_1, v^2 \to B_2, v^3 \to B_1$; (2) $v^1 \to B_1, v^2 \to B_1, v^3 \to B_2$; (3) $v^1 \to B_2, v^2 \to B_1, v^3 \to B_1$.

- *Type 2*: two of v^1, v^2, v^3 are adjacent to vertices in B_2 (3 cases similar to Type 1).

First we choose parts B_1 and B_2 in G; there are $K = \binom{n}{k}\binom{k}{l}$ possibilities of doing this. Then we choose h edges spanning $B_1 \cup B_2$ (there are $H = \binom{3l}{h}\binom{3(k-l)}{h}h!$ ways of doing that) and fix them.

Let t_0, t_1, t_2 be the number of vertices of Type 0, 1, 2. The number of edges between B_1 and C is $3l-h$, the number of edges between B_2 and C is $3(k-l)-h$, so the following system of equations must hold:

$$\begin{aligned} n-k &= t_0+t_1+t_2 \\ 3l-h &= t_0+2t_1+t_2 \\ 3(k-l)-h &= t_0+t_1+2t_2 \end{aligned}$$

Solving it, we get $t_0 = 3n-6k+2h, t_1 = k+3l-h-n, t_2 = 4k-3l-h-n$.

There are

$$T_0 = 6^{t_0}\binom{n-k}{t_0}\binom{3l-h}{t_0}\binom{3(k-l)-h}{t_0}(t_0!)^2\,\mathrm{f}(t_0)$$

ways to choose t_0 vertices of C and to add edges (between them and $B_1 \cup B_2$ and between them mutually) as Type 0.

Having fixed these vertices and edges, there are

$$T_{1,2} = 3^{t_1}\binom{n-k-t_0}{t_1}\binom{3l-h-t_0}{2t_1}\binom{3(k-l)-h-t_0}{t_1}(2t_1)!t_1!\cdot 3^{t_2}t_2!(2t_2)!$$

ways how to add the remaining edges (corresponding to vertices of Type 1 and Type 2).

Let $X_{k,l,h}(G)$ be the number of induced bipartite subgraphs of G with parts of sizes l and $k-l$ spanning over exactly h edges, $G \in \mathcal{G}^*_{3,n}$.

$$\mathbf{E}[X_{k,l,h}] = \frac{KHT_0T_{1,2}}{\mathrm{f}(3n)}$$

Applying Lemma 3.7 we get a bound for the probability $\mathbf{Pr}[A_k]$ that there is a maximum induced bipartite subgraph of G of size k.

$$\mathbf{Pr}[A_k] \leq \frac{\sum_{l,h} \mathbf{E}[X_{k,l,h}]}{3^{0.5(n-k)}}$$

Using similar estimates as in Theorem 3.5, we get

$$\mathbf{E}[X_{k,l,h}] \approx \frac{2^{3n-6k+2h} 3^{2k-0.5n} l^l (k-l)^{k-l}}{n^{0.5n} h^h (3l+k-h-n)^{3l+k-h-n} (4k-3l-h-n)^{4k-3l-h-n} (3n-6k+2h)^{0.5(3n-6k+2h)}}$$

Evaluating the derivative $\frac{\mathbf{d}}{\mathbf{d}l}$ it can be shown that the righthand side term is maximized for $l = 0.5k$. A computer is used check that for all $k \in [0.9082n, n]$ the following holds.

$$\mathbf{Pr}[A_k] \lessapprox 0.99999^n$$

This proves that a random cubic graph contains asymptotically almost surely no induced bipartite subgraph of size at least $0.9082n$. □

3.2 Bipartite density — the upper bound

In this section we show that $\mathrm{bd}(G) \leq 0.9351$ a.a.s. for a random cubic graph. The fact was first announced by McKay [McK82], however the proof never appeared in print. See [Šám06] for weaker estimate $\mathrm{bd}(G) \leq 0.9386$.

Theorem 3.9. $\mathrm{bd}(G) \leq 0.9351$ *a.a.s. for* $G \in_u \mathcal{G}^*_{3,n}$.

Proof. Variable $\frac{3}{2} n \,\mathrm{bd}$ is a random variable such as X_n in Theorem 3.4 (with $c = 2$) so it is strongly concentrated around its expectation. Therefore the statement will be proven even if we show that $\mathrm{bd}(G) \leq 0.9351 - \varepsilon$ (for some $\varepsilon > 0$) a.a.s. for $G \in_u \mathcal{G}^*_{3,n}$, conditioning that G is a simple graph without loops. We denote the ratio of loopless simple cubic graphs over all cubic graphs on n vertices by c_n. Fact 3.2 says that $\lim_{n\to\infty} c_n = e^{-9/4}$.

Until the end of the proof we condition on G being simple without loops.

We find an upper bound for the probability that a random cubic graph G on n vertices has bipartite density at least $\delta = 0.93509$. Let $V(G) = B_1 \cup B_2$ be the partitioning of vertices of G such that bipartite subgraph B with parts B_1, B_2 has maximum edges. Then $\deg_B v \geq 2$ for any $v \in V(G)$ (this is the only point in the proof where G is needed to be without loops). Let $k = |B_1|$, h_i be the number of vertices in B_i (for $i = 1, 2$) having degree 3 in B. Then

$$3h_1 + 2(k - h_1) = 3h_2 + 2(n - k - h_2)$$

which gives $h_2 = 4k - 2n + h_1$. The number of edges of the bipartite graph is $3h_1 + 2(k - h_1) = 2k + h_1$.

$$\mathbf{Pr}[\mathrm{bd}(G) \geq \delta] \leq \sum_{k} \sum_{\substack{B_1 \subseteq V(G) \\ |B_1|=k}} \sum_{\substack{h_1 \\ 2k+h_1 \geq 1.5\delta n}} \mathbf{Pr}[A(B_1, h_1)]$$

where $A(B_1, h_1)$ denotes the event that a bipartite subgraph of G with parts $B_1, V(G)\setminus B_1$ has the maximum number of edges of all bipartite subgraphs of G and there are exactly h_1 vertices of degree three in part B_1.

If $|B_1| = k$ then

$$\begin{aligned}
&\mathbf{Pr}[A(B_1, h_1)] \\
&\quad \leq \frac{\binom{k}{h_1}\binom{n-k}{h_2}(3h_1 + 2(k-h_1))!\, 3^{n-h_1-h_2}\, \mathrm{f}(k-h_1)\, \mathrm{f}(n-k-h_2)}{c_n\, \mathrm{f}(3n)}
\end{aligned}$$

$$\begin{aligned}
&\mathbf{Pr}[(\exists B_1 \subseteq V(G), |B_1| = k) A(B_1, h_1)] \\
&\quad \leq \frac{\binom{n}{k}\binom{k}{h_1}\binom{n-k}{4k-2n+h_1}(2k+h_1)!\, 3^{3n-2h_1-4k}\, \mathrm{f}(k-h_1)\, \mathrm{f}(3n-5k-h_1))}{c_n\, \mathrm{f}(3n)}
\end{aligned}$$

Since the examined bipartite density δ is greater than $\frac{2}{3}$ all the terms of the expression on the right-hand side must tend to infinity as n goes to infinity. We may use the following approximation:

$$\begin{aligned}
&\mathbf{Pr}[(\exists B_1 \subseteq V(G), |B_1| = k) A(B_1, h_1)] \\
&\quad \lessapprox \frac{3^{1.5n-2h_1-4k}(2k+h_1)^{2k+h_1}}{h_1^{h_1}(4k-2n+h_1)^{4k-2n+h_1} n^{0.5n} (k-h_1)^{0.5(k-h_1)} (3n-5k-h_1)^{0.5(3n-5k-h_1)}}
\end{aligned}$$

A computer is used to check that the last term is less or equal to 0.99999^n whenever $2k + h_1 \geq 0.93509 \cdot \frac{3}{2}n$.

Setting $\varepsilon = 0.00001$ completes the proof[1]. □

[1]It could be argued that there exist even smaller positive numbers than 0.00001, so why not to use one (and get a better constant consequently)? We prefer (relatively) nice constants to long and precise ones.

Chapter 4

Algorithm BIPGREEDY

In this section we present randomized Monte Carlo algorithm BIPGREEDY for finding a large induced bipartite subgraph of a random cubic graph. Analysis of it, which we carry out in a later section, shows that it performs quite well. Analysis of the performance of BIPGREEDY gives us also a lower bound on number $\mathrm{bi}(G)$ for a random cubic graph G. We did not find any other probabilistic method to obtain a (nontrivial) lower bound on $\mathrm{bi}(G)$.

BIPGREEDY is most similar to algorithm MINGREEDY presented by Frieze and Suen in [FS94] (see also [FRS95] for some technical details) for finding a large independent set in a random cubic graph.

For comparison, we describe algorithm MINGREEDY first. We start with a cubic graph G at an input and empty independent set I. In each step a vertex v of minimum degree is randomly chosen in graph G, added to independent set I constructed in parallel, and removed with all its neighbors from the graph. This procedure guarantees that independence of the constructed set will never be violated.

MINGREEDY;
Input G;
$I := \emptyset$;
while $G \neq \emptyset$ **do**
  $V_{\min} :=$ set of vertices of minimum degree in G;
  choose $v \in V_{\min}$ uniformly at random;
  $I := I \cup \{v\}$;
  $G := G - \{v\} - \{\text{neighbors of } v\}$;
  remove isolated vertices from G;
end while
Output I;

Algorithm BIPGREEDY works as follows. Input cubic graph is dismounted while the bipartite subgraph is constructed in parallel. At each step vertex of minimum degree is picked randomly and removed from the graph. We add it to that part of the bipartite subgraph in which the property of induceness will not become violated; when there is no

such, we throw the vertex away; if adding to both parts is possible, we add it to one of them at random.

BIPGREEDY;
Input G;
$B_1 := \emptyset$;
$B_2 := \emptyset$;
while $G \neq \emptyset$ **do**
  $V_{\min}$:= set of vertices of minimum degree in G;
  choose $v \in V_{\min}$ uniformly at random;
  if v has no formal neighbor[1] in B_1 nor B_2 **then**
    if random(2)= 0 **then**
      $B_1 := B_1 \cup \{v\}$;
    else
      $B_2 := B_2 \cup \{v\}$;
    end if
  end if
  if v has formal neighbor in B_1 only **then**
    $B_2 := B_2 \cup \{v\}$;
  end if
  if v has formal neighbor in B_2 only **then**
    $B_1 := B_1 \cup \{v\}$;
  end if
  if v has formal neighbors in B_1 and B_2 **then**
    ;
  end if
  $G := G - \{v\}$;
  remove isolated vertices from G;
end while
Output B_1, B_2;

A random variable $\mathrm{bi}_{\mathrm{alg}}(G)$ denotes the number of vertices of the bipartite subgraph produced by BIPGREEDY.

The main result of this section is Theorem 4.16 which states that the algorithm finds a bipartite graph of size at least $(0.75-\varepsilon)n$ on average, i.e. $\mathbf{E}[\mathrm{bi}_{\mathrm{alg}}] \geq (0.75-\varepsilon)n$ (where n is the number of vertices of the input graph and n is big). It is an easy consequence (Corrolary 4.17) that $\mathrm{bi}(G) > (0.75-\varepsilon)n$ a.a.s. for $G \in_u \mathcal{G}^*_{3,n}$.

Analogously, Frieze and Suen proved that MINGREEDY finds an independent set of size at least $(6\ln(3/2) - 2 - \varepsilon)n \doteq 0.4327n$ on average. This means that $\alpha(G) > (6\ln(3/2) - 2 - \varepsilon)n$ a.a.s. for $G \in_u \mathcal{G}^*_{3,n}$.

Both algorithms run in linear time. In the next section we make a short excursion

[1] a *formal neighbor* of a vertex v is a vertex $u \in V(G(0))$ such that $\{u,v\} \in E(G(0))$ and $u \notin V(G(t))$ (where t is the current time)

into complexity theory. We show that finding a maximum induced bipartite subgraph is computationally a hard problem.

4.1 Complexity of finding a maximum induced bipartite subgraph

Decision problem INDEPENDENT SET(G, k) asks whether there is an independent set of size at least k in cubic graph G. Decision problem BIPARTITE SUBGRAPH(G, k) asks whether there is an induced bipartite graph of size at least k in cubic graph G.

Theorem 4.1. *(a) INDEPENDENT SET is NP-complete.*

(b) BIPARTITE SUBGRAPH is NP-complete.

Proof. (a) This is shown in [GJS76] by reduction from 3SAT.

(b) Obviously, BIPARTITE SUBGRAPH $\in$ NP. We make a polynomial time reduction from NAE 3SAT. NAE 3SAT is a problem which gets a boolean formula ϕ such, that ϕ is in conjunctive normal form and in each elementary disjunction at most three literals are used, on the input. The question is, whether it is possible to assign the variables values in such a way, that the formula is satisfied and there is no a clause in which all three literals are set to 1. It is well known that NAE 3SAT is NP-complete.

Let ϕ be an input formula for NAE 3SAT. First observe that we may assume that all the clauses in the input formula consist of exactly three literals (with possible repetition of the literals in clauses). Let α_x be the number of occurrences of variable x, β_x be the number of its negations $\neg x$ in the formula, $\gamma_x = \max(\alpha_x, \beta_x)$. We may assume that $\gamma_x > 1$ for every variable x. If this was not true, we can work with equivalent problem of deciding whether $\phi \wedge \phi \in$ NAE 3 SAT. Now we describe how we construct a graph G from formula ϕ. For each variable x we construct a cycle C_x of length $2\gamma_x$ and we properly color vertices of C_x with red and blue. For each clause L we construct a triangle T_L. Each vertex in the triangle represents one literal of the clause. We connect sequentially vertices of triangles T_L to vertices of cycles C_x. If the literal is variable x (resp. negation $\neg x$), we connect it with a red (resp. blue) vertex of C_x which still has degree two. After this, we all the vertices of the original triangles T_L have degree three. Vertices of the original cycles C_x are of degree two or three. We add gadgets depicted on Figure 4.1 to those of degree two.

Let c be the number of clauses. We shall show that for a formula ϕ (with assumptions as above), $\phi \in$ NAE 3SAT if and only if $(G, 2c + 2\sum_x \alpha_x + \beta_x + 6|\alpha_x - \beta_x|) \in$ BIPARTITE SUBGRAPH.

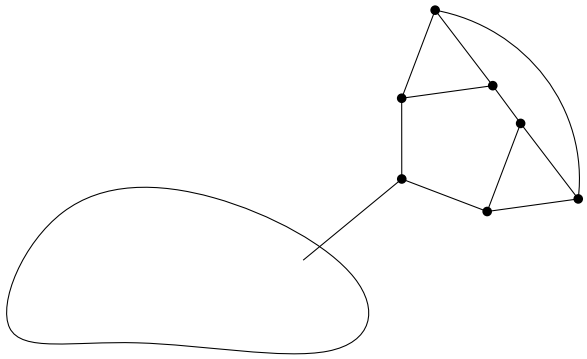


FIGURE 4.1: Gadgets added to vertices of degree two

Let $\phi \in$ NAE 3SAT. We find an induced bipartite subgraph of G of size $2c + 2\sum_x \alpha_x + \beta_x + 6|\alpha_x - \beta_x|$ as follows. Let $X = (x_1, x_2, \ldots, x_n)$ be the satisfying assignment of the values to the variables. In each clause L we pick one literal t_L which is 1 with respect to X and one literal f_L which is 0 with respect to X. Vertices of part B_1 of the bipartite graph will be all the blue vertices, vertices which are corresponding to literals t_L and some of the vertices of the gadgets. Vertices of part B_2 of the bipartite graph will all the red vertices, vertices which are corresponding to literals f_L and some of the vertices of the gadgets. The distribution of the vertices of the gadget into the bipartite graph is shown on Figure 4.2. It is easy to check that we get an induced bipartite graph of demanded size.

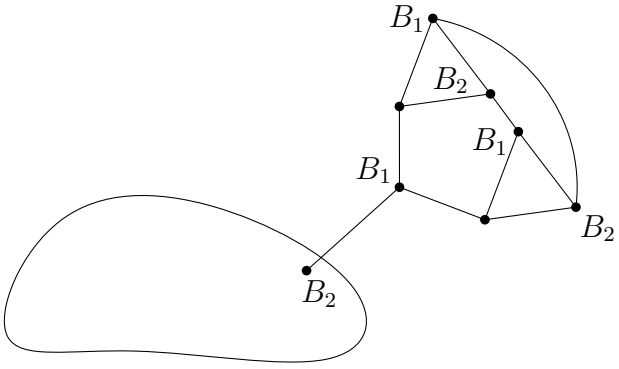


FIGURE 4.2: Distribution of the vertices of a gadget into the parts of the bipartite graph. The case when the gadget is attached to a vertex in part B_2 is depicted. Analogous assignment can be done if the associated vertex is in part B_1.

On the other hand, knowing that there is an induced bipartite subgraph of G of size $2c+2\sum_x \alpha_x+\beta_x+6|\alpha_x-\beta_x|$, we know that $\phi \in$ NAE 3SAT. This follows quite straightforward by a reverse argument. The only thing one has to observe is that there are always at most six vertices of each gadget present in the bipartite graph.

An example of the reduction is shown on Figure 4.3.

□

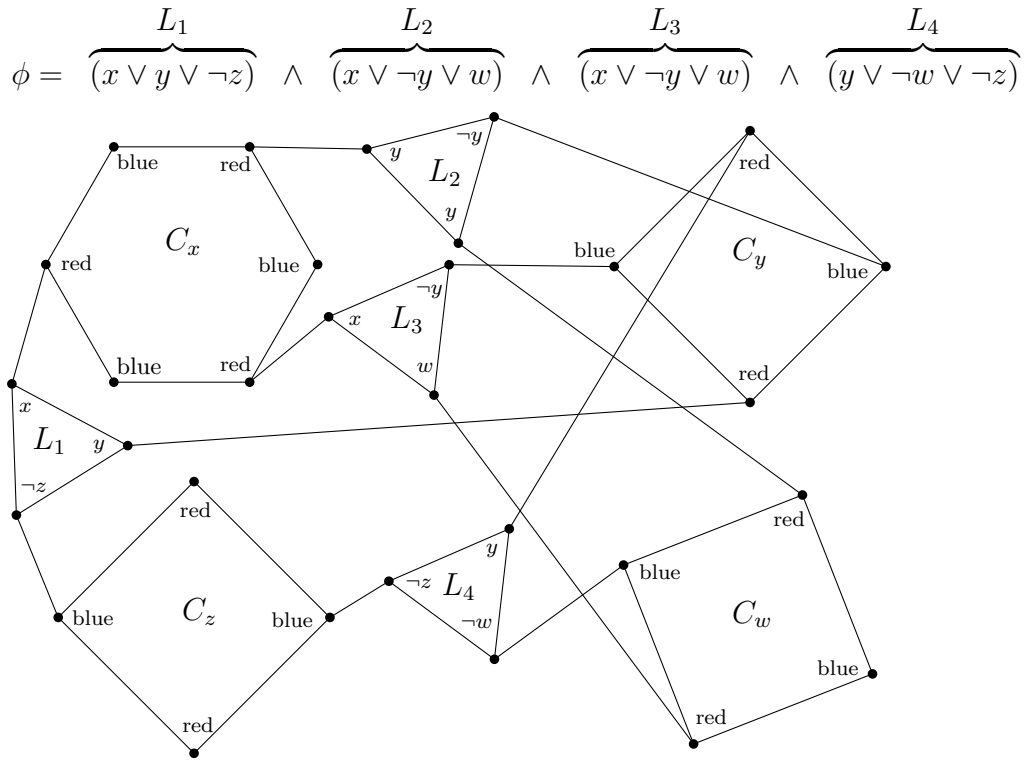


FIGURE 4.3: An example of reduction of NAE 3SAT to BIPARTITE SUBGRAPH. Gadgets which are to be added to vertices of degree two are not depicted for lucidity.

The author is not acquainted with any result concerning approximability of the problem. Thanks to Peter Golovach for discussion on the subject and for appointing the idea of the proof of Theorem 4.1.

4.2 Computer simulation

The following table shows mean value and variance of bipartite ratio $\frac{\mathrm{bi}_{\mathrm{alg}}(G)}{n}$ found by computer. For each number of vertices 1,000 or 10,000 random cubic graphs were generated and for each of the graphs one Monte Carlo run was executed.

number of vertices	1,000	10,000	100,000	1,000,000
number of graphs generated	10,000	10,000	10,000	1,000
$\mathbf{E}[\frac{\mathrm{bi}_{\mathrm{alg}}}{n}]$	.81681	.81778	.81796	.81798
$\mathbf{Var}[\frac{\mathrm{bi}_{\mathrm{alg}}}{n}]$	5.59E-05	5.58E-06	5.36E-07	5.58E-08

4.3 Probabilistic analysis

By probabilistic analysis we mean that we investigate how the algorithm performs typically, i.e. given a random cubic graph as an input. We emphasize that we do not analyze how the algorithm behaves given a fixed graph on the input.

The probability space used for the analysis will be product space $\Omega_n = \mathcal{G}^*_{3,n} \times (\{0,1\}_u)^{\mathbb{N}}$. Term $\mathcal{G}^*_{3,n}$ corresponds to the input we get, term $(\{0,1\}_u)^{\mathbb{N}}$ represents a random number generator which is used by the algorithm. We promised in Preliminaries that only countable probability spaces will be used; which Ω_n is not. However, it can be checked from the design of the algorithm that for a fixed number of vertices only a bounded number of bounded random numbers is needed, so we may use $\Omega_n = \mathcal{G}^*_{3,n} \times (\{0,1\}_u)^{f(n)}$ (for some function f) as well.

In the analysis we investigate behavior of several chains defined on the input graph. One step in the chains is a run of the '**while** $G \neq \emptyset$ **do**' block. We number the steps starting from zero. We denote $G(t)$ the graph (in the sense of the pairing model) at the t-th step of the algorithm. By $N_1(t), N_2(t), N_3(t)$ we denote the number of vertices of degree 1, 2, 3 in graph $G(t)$, $M(t)$ denotes the number of edges in the graph. We set $p_i(t) = iN_i(t)/(2M(t))$ for $i = 1, 2, 3$ and $N(t) = N_1(t) + N_2(t) + N_3(t)$.

Overview of the analysis. The first step in the analysis is to show (Lemma 4.5) that during the run of the algorithm (up to a minor time interval near the end) it holds with very high probability that

$$N_3(t) \doteq \frac{2^{3/2} M(t)^{3/2}}{3^{3/2} n^{1/2}} \tag{4.1}$$

We give a simple reason here which should show that this is plausible. We liken the stochastic process to a differential equation. This method was used several times in analysis of randomized greedy algorithms before, some other applications can be found in [Wor99a]. It is easy to show, that it is sufficient to analyze the run only for connected graphs. Then, except the first step, vertex of degree three in never removed. The first step has only little impact on values $M(t)$, $N_3(t)$ and hence may be omitted from the analysis. We think of the process as a sequence of step-by-step edge removals. Let the graph have at certain time τ (τ does not have to be an integer—we subdivided original steps; each step is divided so that M decreases by 1 in each substep) $M(\tau)$ edges and $N_3(\tau)$ vertices of degree three. When vertex v was picked as a vertex of minimum degree, $\deg v < 3$ and edge $\{u, v\}$ is currently being removed, the number of vertices of degree three decreases if and only if $\deg u = 3$. The probability that this happens is $p_3(\tau) + O(1/M(t))$. The differential equation which approximately describes this is

$$\frac{\mathbf{d}N_3}{\mathbf{d}M} = \frac{3N_3}{2M}$$

Solving it with initial condition $M(0) = \frac{3}{2}n$, $N_3(0) = n$ we get (4.1).

In Lemma 4.8 and Lemma 4.9 we show what is the probability that a vertex of degree 2 at certain time t is picked by the algorithm, namely it is very close to $p_3(t)/(2-p_3(t))$. This is done by examining the transition probabilities for states $N_1 = 0$, $N_1 = 1$, $N_1 = 2$ (which are the only three possible values of N_1 as shown in Proposition 4.6). We approximate

the process $\{N_1(t)\}_t$ by a Markov chain with transition matrix P,

$$P = \begin{pmatrix} p_3(t)^2 & 2p_2(t)p_3(t) & p_2(t)^2 \\ p_3(t) & p_2(t) & 0 \\ 0 & p_3(t) & p_2(t) \end{pmatrix}$$

The stationary distribution of the chain is

$$\overline{\pi} = \left(\frac{p_3(t)}{2 - p_3(t)}, \frac{1 - p_3(t)^2}{2 - p_3(t)}, \frac{(1 - p_3(t))^2}{2 - p_3(t)} \right)$$

Lemmas 4.10 and 4.11 give estimates on how many steps it takes to remove certain number of edges. The estimates are straightforward since we now know the ratio between how many times a vertex of degree one and degree two is picked. The two lemmas give as a consequence Lemma 4.12 which states how many vertices degree 1 are picked during certain intervals on average.

In Lemma 4.15 we show that the number of vertices in $G(t)$ which have a neighbor already placed in B_1 and the number of vertices in $G(t)$ which have a neighbor already placed in B_2 are likely to be similar. We denote the numbers of vertices of degree 2 in $G(t)$ having neighbors in B_1 and B_2 by $K_1(t)$ and $K_2(t)$, respectively. The proof of Lemma 4.15 is somewhat technical but the idea is simple. We shall sketch here why it is not probable (in the rigorous proof we give bound on the probability $O(1/n^2)$) that there exists time t_1 such that

$$K_1(t_1) - K_2(t_1) > \varepsilon n$$

Let t_1 be any such time. Let t_0 be the first time such that $K_1(t) - K_2(t) > \varepsilon n/2$ for all $t \in [t_0, t_1]$. Since $n \to \infty$, $t_1 - t_0$ must be big. Let $t \in [t_0, t_1]$ be any time.

(a) If vertex $v(t)$ chosen at time t has degree 2, then it is more likely that its neighbor is in B_1. If so then $v(t)$ will be inserted to B_2 and we have $\Delta K_1 = -1$ at the moment. There are three cases according to the degrees of the neighbors of $v(t)$

- Both neighbors have degree 2. Then $\Delta K_2 = 2$.
- One neighbor has degree 2 and one has degree 1. Then $\Delta K_2 = 1$.
- Both neighbors have degree 2. Then $\Delta K_2 = 0$.

When $v(t)$ has a neighbor in B_2 we get a symmetric situation in terms of ΔK_1 and ΔK_2 (with the same probabilities of corresponding opposite events). Hence we have that $\Delta(K_1 - K_2) < 0$ on average.

(b) If a vertex of degree 1 with a neighbor of degree 1 is picked, then $\Delta(K_1 - K_2) = 0$.

(c) When it happens that $v(t)$ had degree 1 and in the next step there are no vertices of degree 1 in $G(t+1)$, the situation gets more complicated. One has to trace back the 1-chain (1-chain is a sequence of steps in which the algorithm chose a vertex of degree 1, such that each vertex in the 1-chain neighbors with the consecutive

one) to its beginning and find how it emerged[2]. It turns out that $\Delta(K_1 - K_2) > 0$ on average, but still alternation of cases (a), (b) and (c) (in ratio given by the probabilities of the cases) gives total outcome $\Delta(K_1 - K_2) < 0$.

Since $t_1 - t_0$ is big, we have that the 'on average' stuff really behaves how expected and thus the initial disproportion $K_1(t_0) - K_2(t_0) > \varepsilon n/2$ was much more likely to decrease then to increase.

Lemma 4.15 was the last step needed in our analysis. In Theorem 4.16 we show that BIPGREEDY returns an induced bipartite subgraph of size at least $(0.75 - \varepsilon)n$ on average. This is done by showing that at most $(0.25 + \varepsilon)n$ vertices are not included in the bipartite graph. A vertex is not included in the bipartite graph either because it had degree 0 or because it was neighboring with vertices of the both parts of the bipartite graph. We show in Lemma 4.7 that there are at most $\ln n$ degree-zero vertices with high probability. To cope with vertices which have formal neighbors in the both parts of the bipartite graph, Lemma 4.12 and Lemma 4.15 are used. By Lemma 4.12 we know how many vertices of degree 1 were picked at each stage of the algorithm, by Lemma 4.15 we know that the case that both formal neighbors of any such vertex were in the same part of the bipartite graph is nearly the same as that the neighbors were in different parts. This suffices to prove the theorem.

The following lemma and its consequence stated in Lemma 4.3 are essential for any further analysis. We note that the statements are not true when a random graph at the input is replaced by a fixed one, which is the reason why we are not able to give analysis of performance of BIPGREEDY for a fixed graph.

Symbol $RS(i)$ denotes a distribution of a random uniform subset of size i of set $\{1, 2, 3\}$, $RS(i) \sim \binom{\{1,2,3\}}{i}_u$.

Lemma 4.2. *In Ω_n, given $(N_1(t), N_2(t), N_3(t))$ at any time t, $G(t)$ has the same distribution as a random matching on set $S \subseteq \{1, 2, \ldots, n\} \times \{1, 2, 3\}$, $S = V_1 \cup V_2 \cup V_3$, where V_1, V_2, V_3 were obtained in the following way. Take a triple of disjoint sets $W_1, W_2, W_3 \subseteq \{1, 2, \ldots, n\}$, $|W_i| = N_i(t)$ randomly uniformly from the set of all such triples. Then define $V_i = \bigcup_{w \in W_i} \{w\} \times X(w)$, where $X(w) \overset{ind}{\sim} RS(i)$, $i = 1, 2, 3$.*

Proof. A random graph $G \in_u \mathcal{G}^*_{3,n}$ can be constructed as follows. At the beginning we take set $\{1, 2, \ldots, n\} \times \{1, 2, 3\}$ of all elements of the future matching. We repeatedly pick an arbitrary (random, or given by a certain rule, for example) unmatched element x and match it with an element which was chosen randomly uniformly from the set of currently unmatched elements different from x.

We shall construct a random graph G using this procedure. We start with unmatched elements $F = \{1, 2, \ldots, n\} \times \{1, 2, 3\}$. At each step we pick a vertex $v = (v^1, v^2, v^3)$ in G

[2]Well, one cannot expect any results tracing the algorithm backwards; one can only trace it forward and use essential Lemma 4.2. The description given here should however give some feeling what we will be trying for later on.

uniformly at random such that the degree of v is the maximum degree over all vertices of G except vertices of degree three. Out of v^1, v^2, v^3 there are one, two or three elements which are currently unpaired. We take the unpaired elements in sequence and for each element we take a random uniform unpaired element in F and establish a pairing between these two elements.

The described way of generating a random graph was not arbitrary; it is an easy observation that we constructed the graph in reverse to step-by-step dismounting the edges by BIPGREEDY. The elements of the graph which remained paired at a certain step during the run of BIPGREEDY are exactly those which are yet to be paired in the parallel construction and thus the lemma holds. □

Lemma 4.3. *In Ω_n, $\{(N_1(t), N_2(t), N_3(t))\}_t$ is a Markov chain with initial state $(0, 0, n)$.*

Proof. Follows directly from Lemma 4.2. □

We say that graph G is *nearly-connected* if G is connected or has exactly two components, one of which has at most two vertices.

Proposition 4.4. $\mathbf{Pr}[G \in_u \mathcal{G}^*_{3,n}$ *is nearly-connected*$] = 1 - O(1/n^2)$.

Proof.

$$\begin{aligned}
&\mathbf{Pr}[G \in_u \mathcal{G}^*_{3,n} \text{ is not nearly-connected}] \\
&\leq \sum_{\substack{A \subseteq V(G) \\ 4 \leq |A| \leq n-4}} \mathbf{Pr}[\text{there are no edges between } A \text{ and } V(G) \setminus A] \\
&= \sum_{k=4}^{n-4} \binom{n}{k} \cdot \frac{\mathrm{f}(3k)\,\mathrm{f}(3(n-k))}{\mathrm{f}(3n)} \\
&= O\left(\frac{1}{n^2}\right)
\end{aligned}$$

□

Lemma 4.5. *For any $\varepsilon > 0$,*

$$\mathbf{Pr}\left[\exists t : M(t) > n^{1/2} \ln^3 n \text{ and } \left|\frac{3^{3/2} N_3(t) n^{1/2}}{2^{3/2} M(t)^{3/2}} - 1\right| > \varepsilon\right] = O\left(\frac{1}{n^2}\right) \tag{4.2}$$

Proof. By Proposition 4.4 we may suppose that our graph G is nearly-connected. If G was not connected (i.e. it consisted of two components H_1 and H_2, H_1 having two vertices) we let the algorithm run on G (we denote the chains $N_{3,G}(t), M_G(t)$) and on H_2 (the chains are $N_{3,H_2}(t), M_{H_2}(t)$). Then obviously:

$$\begin{aligned}
N_{3,H_2}(t) \leq N_{3,G}(t) \leq N_{3,H_2}(t) + 2 \qquad &\text{in distribution} \\
M_{H_2}(t) \leq M_G(t) \leq M_{H_2}(t) + 3 \qquad &\text{in distribution}
\end{aligned}$$

so the examined ratio changes only little by replacing chains corresponding to G with ones corresponding to H_2. Hence it is sufficient to prove the result only for connected graphs.

We make a finer division of steps made by the algorithm: each step is one edge removal. The refined chains $M(t), N_3(t)$ will be denoted by $\widetilde{M}(\tau), \widetilde{N_3}(\tau)$.

Suppose that in a certain step τ a vertex $v \in V_{\min}$ was picked and edge $e = \{u, v\}$ is removed. Except the case $\tau = 0$ it holds $\deg v < 3$. The probability that u is of degree three is

$$p(\tau) = \frac{3\widetilde{N_3}(\tau) + O(1)}{2\widetilde{M}(\tau)} \tag{4.3}$$

Let $h = \lfloor n^{1/4} \rfloor$, for $i = 0, 1, \ldots, \lfloor \frac{3}{2} n^{3/4} \rfloor$ define $m_i = \frac{3}{2}n - ih$. Let z_i be the number of vertices of degree three in G at the first time $\widehat{\tau}_i$ when $\widetilde{M}(\widehat{\tau}_i) \leq m_i$ and let $\mathcal{E}_i$ be the event that

$$\frac{3^{3/2} z_i n^{1/2}}{2^{3/2} m_i^{3/2}} - 1 = O\left(\sum_{j=0}^{i-1} \frac{n^{3/8} \ln^{1/2} n}{m_j^{5/4}} \right) \tag{4.4}$$

We first prove that for i such that $m_i \geq n^{1/2} \ln^3 n$,

$$\mathbf{Pr}\left[\bigcap_{j \leq i} \mathcal{E}_j \right] = 1 - O\left(\frac{i}{n^4} \right) \tag{4.5}$$

Note that the O term on the right hand-side of (4.4) tends to zero as n goes to infinity, since

$$\begin{aligned}
&\sum_{\substack{j \\ 3/2 n - j n^{1/4} \geq n^{1/2} \ln^3 n}} \frac{n^{3/8} \ln^{1/2} n}{(\frac{3}{2}n - j n^{1/4})^{5/4}} \\
&\leq n^{1/16} \ln^{1/2} n \sum_{\substack{j \\ \frac{3}{2} n - j n^{1/4} \geq n^{1/2} \ln^3 n}} \frac{1}{\left(\frac{3}{2} n^{3/4} - j\right)^{5/4}} \\
&= n^{1/16} \ln^{1/2} n \sum_{n^{1/4} \ln^3 n \leq j \leq \frac{3}{2} n^{3/4}} \frac{1}{j^{5/4}} \\
&= (1 + o(1)) n^{1/16} \ln^{1/2} n \int_{n^{1/4} \ln^3 n}^{\frac{3}{2} n^{3/4}} \frac{1}{x^{5/4}} \mathbf{d}x \\
&= O\left(\frac{1}{\ln^{1/4} n} \right)
\end{aligned}$$

Thus proving (4.5) would imply that for any $\varepsilon' > 0$ it holds that

$$\mathbf{Pr}\left[\exists i : m_i > n^{1/2} \ln^3 n \text{ and } \left| \frac{3^{3/2} z_i n^{1/2}}{2^{3/2} m_i^{3/2}} - 1 \right| > \varepsilon' \right] = O\left(\frac{1}{n^3} \right) \tag{4.6}$$

We prove (4.5) by induction. The equation is true for $i = 0$. Assume that the hypothesis holds for $i-1$ and $m_i \geq n^{1/2} \ln^3 n$. We shall show that

$$\mathbf{Pr}\left[\overline{\mathcal{E}_i} \mid \mathcal{E}_{i-1}\right] = O\left(\frac{1}{n^4}\right) \tag{4.7}$$

Denote $\Delta z_i = z_{i-1} - z_i$. Until the end of the proof of induction step we condition on $\mathcal{E}_{i-1}$. We substitute into (4.3) and get for all τ such that $\tau_{i-1} \leq \tau \leq \tau_i$:

$$p(\tau) = \frac{3z_{i-1} + O(h)}{2m_{i-1} + O(h)} = \frac{3z_{i-1}}{2m_{i-1}}\left(1 + O\left(\frac{n^{3/4}}{m_{i-1}^{3/2}}\right)\right)$$

Define

$$p_{\text{approx}} = \frac{3z_{i-1}}{2m_{i-1}}$$

Variables X_τ (for $\tau = \widehat{\tau_{i-1}}, \ldots, \widehat{\tau_i} - 1$) are Bernoulli variables with success probability very close to p_{approx}. Thus we may expect a Chernoff-type concentration result for $\Delta z_i = \sum_{\tau=\widehat{\tau_{i-1}}}^{\widehat{\tau_i}-1} X_\tau$. But variables X_τ are moderately dependent; $p(\tau)$ depends on $p(\widehat{\tau_{i-1}})$ and $\sum_{j=\widehat{\tau_{i-1}}}^{\tau-1} X_j$. Janson-Suen or Talagrand inequality seem not to help in this case. What helps is to bound each X_t by X_t^{upper} from above and by X_t^{lower} from below, introduce variables $\Delta z_i^{\text{upper}} = \sum_{\tau=\widehat{\tau_{i-1}}}^{i} X_\tau^{\text{upper}}$ and $\Delta z_i^{\text{lower}} = \sum_{\tau=\widehat{\tau_{i-1}}}^{i} X_\tau^{\text{lower}}$ and then to show that $\Delta z_{\widehat{\tau_i}-1}^{\text{upper}}$ and $\Delta z_{\widehat{\tau_i}-1}^{\text{lower}}$ are concentrated variables with expectations close to each the other.

Let K be such that

$$p(\tau) \in \left[\frac{3z_{i-1}}{2m_{i-1}}\left(1 - K\left(\frac{n^{3/4}}{m_{i-1}^{3/2}}\right)\right), \frac{3z_{i-1}}{2m_{i-1}}\left(1 + K\left(\frac{n^{3/4}}{m_{i-1}^{3/2}}\right)\right)\right]$$

Denote

$$\begin{aligned} p^{\text{upper}} &= \frac{3z_{i-1}}{2m_{i-1}}\left(1 + K\left(\frac{n^{3/4}}{m_{i-1}^{3/2}}\right)\right) \\ p^{\text{lower}} &= \frac{3z_{i-1}}{2m_{i-1}}\left(1 - K\left(\frac{n^{3/4}}{m_{i-1}^{3/2}}\right)\right) \end{aligned}$$

For each $\tau = \widehat{\tau_{i-1}}, \ldots, \widehat{\tau_i} - 1$ define $X_\tau^{\text{upper}} \overset{\text{ind}}{\sim} Be(p^{\text{upper}})$ and $X_\tau^{\text{lower}} \overset{\text{ind}}{\sim} Be(p^{\text{lower}})$. Now X_τ^{upper} and X_τ^{lower} are independent random variables, for which

$$\begin{aligned} X_\tau^{\text{lower}} &\leq X_\tau \leq X_\tau^{\text{upper}} && \text{in distribution} \\ \Delta z_i^{\text{lower}} &\leq \Delta z_i \leq \Delta z_i^{\text{upper}} && \text{in distribution} \end{aligned}$$

The expected values of the sums bounding Δz_i are

$$\mathbf{E}[\Delta z_i^{\text{upper}}] = hp^{\text{upper}} = \frac{3n^{1/4} z_{i-1}}{2m_{i-1}}\left(1 + K\left(\frac{n^{3/4}}{m_{i-1}^{3/2}}\right)\right) \tag{4.8}$$

$$\mathbf{E}[\Delta z_i^{\text{lower}}] = hp^{\text{lower}} = \frac{3n^{1/4} z_{i-1}}{2m_{i-1}}\left(1 - K\left(\frac{n^{3/4}}{m_{i-1}^{3/2}}\right)\right) \tag{4.9}$$

We apply Chernoff Inequality 2.2 to $\Delta z_i^{\text{upper}}$. We shall use that $\mathbf{Var}[\Delta z_i^{\text{upper}}] \leq \mathbf{E}[\Delta z_i^{\text{upper}}]$.

$$\begin{aligned}
&\mathbf{Pr}\left[|\Delta z_i^{\text{upper}} - \mathbf{E}[\Delta z_i^{\text{upper}}]| \geq \sqrt{10hp^{\text{upper}} \ln n} \;\middle|\; \mathcal{E}_{i-1}\right] \\
&\leq \; 2\exp\left(-\frac{10hp^{\text{upper}} \ln n}{\mathbf{Var}[\Delta z_i^{\text{upper}}] + \sqrt{10hp^{\text{upper}} \ln n}/3} \cdot (1+o(1))\right) \\
&\leq \; 2\exp\left(-\frac{10hp^{\text{upper}} \ln n}{2hp^{\text{upper}}} \cdot (1+o(1))\right) = O\left(\frac{1}{n^4}\right)
\end{aligned}$$

Similar computations can be carried out for $\Delta z_i^{\text{lower}}$.

Combining (4.8) and (4.9) with the concentration results we get that, with probability $1 - O(1/n^4)$, it holds:

$$\frac{\Delta z_i}{z_{i-1}} = \frac{3n^{1/4}}{2m_{i-1}} + O\left(\frac{n}{m_{i-1}^{5/2}}\right) + O\left(\frac{n^{1/8}\ln^{1/2} n}{m_{i-1}^{1/2} z_{i-1}^{1/2}}\right) = \frac{3n^{1/4}}{2m_{i-1}} + O\left(\frac{n^{3/8}\ln^{1/2} n}{m_{i-1}^{5/4}}\right)$$

We shall verify (4.4):

$$\begin{aligned}
z_i = &\; z_{i-1} - \Delta z_i \\
= &\; z_{i-1}\left(1 - \frac{3n^{1/4}}{2m_{i-1}} + O\left(\frac{n^{3/8}\ln^{1/2} n}{m_{i-1}^{5/4}}\right)\right)
\end{aligned}$$

We use approximation by Generalized binomial formula

$$\left(\frac{m_i}{m_{i-1}}\right)^{3/2} = 1 - \frac{3n^{1/4}}{2m_{i-1}} + O\left(\frac{n^{1/2}}{m_{i-1}^2}\right)$$

and get

$$z_i = z_{i-1}\left(\frac{m_i}{m_{i-1}}\right)^{3/2}\left(1 + O\left(\frac{n^{3/8}\ln^{1/2} n}{m_{i-1}^{5/4}}\right)\right)$$

Substituting z_{i-1} using (4.4) gives us the desired relation for z_i

$$\begin{aligned}
z_i = &\; \left(1 + O\left(\sum_{j=0}^{i-2} \frac{n^{3/8}\ln^{1/2} n}{m_j^{5/4}}\right)\right) \cdot \frac{2^{3/2} m_i^{3/2}}{3^{3/2} n^{1/2}} \cdot \left(1 + O\left(\frac{n^{3/8}\ln^{1/2} n}{m_{i-1}^{5/4}}\right)\right) \\
= &\; \left(1 + O\left(\sum_{j=0}^{i-1} \frac{n^{3/8}\ln^{1/2} n}{m_j^{5/4}}\right)\right) \cdot \frac{2^{3/2} m_i^{3/2}}{3^{3/2} n^{1/2}}
\end{aligned}$$

and proof of (4.5) is completed.

What only remains to be shown is that the examined ratio is almost constant everywhere, not only in node points $\widehat{\tau}_i$. Let $\varepsilon > 0$ be arbitrary. Assume that for all i such that $m_i \geq n^{1/2}\ln^3 n$ it holds

$$\left|\frac{3^{3/2} z_i n^{1/2}}{2^{3/2} m_i^{3/2}} - 1\right| < \frac{\varepsilon}{2}$$

Let t be such, that $M(t) \geq n^{1/2} \ln^3 n$. (Actually, we need a stronger hypothesis on t, $(*)$ $M(t) \geq n^{1/2} \ln^3 n + n^{1/4}$, but this is only a technical detail. The proof of (4.6) would also work weakening the condition on the number of edges to $m_i \geq 0.9 n^{1/2} \ln^3 n$ instead of $m_i \geq n^{1/2} \ln^3 n$; and then $(*)$ holds.) Let $\widehat{\tau_i}$ and $\widehat{\tau_{i+1}}$ be the neighboring node points, $m_i \leq M(t) \leq m_{i-1}$, $z_i \leq N_3(t) \leq z_{i-1}$. We shall only show that

$$\frac{3^{3/2} N_3(t) n^{1/2}}{2^{3/2} M(t)^{3/2}} - 1 < \varepsilon$$

The other inequality is analogous.

$$\begin{aligned} \frac{3^{3/2} N_3(t) n^{1/2}}{2^{3/2} M(t)^{3/2}} - 1 &\leq \frac{3^{3/2} z_{i-1} n^{1/2}}{2^{3/2} m_i^{3/2}} - 1 = \frac{3^{3/2} z_{i-1} n^{1/2}}{2^{3/2} m_{i-1}^{3/2}} \cdot \left(\frac{m_{i-1}}{m_i} \right)^{3/2} - 1 \\ &= \frac{3^{3/2} z_{i-1} n^{1/2}}{2^{3/2} m_{i-1}^{3/2}} \cdot (1 + o(1)) - 1 \leq \left(1 + \frac{\varepsilon}{2} \right) + o(1) \end{aligned}$$

□

Equivalent way of writing $\left| \frac{3^{3/2} N_3(t) n^{1/2}}{2^{3/2} M(t)^{3/2}} - 1 \right| \leq \varepsilon$ is $\left| p_3(t) - \sqrt{\frac{2M(t)}{3n}} \right| \leq \varepsilon \sqrt{\frac{2M(t)}{3n}}$. Thus Lemma 4.5 provides us with a good estimate for $p_3(t)$ depending on the current number of edges $M(t)$.

Note that the proof of Lemma 4.5 could be strengthened; the error probability could be $O(1/n^3)$ instead of $O(1/n^2)$. The only thing one has to do is to weaken the notion of nearly-connectedness.

Proposition 4.6. *$N_1(t) \leq 2$ for all t.*

Proof. By induction on t. For $t = 0$ the number of vertices of degree one is zero.

Suppose by induction hypothesis that $N_1(t-1) \leq 2$ for some $t > 0$. We distinguish two cases:

- $N_1(t-1) = 0$. The minimum degree of the graph is either two or three. In each of these cases not more than two vertices of degree one emerge.
- $N_1(t-1) > 0$. In step t a vertex of degree one is picked (and deleted from the graph) and at most one vertex of degree one emerges. The number of vertices of degree one does not increase.

□

We denote by J_ε a random variable which is the number of times such that $N_1(t) = 2$ and $N_1(t+1) = 0$ in time interval where $p_2(t), p_3(t) > \varepsilon$, $N(t) > \varepsilon n$. Variable J_ε counts how many times a vertex of degree one, such that its only neighbor has degree one, was picked.

Lemma 4.7. *Let $\varepsilon > 0$ be fixed. Then*

$$\mathbf{Pr}\left[J_\varepsilon(G) > \ln n;\ G \in_u \mathcal{G}^*_{3,n}\right] = O\left(\frac{1}{n^2}\right)$$

Proof. Probability that we picked at a certain step t (with restriction on t given by the hypotheses) a vertex of degree one, such that its only neighbor has degree one is at most $\frac{1}{\varepsilon n}$. The probability that this happens in at least $\ln n$ steps can be estimated by

$$\mathbf{Pr}\left[J_\varepsilon(G) > \ln n;\ G \in_u \mathcal{G}^*_{3,n}\right] \leq \binom{n}{\lceil \ln n \rceil}\left(\frac{1}{\varepsilon n}\right)^{\ln n} \leq \left(\frac{en}{\ln n}\right)^{\lceil \ln n \rceil}\left(\frac{1}{\varepsilon n}\right)^{\ln n} = O\left(\frac{1}{n^2}\right)$$

□

Lemma 4.8. *Let $\varepsilon > 0$ and $k \in \mathbb{N}$ be fixed. With probability $1 - O(1/n^2)$ the following holds.*

For any time t_0 such that

- *$p_2(t_0), p_3(t_0) > 2\varepsilon$,*
- *$N(t_0) > 2\varepsilon n$, and*
- *a vertex of degree one with neighbor of degree one is never picked by the algorithm in time interval $[t_0, t_0 + (k+1)\ln n]$.*

it holds that

$$\mathbf{Pr}[N_1(t) = 0] \leq (1+\varepsilon) \cdot \frac{p_3(t_0)}{2 - p_3(t_0)}$$

for any $t \in [t_0 + \ln n, t_0 + (k+1)\ln n]$.

Proof. We may suppose (for the same reason as in the proof of Lemma 4.5) that G is connected. This does not occur with probability $O(1/n^2)$.

By Proposition 4.6, $N_1 \leq 2$. We examine transition probabilities of states $N_1 = 0, N_1 = 1, N_1 = 2$. At time $t > 0$ they are given by the matrix

$$P(t) = \overbrace{\begin{matrix} N_1 = 0 & N_1 = 1 & N_1 = 2 \end{matrix}}^{\text{}} \left.\begin{pmatrix} p_3(t)^2 + o & 2p_2(t)p_3(t) + o & p_2(t)^2 + o \\ p_3(t) + o & p_2(t) + o & 0 \\ 0 & p_3(t) + o & p_2(t) + o \end{pmatrix}\right\} \begin{matrix} N_1 = 0 \\ N_1 = 1 \\ N_1 = 2 \end{matrix}$$

where symbol o is an abbreviation for $O(1/M(t))$.

At each step of the algorithm both p_2 and p_3 increase by at most $2/N(t)$. This means that for every time t between t_0 and $t_0 + \Delta t$ it holds that $p_2(t) \leq p_2(t_0) + 2\Delta t/N(t_0 + \Delta t)$, $p_3(t) \leq p_3(t_0) + 2\Delta t/N(t_0 + \Delta t)$.

Set $\Delta_1 t = \lceil \ln n \rceil$, $\Delta_2 t = k\lceil \ln n \rceil$. By $\pi(t)$ we denote the distribution of states $N_1 = 0, N_1 = 1, N_1 = 2$ at time t (so $\mathbf{Pr}[N_1(t) = 0] = \pi^{(N_1=0)}(t)$). We shall show

that, irrespectively to the starting distribution $\pi(t_0)$, $\pi(t)$ is very close to the stationary distribution of a Markov chain given by transition matrix P_0,

$$P_0 = \begin{pmatrix} p_3(t_0)^2 & 2p_2(t_0)p_3(t_0) & p_2(t_0)^2 \\ p_3(t_0) & p_2(t_0) & 0 \\ 0 & p_3(t_0) & p_2(t_0) \end{pmatrix}$$

in time interval $t \in [t_0 + \Delta_1 t, t_0 + \Delta_2 t]$. Note that the stationary distribution of the Markov chain is

$$\overline{\pi} = \left(\frac{p_3(t_0)}{2 - p_3(t_0)}, \frac{1 - p_3(t_0)^2}{2 - p_3(t_0)}, \frac{(1 - p_3(t_0))^2}{2 - p_3(t_0)} \right)$$

Direct computation shows that, under the above hypotheses, at every time t between t_0 and $t_0 + \Delta_2 t$ it holds

$$P(t) \leq (1 + \delta) P_0$$

where $\delta = \frac{2k \ln n}{n}$. By Theorem 2.6 we know that, given starting distribution $\pi(t_0)$, it holds for any Δt that

$$\| \pi(t_0) P_0^{\Delta t} - \overline{\pi} \| \leq (1 - \varepsilon^2)^{\Delta t}$$

which gives for all t between $t_0 + \Delta_1 t$ and $t_0 + \Delta_2 t$ that

$$\| \pi(t_0) P_0^{t - t_0} - \overline{\pi} \| \leq \frac{1}{n^{\varepsilon^2}}$$

Since $\pi(t) \leq (1 + \delta)^{t - t_0} \pi(t_0) P_0^{t - t_0}$, for $t \leq t_0 + \Delta_2 t$, we have that

$$\pi(t) \leq \left(1 + \frac{4k^2 \lceil \ln n \rceil^2}{n} + O\left(\frac{\ln^3 n}{n^2} \right) \right) \pi(t_0) P_0^{t - t_0}$$

Thus, we were able to bound $\pi^{(N_1 = 0)}(t)$ for all times t between $t_0 + \Delta_1 t$ and $t_0 + \Delta_2 t$.

$$\begin{aligned} \pi^{(N_1 = 0)}(t) &\leq \left(1 + \frac{k^2 \lceil \ln n \rceil^2}{n} + O\left(\frac{\ln^3 n}{n^2} \right) \right) \cdot \frac{p_3(t_0)}{2 - p_3(t_0)} + \frac{1}{n^{\varepsilon^2}} \\ &\leq (1 + \varepsilon) \cdot \frac{p_3(t_0)}{2 - p_3(t_0)} \end{aligned}$$

for sufficiently large n. □

Lemma 4.9. *Let $\varepsilon > 0$ and $k \in \mathbb{N}$ be fixed. With probability $1 - O(1/n^2)$ the following holds.*

For any time t_0 such that

- $p_2(t_0), p_3(t_0) > 2\varepsilon$,
- $N(t_0) > 2\varepsilon n$, *and*
- *a vertex of degree one with neighbor of degree one is never picked by the algorithm in time interval $[t_0, t_0 + (k + 1) \ln n]$.*

it holds that

$$\mathbf{Pr}[N_1(t) = 0] \geq (1 - \varepsilon) \cdot \frac{p_3(t_0)}{2 - p_3(t_0)}$$

for any $t \in [t_0 + \ln n, t_0 + (k+1)\ln n]$.

Proof. Analogously to proof of Lemma 4.8. □

We denote t_i the first time when $M \leq \frac{3}{2}n\left(1 - \frac{i}{h}\right)$.

Lemma 4.10. *Let* $\varepsilon > 0$ *be arbitrary. Let* t_i *be such that* $p_2(t_i) > 2\varepsilon, p_3(t_{i+1}) > 2\varepsilon$ *and* $N(t_i) > 2\varepsilon n$. *Then*

$$\mathbf{E}[t_{i+1} - t_i] \leq (1 + \varepsilon) \cdot \frac{3n\left(2 - \sqrt{1 - \frac{i}{h}}\right)}{4h}$$

Proof. We set $k \in \mathbb{N}$ large and $\varepsilon' > 0$ small.

Leaving out finitely many events of probability $O(1/n^2)$ will not effect the linear term of the expectation, hence we may omit them. Again we suppose that G is connected and distribution of N_1 behaves as described in Lemma 4.8. Suppose, by Lemma 4.5, that $\left|p_3(t) - \sqrt{2M(t)/(3n)}\right| < \varepsilon'\sqrt{2M(t)/(3n)}$ for all t where the conditions of Lemma hold.

At each time $t > 0$ the number of edges in the graph decreases by one (if $N_1(t) > 0$) or two (if $N_1(t) > 0$). We split interval $[t_i, t_{i+1}]$ into subintervals $[\tau_1, \tau_2], [\tau_2, \tau_3], \ldots, [\tau_{l-1}, \tau_l]$, each of them, except possibly the last one, of length $(k+1)\lceil \ln n \rceil$. Number of subintervals $l - 1$ is $\Theta(n/\ln n)$. By Lemma 4.7 we can suppose that the number of intervals $[\tau_s, \tau_{s+1}]$ in which there is picked a vertex of degree one with a neighbor of degree one is at most $\varepsilon' l$. We apply Lemma 4.9 for the remaining intervals. This gives us a lower bound on the expected number of times $R_{i,q}$ when, within interval $[\tau_q, \tau_{q+1}]$, state $N_1 = 0$ was visited.

$$\begin{aligned}
\mathbf{E}[R_{i,q}] &\geq k\lceil \ln n \rceil (1 - \varepsilon') \cdot \frac{p_3(\tau_q)}{2 - p_3(\tau_q)} \\
&\geq k\lceil \ln n \rceil (1 - 5\varepsilon') \cdot \frac{p_3(t_i)}{2 - p_3(t_i)}
\end{aligned}$$

for interval $[\tau_q, \tau_{q+1}]$ in which a vertex of degree one with a neighbor of degree one was not picked.

Summing these bounds we get a bound for the expected number of times R_i when, within interval $[t_i, t_{i+1}]$, state $N_1 = 0$ was visited. In the computation we omit the intervals in which a vertex of degree one with a neighbor of degree one was not picked and also we omit the starting stage of length $\lceil \ln n \rceil$ in each interval.

$$\begin{aligned}
\mathbf{E}[R_i] &\geq (1 - \varepsilon')\frac{k(t_{i+1} - t_i)}{k+1}(1 - 5\varepsilon') \cdot \frac{p_3(t_i)}{2 - p_3(t_i)} \\
&\geq \frac{k(t_{i+1} - t_i)}{k+1}(1 - 8\varepsilon') \cdot \frac{\sqrt{1 - \frac{i}{h}}}{2 - \sqrt{1 - \frac{i}{h}}}
\end{aligned}$$

This gives us (for sufficiently large k, n and sufficiently small ε')

$$\mathbf{E}[R_i] \geq (1-\varepsilon)\ (t_{i+1}-t_i) \cdot \frac{\sqrt{1-\frac{i}{h}}}{2-\sqrt{1-\frac{i}{h}}} \tag{4.10}$$

It is an easy observation that $(1+o(1))(t_{i+1}-t_i)+R_i = \frac{3n}{2h}$. Plugging (4.10) into this, we get

$$\mathbf{E}[t_{i+1}-t_i] \leq (1+\varepsilon) \cdot \frac{3n\left(2-\sqrt{1-\frac{i}{h}}\right)}{4h}$$

□

Lemma 4.11. *Let $\varepsilon > 0$ be arbitrary. Let t_i be such that $p_2(t_i) > 2\varepsilon, p_3(t_{i+1}) > 2\varepsilon$ and $N(t_i) > 2\varepsilon n$. Then*

$$\mathbf{E}[t_{i+1}-t_i] \geq (1-\varepsilon) \cdot \frac{3n\left(2-\sqrt{1-\frac{i}{h}}\right)}{4h}$$

Proof. Analogously to proof of Lemma 4.10. □

Lemma 4.12. *Let $\varepsilon > 0$ be arbitrary. Let t_i be such that $p_2(t_i) > 2\varepsilon, p_3(t_{i+1}) > 2\varepsilon$ and $N(t_i) > 2\varepsilon n$. Let Y_i be the number of times when a vertex of degree 1 is picked by the algorithm within interval $[t_i, t_{i+1}]$. Then*

$$(1-\varepsilon)\frac{3n}{2h}\left(1-\sqrt{1-\frac{i}{h}}\right) \leq \mathbf{E}[Y_i] \leq (1+\varepsilon)\frac{3n}{2h}\left(1-\sqrt{1-\frac{i}{h}}\right)$$

Proof. We obtained estimates for $t_{i+1}-t_1$ and the number of times R_1 when $N_1 = 0$ in Lemmas 4.10 and 4.11. The statement follows from an obvious equality $Y_i = t_{i+1} - t_1 - R_i$. □

We say that vertex v of graph $G(t)$ is *1-fixed* if it has exactly one neigbor, which is already removed; and this neighbor was placed into part B_1. If the neighbor was placed into B_2 instead, we say that the vertex is *2-fixed*. If a vertex v has exactly one neighbor, which is already removed from $G(t)$ and this neighbor was not added to the bipartite graph, we say that v is *0-fixed*. Similarly, we say that vertex v of graph $G(t)$ whose exactly two neighbors were removed from the graph is *0,0-fixed*, *0,1-fixed*, *0,2-fixed*, *1,1-fixed*, *1,2-fixed* or *2,2-fixed*.

Variables $K_1(t), K_2(t)$ and $K_0(t)$ are the numbers of 1-fixed, 2-fixed and 0-fixed vertices in $G(t)$.

Another notion which will come in handy is the notion of a *1-chain*. We say that times $t'_1 < t'_2 < \ldots < t'_l$ form a 1-chain if

- The degree of vertex $v(t'_i)$ picked by the algorithm at time t'_i $(i > 1)$ changed at time t'_{i-1} from 2 to 1. In particular, this means that $v(t'_{i-1})$ and $v(t'_i)$ were adjacent.
- There is no $t'_0 < t'_1$ such that $t'_0, t'_1, t'_2, \ldots, t'_l$ satisfy the above condition of a 1-chain.

- There is no $t'_{l+1} > t'_l$ such that $t'_1, t'_2, \ldots, t'_l, t'_{l+1}$ satisfy the above conditions of a 1-chain.

Vertices $\{v(t'_i)\}_{i=1}^{l}$ are the *vertices of 1-chain*. *Length* of the 1-chain is l. Time t'_1 is the *starting time* and $v(t'_1)$ is the *starter*.

We say that vertex $v(t)$ picked by the algorithm at time t is in state (i) A_0, (ii) A_1 or (iii) A_2 if $\deg_{G(t)} v(t) = 1$ and (i) $v(t)$ will not be added to the bipartite graph, (ii) will be added to part B_1 or (iii) will be added to part B_2, respectively. Observe that $v(t)$ is a vertex of a 1-chain if and only if $\deg_{G(t)} v(t) = 1$.

Proposition 4.13. *Let $\varepsilon \in (0,1)$ be arbitrary. Let t_s be the first time when $p_3(t_s) \leq 2\varepsilon$. Let $C_\varepsilon(n) = \lceil \frac{3}{\varepsilon} \ln n \rceil$. Then the probability that there exists $t_0 \leq t_s$ such, that for all t, $t \in [t_0, t_0 + C_\varepsilon(n)]$ it holds* $\operatorname{mindeg}(G(t)) = 1$ *is $O(1/n^2)$.*

Proof. Let t_0 be time as in the hypothesis. We condition the computations on p_3 behaving as described in Lemma 4.2.

Since one cannot get from state $N_1 = 1$ to state $N_1 = 2$ avoiding state $N_1 = 0$, there must be an integer k such that $N_1(t) = 2$ for all $t \in [t_0, t_0 + k]$ and $N_1(t) = 1$ for all $t \in [t_0 + k + 1, t_0 + C_\varepsilon(n)]$.

Recall that the transition probabilities of states $N_1 = 0, N_1 = 1, N_1 = 2$ were examined in the proof of Lemma 4.8.

$$
\begin{aligned}
&\mathbf{Pr}\left[\exists t_0 : t_0 \leq t_{\mathrm{s}} \text{ and } \operatorname{mindeg}(G(t)) = 1 \text{ for all } t \in [t_0, t_0 + C_\varepsilon(n)]\right] \\
&\leq \sum_{t \leq t_{\mathrm{s}}} \sum_{k=0}^{C_\varepsilon(n)} \prod_{t=t_0}^{t_0+k} \left(p_2(t) + O\left(\frac{1}{M(t)}\right)\right) \prod_{t=t_0+k+1}^{C_\varepsilon(n)} \left(p_2(t) + O\left(\frac{1}{M(t)}\right)\right) \\
&\leq nC_\varepsilon(n) \left(1 - \frac{3}{2}\varepsilon\right)^{C_\varepsilon(n)} \\
&= O\left(\frac{1}{n^2}\right)
\end{aligned}
$$

□

Proposition 4.14. *Let $\{Q^*_t\}_{t \in \mathbb{N}_0}$ be a Markov chain with states S_1, S_2 and S_3, transition matrix*

$$
P = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & 0 \\ 0 & \alpha_{3,2} & \alpha_{3,3} \end{pmatrix}
$$

and initial state $Q^*_0 = S_1$. Let $\{Q_t\}_{t \in \mathbb{N}_0}$ be a discrete time stochastic process with states S_1, S_2 and S_3, initial state $Q_0 = S_1$ such, that at every time t the transition probabilities

satisfy the following conditions

$$\begin{aligned}
\mathbf{Pr}[Q_{t+1} = S_1 | Q_t = S_1] &\geq \alpha_{1,1} \\
\mathbf{Pr}[Q_{t+1} = S_2 | Q_t = S_1] &\leq \alpha_{1,2} \\
\mathbf{Pr}[Q_{t+1} = S_3 | Q_t = S_1] &\leq \alpha_{1,3} \\
\mathbf{Pr}[Q_{t+1} = S_1 | Q_t = S_2] &\geq \alpha_{2,1} \\
\mathbf{Pr}[Q_{t+1} = S_2 | Q_t = S_2] &\leq \alpha_{2,2} \\
\mathbf{Pr}[Q_{t+1} = S_3 | Q_t = S_2] &= 0 \\
\mathbf{Pr}[Q_{t+1} = S_1 | Q_t = S_3] &= 0 \\
\mathbf{Pr}[Q_{t+1} = S_2 | Q_t = S_3] &\geq \alpha_{3,2} \\
\mathbf{Pr}[Q_{t+1} = S_3 | Q_t = S_3] &\leq \alpha_{3,3}
\end{aligned}$$

Let $R^*(t)$ be a random variable counting how many times $Q^*_\tau = S_1$ for $\tau \in [0,t]$. Let $R(t)$ be a random variable counting how many times $Q_\tau = S_1$ for $\tau \in [0,t]$.

Then

$$R^*(t) \leq R(t) \qquad \textit{in distribution}$$

for every $t \in \mathbb{N}_0$.

Proof. Let random variables T^* and T indicate the first time $t > 0$ when $Q^*_t = S_1$ and $Q_t = S_1$, respectively. The statement will be proven by showing that

$$T^* \geq T \qquad \text{in distribution}$$

This can be easily done by coupling. □

Lemma 4.15. *For any $\varepsilon \in \left(0, \frac{1}{100}\right)$,*

$$\mathbf{Pr}\left[\exists t : p_2(t') < 1-\varepsilon \textit{ for all } t' \in [0,t] \textit{ and } |K_1(t) - K_2(t)| > \varepsilon n\right] = O\left(\frac{1}{n^2}\right) \quad (4.11)$$

Proof. We shall assume, as usually, that rare events do not occur. We assume that a vertex of degree 2 is picked at least once in every $C_\varepsilon(n)$ steps (by Proposition 4.13); we assume that G is connected and that p_3 behaves as described by Lemma 4.2.

In the proof we use some inequalities which become true only when n is large enough, $n \geq n_0(\varepsilon)$.

Define $h = \left\lfloor \frac{n}{\ln n} \right\rfloor$ and for $i = 0, 1, \ldots, \left\lfloor \frac{3}{2} \ln n \right\rfloor$ define $m_i = \frac{3}{2}n - ih$. Let $\mathcal{E}_i$ be an event that $|K_1(\widehat{\tau}_i) - K_2(\widehat{\tau}_i)| \leq \varepsilon n$ where $\widehat{\tau}_i$ is the first time when $M(\widehat{\tau}_i) \leq m_i$. Set $i_{\text{last}} = \min\left\{ \left\lfloor \frac{3}{2} \ln n \right\rfloor, i \ : p_2(t') \geq 1-\varepsilon \text{ for some } t' \in [0, \widehat{\tau}_i] \right\}$. We shall show that

$$\mathbf{Pr}\left[\overline{\mathcal{E}_i} \mid \mathcal{E}_{i-1}\right] = O\left(\frac{1}{n^4}\right) \quad (4.12)$$

We postpone the proof of (4.12) and conclude from it the statement of the theorem.

Event $\mathcal{E}_0$ holds with probability 1. Combining this with (4.12) we have that

$$\mathbf{Pr}\left[\bigcap_{i=0}^{i_{\text{last}}} \mathcal{E}_i\right] = 1 - O\left(\frac{1}{n^2}\right)$$

This means that with probability $1-O(1/n^2)$ it holds for every time $\widehat{\tau_i}$ that $|K_1(\widehat{\tau_i}) - K_2(\widehat{\tau_i})| \leq \varepsilon n$. But since $\widehat{\tau_i} - \widehat{\tau_{i-1}} \leq \frac{n}{\ln n}$ and number $|K_1 - K_2|$ changes by at most 3 in each step, we have that $|K_1(t) - K_2(t)| \leq 2\varepsilon n$ at any time t such that $M(t) > \frac{3}{2}n - o(n)$ and $p_2 < 1 - \varepsilon$ on $[0, t]$. By similar consideration we have that $|K_1(t) - K_2(t)| \leq 6\varepsilon n$ at any time t which proves the statement of the theorem (up to factor 6; which is, of course, not fatal).

We shall now prove (4.12). Suppose that $\mathcal{E}_{i-1}$ holds and $\mathcal{E}_i$ does not. Because $\mathcal{E}_i$ does not hold, we have that

$$|K_1(t) - K_2(t)| > \varepsilon n - \frac{\varepsilon^2 n}{100}$$

for all $t \in [\widehat{\tau_{i-1}}, \widehat{\tau_i}]$. Without loss of generality we assume that

$$K_1(t) - K_2(t) > \varepsilon n - \frac{\varepsilon^2 n}{100}$$

for all $t \in [\widehat{\tau_{i-1}}, \widehat{\tau_i}]$. By the same reasoning we know that at every time t

$$|K_1(t) - K_2(t)| < \varepsilon n + \frac{\varepsilon^2 n}{100}$$

We divide the rest of the proof into several steps:

(⋆**1**) We count how many times state $N_1 = 0$ was visited.

(⋆**2**) We estimate the probability that a 1-chain with starter in state A_1 or A_2 emerges at certain time.

(⋆**3**) We count the 1-chains having starter in state A_i and the last visited state A_j.

(⋆**4**) We show that outside the 1-chains the difference $K_1 - K_2$ tends to decrease.

(⋆**5**) The final computation gives (4.12).

Parts (⋆**1**), (⋆**2**), (⋆**3**) and (⋆**4**) are somewhat technical. Each shows that some random variable behaves nearly the same way as if many events (which are actually slightly dependant) in the run of the algorithm are considered mutually independent. The reader may find it more understandable to read part (⋆**5**) first (and look up definitions of some symbols in the previous parts).

(⋆**1**):

Set $\Delta\tau = \widehat{\tau_i} - \widehat{\tau_{i-1}}$. Clearly, $\Delta\tau = \Theta\left(\frac{n}{\ln n}\right)$. Define τ^* as the first time after $\widehat{\tau_{i-1}}$ such that $N_1(\tau^*) = 0$. By assumption made on the beginning of the proof, $\tau^* \leq \widehat{\tau_{i-1}} + C_\varepsilon(n)$

where $C_\varepsilon(n) = \lceil \frac{3}{\varepsilon} \ln n \rceil = o(\Delta\tau)$. Let R be a random variable, which counts how many times the state $N_1 = 0$ was visited within interval $[\tau^*, \widehat{\tau_i} - 1]$. Let R^* be a random variable, which counts how many times the state S_1 was visited within interval $[\tau^*, \widehat{\tau_i} - 1]$ for a Markov chain $\{Q_t^*\}_{t\in\mathbb{N}_0}$ with states S_1, S_2 and S_3 and with corresponding transition matrix P_0,

$$P_0 = \begin{pmatrix} p_3(\widehat{\tau_{i-1}})^2 - \frac{\varepsilon^2}{4} & 2p_2(\widehat{\tau_{i-1}})p_3(\widehat{\tau_{i-1}}) + \frac{\varepsilon^2}{8} & p_2(\widehat{\tau_{i-1}})^2 + \frac{\varepsilon^2}{8} \\ p_3(\widehat{\tau_{i-1}}) - \frac{\varepsilon^2}{8} & p_2(\widehat{\tau_{i-1}}) + \frac{\varepsilon^2}{8} & 0 \\ 0 & p_3(\widehat{\tau_{i-1}}) - \frac{\varepsilon^2}{8} & p_2(\widehat{\tau_{i-1}}) + \frac{\varepsilon^2}{8} \end{pmatrix}$$

As n goes to infinity, the differences

$$\max_{t\in[\tau_{i-1},\tau_i]}\{p_2(t)\} - \min_{t\in[\tau_{i-1},\tau_i]}\{p_2(t)\}$$

and

$$\max_{t\in[\tau_{i-1},\tau_i]}\{p_3(t)\} - \min_{t\in[\tau_{i-1},\tau_i]}\{p_3(t)\}$$

tend uniformly to zero. This means that we may use chain $\{Q_t^*\}_{t\in\mathbb{N}_0}$ to get a lower bound on R. Proposition 4.14 gives us

$$R^* \leq R \qquad \text{in distribution}$$

Lower bound on R^* is a straightforward application of Lemma 2.8; process runs for time at least $\left(1 - \frac{\varepsilon^2}{4}\right)\Delta\tau$ and constants of the Lemma are $p = 0$, $q = p_3(\widehat{\tau_{i-1}}) - \frac{\varepsilon^2}{8}$ and random variable Z is defined by

$$Z = \begin{cases} 0 & \text{with probability } p_3(\widehat{\tau_{i-1}})^2 - \frac{\varepsilon^2}{4} \\ 1 & \text{with probability } 2p_2(\widehat{\tau_{i-1}})p_3(\widehat{\tau_{i-1}}) + \frac{\varepsilon^2}{8} \\ 2 & \text{with probability } p_2(\widehat{\tau_{i-1}})^2 + \frac{\varepsilon^2}{8} \end{cases}$$

Thus we have $\mathbf{E}[Z] = 2p_2(\widehat{\tau_{i-1}}) + \frac{3\varepsilon^2}{8} + O(1/N(\widehat{\tau_{i-1}}))$. Direct substitution gives

$$\mathbf{Pr}\left[R^* < \frac{\left(1 - \frac{\varepsilon^2}{4}\right)\Delta\tau\left(p_3(\widehat{\tau_{i-1}}) - \frac{\varepsilon^2}{8}\right)}{2p_2(\widehat{\tau_{i-1}}) + p_3(\widehat{\tau_{i-1}}) + \frac{\varepsilon^2}{4} + O\left(\frac{1}{N(\widehat{\tau_{i-1}})}\right)} - 4\ln n\sqrt{\left(1 - \frac{\varepsilon^2}{4}\right)\Delta\tau}\right] = O\left(\frac{1}{n^4}\right)$$

which yields

$$\mathbf{Pr}\left[R^* < \frac{\left(1 - \frac{\varepsilon}{4}\right)\ \Delta\tau\ p_3(\widehat{\tau_{i-1}})}{2 - p_3(\widehat{\tau_{i-1}})}\right] = O\left(\frac{1}{n^4}\right) \tag{4.13}$$

($\star\mathbf{2}$):

In the next denotion of 1-chains $\{\mathcal{C}_\kappa^0\}_\kappa$, $\{\mathcal{C}_\kappa^1\}_\kappa$ and $\{\mathcal{C}_\kappa^2\}_\kappa$ we restrict ourselves on the 1-chains which start and finish within interval $[\tau^*, \widehat{\tau_i}]$. Let $\mathcal{C}_1^0, \mathcal{C}_2^0, \ldots, \mathcal{C}_{c_0}^0$, $\mathcal{C}_1^1, \mathcal{C}_2^1, \ldots, \mathcal{C}_{c_1}^1$ and $\mathcal{C}_1^2, \mathcal{C}_2^2, \ldots, \mathcal{C}_{c_2}^2$ be all the 1-chains with starter in state A_0, A_1 and A_2, respectively.

A chain starting at time t with starter in state A_1 could have emerged only if $N_1 = 0$ in the previous step $t-1$. In this case the expected number of starters in state A_1 (there can be 0, 1 or 2 of them) emerging at time t is

$$\frac{(K_0(t)+2K_1(t))(K_0(t)+K_2(t))}{(K_0(t)+K_1(t)+K_2(t))^2}\cdot\left(1+O\left(\frac{1}{N(t)}\right)\right)p_2(t)$$

We shall use bounds on this probability which are true on the whole interval $[\widehat{\tau_{i-1}},\widehat{\tau_i}]$. The bound from above can be set to

$$\frac{(K_0(\widehat{\tau_{i-1}})+2K_1(\widehat{\tau_{i-1}}))(K_0(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))}{(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))^2}\cdot\left(1+\frac{\varepsilon^2}{100}\right)p_2(\widehat{\tau_{i-1}})$$

and the bound from below to

$$\frac{(K_0(\widehat{\tau_{i-1}})+2K_1(\widehat{\tau_{i-1}}))(K_0(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))}{(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))^2}\cdot\left(1-\frac{\varepsilon^2}{100}\right)p_2(\widehat{\tau_{i-1}})$$

Similarly, the expected number of starters in state A_2 emerging at time t is

$$\frac{(K_0(t)+2K_2(t))(K_0(t)+K_1(t))}{(K_0(t)+K_1(t)+K_2(t))^2}\cdot\left(1+O\left(\frac{1}{N(t)}\right)\right)p_2(t)$$

with bounds

$$\frac{(K_0(\widehat{\tau_{i-1}})+2K_2(\widehat{\tau_{i-1}}))(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}}))}{(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))^2}\cdot\left(1+\frac{\varepsilon^2}{100}\right)p_2(\widehat{\tau_{i-1}})$$

from above and

$$\frac{(K_0(\widehat{\tau_{i-1}})+2K_2(\widehat{\tau_{i-1}}))(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}}))}{(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))^2}\cdot\left(1-\frac{\varepsilon^2}{100}\right)p_2(\widehat{\tau_{i-1}})$$

from below.

($\star\mathbf{3}$):

Let $\{t'_j\}_{j=1}^{l}$ form a 1-chain $\mathcal{C}$, $\mathcal{C}\in\{\mathcal{C}^0_\kappa\}_\kappa\cup\{\mathcal{C}^1_\kappa\}_\kappa\cup\{\mathcal{C}^2_\kappa\}_\kappa$. We glance at the approximation of the processes with states A_0, A_1 and A_2 to see what is happening. We denote the matrix of transition probabilities of the 1-chain at time t by $C(t)$ (The notion of transition matrix does not apply for this case—it is not clear what $C(t)$ should stand for. Entry $C(t)_{\iota,j}$ is defined as the probability that the state of the 1-chain in the step after time t will be the one corresponding to column j, conditioned on what was happening in the graph until time $t-1$ and on that the vertex chosen at time t is in state corresponding to row ι. Note, that matrices $\{C(t)\}_t$ are not independent.), C^{appr} is the approximation by a Markov chain $\{L_t\}_{t\in\mathbb{N}}$,

$$C^{\text{appr}}=\begin{pmatrix} 0 & p_2\cdot\frac{K_0+2K_2}{2(K_0+K_1+K_2)} & p_2\cdot\frac{K_0+2K_1}{2(K_0+K_1+K_2)} & p_3\\ p_2\cdot\frac{K_2}{K_0+K_1+K_2} & 0 & p_2\cdot\frac{K_0+K_1}{K_0+K_1+K_2} & p_3\\ p_2\cdot\frac{K_1}{K_0+K_1+K_2} & p_2\cdot\frac{K_0+K_2}{K_0+K_1+K_2} & 0 & p_3\\ 0 & 0 & 0 & 1\end{pmatrix}$$

where the last row and column of the matrix represent a termination of the chain and symbols p_2, p_3, K_0, K_1, K_2 are shortcuts for $p_2(t'_1)$, $p_3(t'_1)$, $K_0(t'_1)$, $K_1(t'_1)$, $K_2(t'_1)$. The described approximative chain is an absorbing Markov chain; its fundamental matrix is

$$N = \frac{1}{\Upsilon} \cdot \cdot \begin{pmatrix} 1-p_2^2(K_0+K_1)(K_0+K_2)/\Gamma^2 & p_2(\frac{1}{2}K_0+K_2)/\Gamma+p_2^2(\frac{1}{2}K_0+K_1)(K_0+K_2)/\Gamma^2 & p_2(\frac{1}{2}K_0+K_1)/\Gamma+p_2^2(\frac{1}{2}K_0+K_2)(K_0+K_1)/\Gamma^2 \\ p_2K_2/\Gamma+p_2^2K_1(K_0+K1)/\Gamma^2 & 1-p_2^2K_1(\frac{1}{2}K_0+K_1)/\Gamma^2 & p_2(K_0+K_1)/\Gamma+p_2^2K_2(\frac{1}{2}K_0+K_1)/\Gamma^2 \\ p_2K_1/\Gamma+p_2^2K_2(K_0+K_2)/\Gamma^2 & p_2(K_0+K_2)/\Gamma+p_2^2K_1(\frac{1}{2}K_0+K_2)/\Gamma^2 & 1-p_2^2K_2(\frac{1}{2}K_0+K_2)/\Gamma^2 \end{pmatrix}$$

where $\Gamma = K_0+K_1+K_2$ and $\Upsilon = 1-p_2^2((K_0+K_1)(K_0+K_2)+K_1(\frac{1}{2}K_0+K_1)+K_2(\frac{1}{2}K_0+K_2))/\Gamma^2-p_2^3(K_1(\frac{1}{2}K_0+K_2)(K_0+K_1)+K_2(\frac{1}{2}K_0+K_1)(K_0+K_2))/\Gamma^3$. We shall investigate probabilities $\alpha_{\iota,1}$ and $\alpha_{\iota,2}$,

$$\alpha_{\iota,1} = \mathbf{Pr}[\text{the last transient state of } \mathcal{C} \text{ was } A_1 \mid \text{the starter of } \mathcal{C} \text{ is in state } A_\iota]$$
$$\alpha_{\iota,2} = \mathbf{Pr}[\text{the last transient state of } \mathcal{C} \text{ was } A_2 \mid \text{the starter of } \mathcal{C} \text{ is in state } A_\iota]$$

for $\iota = 0, 1, 2$. First we look at the approximation of the probabilities. Let

$$\begin{aligned} \alpha_{\iota,1}^{\text{appr}} &= \mathbf{Pr}[\text{the last visited transient state of } \{L_t\}_t \text{ was } A_1 \mid L_1 = A_\iota] \\ \alpha_{\iota,2}^{\text{appr}} &= \mathbf{Pr}[\text{the last visited transient state of } \{L_t\}_t \text{ was } A_2 \mid L_1 = A_\iota] \end{aligned}$$

By Proposition 2.10 we have $\alpha_{\iota,1}^{\text{appr}} = p_3 N_{\iota,2}$ and $\alpha_{\iota,2}^{\text{appr}} = p_3 N_{\iota,3}$. It can be easily seen that $(1-8\ln n/(\varepsilon n))C_0 \le C(t) \le (1+8\ln n/(\varepsilon n))C_0$ for every time t of the chain $\mathcal{C}$. It holds

$$\alpha_{\iota,j} \le \left(1+\frac{8\ln n}{\varepsilon n}\right)^{\lceil \frac{3}{\varepsilon}\ln n \rceil} \alpha_{\iota,j}^{\text{appr}} + O\left(\frac{1}{n^2}\right) = (1+o(1))\alpha_{\iota,j}^{\text{appr}} + O\left(\frac{1}{n^2}\right) \qquad (4.14)$$

and

$$\alpha_{\iota,j} \ge \left(1-\frac{8\ln n}{\varepsilon n}\right)^{\lceil \frac{3}{\varepsilon}\ln n \rceil} \alpha_{\iota,j}^{\text{appr}} - O\left(\frac{1}{n^2}\right) = (1+o(1))\alpha_{\iota,j}^{\text{appr}} - O\left(\frac{1}{n^2}\right) \qquad (4.15)$$

(Terms $O\left(\frac{1}{n^2}\right)$ are here, because we conditioned on the lenght of each 1-chain; we allow it to be of length $\lceil \frac{3}{\varepsilon}\ln n \rceil$ at most). Denote symbols $o(1)$ used in equations (4.14) and (4.15) by $\frac{1}{2}g_1(n)$ and $\frac{1}{2}g_2(n)$, respectively (g_1 is a positive function and g_2 is negative).

Define C^0 as

$$C^0 = \begin{pmatrix} 0 & \frac{p_2(\widehat{\tau_{i-1}})\cdot(K_0(\widehat{\tau_{i-1}})+2K_2(\widehat{\tau_{i-1}}))}{2(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))} & \frac{p_2(\widehat{\tau_{i-1}})\cdot(K_0(\widehat{\tau_{i-1}})+2K_1(\widehat{\tau_{i-1}}))}{2(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))} & p_3(\widehat{\tau_{i-1}}) \\ \frac{p_2(\widehat{\tau_{i-1}})\cdot K_2(\widehat{\tau_{i-1}})}{K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}})} & 0 & \frac{p_2(\widehat{\tau_{i-1}})\cdot(K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}}))}{K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}})} & p_3(\widehat{\tau_{i-1}}) \\ \frac{p_2(\widehat{\tau_{i-1}})\cdot K_1(\widehat{\tau_{i-1}})}{K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}})} & \frac{p_2(\widehat{\tau_{i-1}})\cdot(K_0(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}}))}{K_0(\widehat{\tau_{i-1}})+K_1(\widehat{\tau_{i-1}})+K_2(\widehat{\tau_{i-1}})} & 0 & p_3(\widehat{\tau_{i-1}}) \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and let N^0 be a fundamental matrix which would correspond to a Markov chain with transition matrix C^0. It can be checked that for all chains $\{L_t\}_{t\in\mathbb{N}}$ (as above) with transition matrices C^{appr} it holds for $\iota = 1, 2, 3$ that

$$\begin{aligned} N_{1,\iota} &\le (1+f_1(n))N_{1,\iota}^0 \\ N_{2,\iota} &\ge (1-f_2(n))N_{2,\iota}^0 \end{aligned}$$

where $f_1(n) = o(1)$, $f_2(n) = o(1)$. Let $r_{\iota,j}$ ($\iota = 0, 1, 2, j = 1, 2$) be the number of 1-chains with starter in state A_i and last visited transient state A_j. Let $r_{\iota,1}^{\text{upper}}$ be a sum of c_1 independent Bernoulli variables with success probability

$$(1 + g_1(n))(1 + f_{\text{p}}(n))(1 + f_1(n))p_3(\widehat{\tau_{i-1}})N_{1,\iota}^0$$

where $f_{\text{p}}(n) = o(1)$. Let $r_{\iota,2}^{\text{lower}}$ be a sum of c_2 independent Bernoulli variables with success probability

$$(1 + g_2(n))(1 - f_{\text{p}}(n))(1 - f_2(n))p_3(\widehat{\tau_{i-1}})N_{1,\iota}^0$$

We have that

$$\begin{aligned} r_{\iota,1} &\leq r_{\iota,1}^{\text{upper}} \qquad \text{in distribution} \\ r_{\iota,2} &\geq r_{\iota,2}^{\text{lower}} \qquad \text{in distribution} \end{aligned}$$

We apply Chernoff bounds 2.2 for $r_{\iota,1}^{\text{upper}}$ and $r_{\iota,2}^{\text{lower}}$

$$\mathbf{Pr}[r_{\iota,1}^{\text{upper}} > (1 + \varepsilon^2)c_\iota p_3(\widehat{\tau_{i-1}})N_{1,\iota}^0] = O\left(\frac{1}{n^4}\right) \tag{4.16}$$

$$\mathbf{Pr}[r_{\iota,2}^{\text{lower}} < (1 - \varepsilon^2)c_\iota p_3(\widehat{\tau_{i-1}})N_{2,\iota}^0] = O\left(\frac{1}{n^4}\right) \tag{4.17}$$

($\star$**4**):

Let L_1 and L_2 be the change of K_1 and K_2, respectively, when restricted to vertices which are not in 1-chain, i.e.

$$L_\iota = \sum_{\substack{t \in [\tau^*, \widehat{\tau_i}] \\ \deg_{G(t)} v(t) = 2}} \Delta K_\iota(t)$$

At each such time t the probabilities that vertex $v(t)$ (yet unchosen at the beginning of step t) will be 0-fixed, 1-fixed or 2-fixed are in ratio $K_1(t) : K_2(t) : K_3(t)$; and the expected number of its neighbors of degree three will be $2p_3(t) + O(1/N(t))$. If vertex $v(t)$ is 1-fixed, for illustration, having two neighbors of degree 3 in $G(t)$, than it is assigned to B_2 and $\Delta K_1(t) = -1$ ($v(t)$ is not 1-fixed any more) and $\Delta K_2 = 2$ (two 2-fixed vertices emerged). In general, the expected changes of K_1 and K_2 at time t are

$$\begin{aligned} \mathbf{E}[\Delta K_1(t)] &= -\frac{K_1(t)}{K_0(t) + K_1(t) + K_2(t)} + 2p_3(t) \cdot \frac{\frac{1}{2}K_0(t) + K_2(t)}{K_0(t) + K_1(t) + K_2(t)} + O\left(\frac{1}{N(t)}\right) \\ \mathbf{E}[\Delta K_2(t)] &= -\frac{K_2(t)}{K_0(t) + K_1(t) + K_2(t)} + 2p_3(t) \cdot \frac{\frac{1}{2}K_0(t) + K_1(t)}{K_0(t) + K_1(t) + K_2(t)} + O\left(\frac{1}{N(t)}\right) \end{aligned}$$

with the variances very moderate. One can bound these variables in distribution from below and from above respectively by independent variables with the same distribution and use Chernoff inequality 2.2. Combining this with an estimate on the number of summands (4.13) and using $\frac{K_1 - K_2}{K_0 + K_1 + K_2} > \varepsilon n - \frac{\varepsilon^2 n}{100}$ we get

$$\mathbf{Pr}\left[L_1 - L_2 > -\frac{\varepsilon\, \Delta\tau\, p_3(\widehat{\tau_{i-1}})}{2(2 - p_3(\widehat{\tau_{i-1}}))}\right] = O\left(\frac{1}{n^4}\right) \tag{4.18}$$

($\star$**5**):

One can bound number $\overline{r} = r_{0,1}^{\text{upper}} + r_{1,1}^{\text{upper}} + r_{2,1}^{\text{upper}} - r_{0,2}^{\text{lower}} - r_{1,2}^{\text{lower}} - r_{2,2}^{\text{lower}}$ from above using Equations (4.16) and (4.17) by

$$\frac{\varepsilon^2 p_3(c_0 + c_1 + c_2)}{10\Upsilon}$$

with error probability $O(1/n^4)$. Assumption $p_2 < 1 - \varepsilon$ and a little high school algebra gives $\Upsilon > \varepsilon$. We have

$$\mathbf{Pr}\left[\overline{r} > \frac{1}{10}\varepsilon p_3(c_0 + c_1 + c_2)\right] = O\left(\frac{1}{n^4}\right) \tag{4.19}$$

Now the crucial observation comes: numbers K_1 and K_2 change by at most 1 during the run of a 1-chain and the possible change happens only the termination step of a 1-chain. In that step a vertex of degree 3 becomes either 0-fixed, 1-fixed or 2-fixed, depending on whether the last vertex of the 1-chain was in state A_0, A_1 or A_2. Equation (4.19) thus states that almost surely (with error probability $O(1/n^4)$) the number $K_1 - K_2$ increases by at most $\frac{1}{10}\varepsilon p_3(c_0 + c_1 + c_2)$ when restricted to 1-chains. By (4.18) $K_1 - K_2$ decreases by at least $\frac{\varepsilon\,\Delta\tau\,p_3(\widehat{\tau_{i-1}})}{2(2-p_3(\widehat{\tau_{i-1}}))}$ when restricted to the complement of 1-chains with probability $1 - O(1/n^4)$. This gives in total

$$K_1(t) - K_2(t)\Big|_{t=\tau^*}^{\widehat{\tau_i}} < -\frac{1}{10}\varepsilon\,\Delta\tau\,p_3(\widehat{\tau_{i-1}}) \tag{4.20}$$

with probability $1 - O(1/n^4)$ and (4.12) follows — we see that we did not have to treat interval $[\widehat{\tau_{i-1}}, \tau^*]$; it is of at most logarithmic length while Equation (4.20) gives us spare in $\Delta(K_1 - K_2)$ of size $\Theta\left(\frac{n}{\ln n}\right)$. □

Theorem 4.16. *For every $\varepsilon > 0$, BIPGREEDY returns bipartite graph with at least $\left(\frac{3}{4} - \varepsilon - o(1)\right) n$ vertices on average given graph $G \in_u \mathcal{G}_{3,n}^*$ on the input.*

Proof. Fix $\eta > 0$ and h a positive integer.

We do the analysis only in the central part of the run, i.e. except a very short initial and final part of the run. What is important is that we can make these parts arbitrarily small, i.e. as small as linear with arbitrarily small constant. Thus what is happening in these parts can be fitted within allowed error εn.

All the vertices of original graph G are included in the bipartite graph, except vertices which were in state A_0 and vertices of degree 0. By Lemma 4.7 there are at most $\ln n$ vertices of degree 0 with probability $1 - O(1/n^2)$, so this number has no impact on the linear term of the expected number of removed vertices. Vertices in state A_0 correspond exactly to 1,2-fixed vertices. Let us count how many of them there were. By Lemma 4.12, the expected number of vertices of degree one picked in the interval $[t_{i-1}, t_i]$ is bounded by

$$\mathbf{E}[Y_i] \leq (1+\eta)\frac{3n}{2h}\left(1 - \sqrt{1 - \frac{i}{h}}\right)$$

We see from Lemma 4.5 that for every $\eta' > 0$ there exists $\beta' > 0$ such that

$$\mathbf{Pr}\left[\exists t \ : \ \eta' n < N(t) < (1-\eta')n \text{ and } N_2(t) < \beta' n\right] = O\left(\frac{1}{n^2}\right)$$

Set β to such a number corresponding to $\eta' = \eta$. We use Lemma 4.15 with allowing deviation K_1 from K_2 to be at most $\eta\beta n$. Thus we may condition on $|K_1(t) - K_2(t)| \leq \eta\beta n$ for all t such that $\eta n < N(t) < (1-\eta)n$. Therefore at each step t, conditioning on that a vertex v of degree one will emerge in this step, the probability that it will be 1,2-fixed is at most $\frac{1}{2} + \eta$ (here we used Lemma 4.2).

The average number of 1,2-fixed vertices (not including 1,2-fixed vertices near the beginning and the end) can be bounded from above by

$$\begin{aligned}
&\left(\frac{1}{2}+\eta\right)\sum_{i=0}^{h}\mathbf{E}[Y_i] \\
&\leq \quad \left(\frac{1}{2}+\eta\right)(1+\eta)\sum_{i=0}^{h}\frac{3n}{2h}\left(1-\sqrt{1-\frac{i}{h}}\right) \\
&\overset{h\to\infty}{=} \quad \left(\frac{1}{2}+\eta\right)(1+\eta)\cdot n\cdot\left(\int_0^1 \frac{3\left(1-\sqrt{1-x}\right)}{2}\mathbf{d}x + o(1)\right) \\
&= \quad \left(\frac{1}{2}+\eta\right)(1+\eta)\cdot n\cdot\left(\frac{1}{2}+o(1)\right)
\end{aligned}$$

Setting η small and h large we see that BIPGREEDY returns a bipartite graph with at least $\left(\frac{3}{4}-\varepsilon\right)n$ vertices (for $n \geq n_0(\varepsilon)$) on average when $G \in \mathcal{G}^*_{3,n}$ is given on the input. □

Corollary 4.17. *For any $\varepsilon > 0$ it a.a.s. holds that for $G \in_u \mathcal{G}^*_{3,n}$*

$$\frac{\mathrm{bi}(G)}{n} > \frac{3}{4} - \varepsilon$$

Proof. Follows from Theorem 4.16 and Theorem 3.4. □

We present the following corollary for the independence number of a random cubic graph just for completeness. The result itself is much weaker then Frieze and Suen's result. There are two reasons for this: we show in Remark 4.19 that the constant in Theorem 4.16 could be considerably improved. Secondly, one should not expect that an algorithm designed for finding a big induced subgraph will excel in finding a big independent set also.

Corollary 4.18. *For any $\varepsilon > 0$ it a.a.s holds that for $G \in_u \mathcal{G}^*_{3,n}$*

$$\frac{\alpha(G)}{n} > \frac{3}{8} - \varepsilon$$

Proof. Follows from Corollary 4.17 by taking independent set I the bigger of parts B_1 and B_2. □

4.4 Concluding remarks

Remark 4.19. The analysis done in Theorem 4.16 does not give asymptotically tight result (i.e. there exists $q > 0.75$ such that $\frac{\mathrm{bi}_{\mathrm{alg}}(G)}{n} \geq q$ a.a.s.). To see this we recall how we bounded the number of vertices excluded from the graph. A vertex is excluded (with probability p) only if it was picked as a degree-one vertex. We were able to bound p by $\frac{1}{2} + \varepsilon$ in Lemma 4.15 by exhibiting that when a vertex becomes degree-one vertex, then it will be with nearly the same probability 1,2-fixed as 1,1-fixed or 2,2-fixed. But there is also a considerable (i.e. linear, during the middle phase) number of 0-fixed vertices; these may turn into 0,0-fixed, 0,1-fixed or 0,2-fixed vertices and will never be excluded. We did not take this fact into account in the analysis.

We believe that tight result is obtainable with only a little more effort (note that all other statements which we used in the proof of Theorem 4.16 had counterparts saying that the constant in that statement is tight). It is quite likely that $q \doteq 0.8719$, as data obtained from computer simulation show.

Remark 4.20. Algorithm BIPGREEDY can be improved. Here we propose a simple improvement BIPGREEDY2. BIPGREEDY2 cares about the neighborhood of vertex v which is currently being placed into the bipartite graph. If v could be placed into both parts of the bipartite graph, then we place it in that part, so as to create less restrictions to the neighboring vertices. Only if none of the choices is favorable we toss a coin. In BIPGREEDY we always tossed a coin at this point.

BIPGREEDY2;
Input G;
$B_1 := \emptyset$;
$B_2 := \emptyset$;
while $G \neq \emptyset$ **do**
  $V_{\min} :=$ set of vertices of minimum degree in G;
  choose $v \in V_{\min}$ uniformly at random;
  if v has no formal neighbor in B_1 nor B_2 **then**
    if placing v to B_1 creates less new restrictions to the neighbors of v then placing it to B_2 **then**
      $B_1 := B_1 \cup \{v\}$;
    end if
    if placing v to B_2 creates less new restrictions to the neighbors of v then placing it to B_1 **then**
      $B_2 := B_2 \cup \{v\}$;
    end if
    if the number of new restrictions to the neighbors of v does not depend on where v is placed **then**
      if random(2)$= 0$ **then**
        $B_1 := B_1 \cup \{v\}$;
      else

$B_2 := B_2 \cup \{v\}$;
end if
end if
end if
if v has formal neighbor in B_1 only **then**
$B_2 := B_2 \cup \{v\}$;
end if
if v has formal neighbor in B_2 only **then**
$B_1 := B_1 \cup \{v\}$;
end if
if v has formal neighbors in B_1 and B_2 **then**
;
end if
$G := G - \{v\}$;
remove isolated vertices from G;
end while
Output B_1, B_2;

Figure 4.4 shows an example when BIPGREEDY and BIPGREEDY2 behave differently; in case (a) BIPGREEDY decides randomly into which part it places v, in case (b) BIPGREEDY2 places v into part B_1, because no new restrictions on the neighbors of v will emerge then.

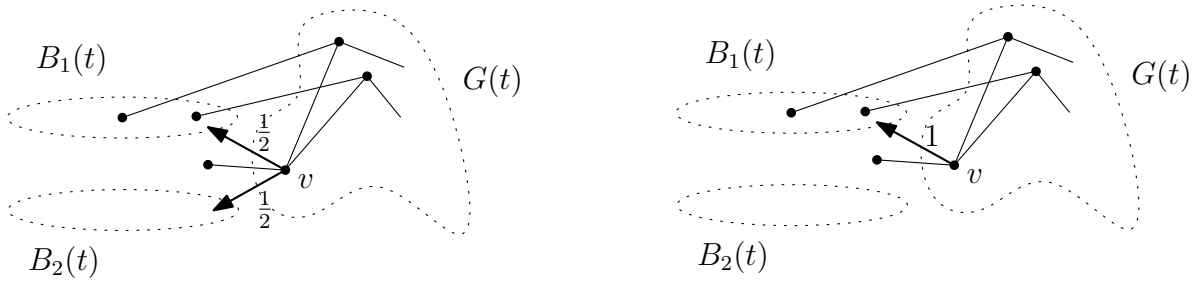


FIGURE 4.4: Placing vertex v (a) at random by BIPGREEDY (b) deterministically by BIPGREEDY2.

We did not analyze BIPGREEDY2. The ideas of the analysis would remain the same but the computation would become more technical and tedious. Another reason which discourages us from the analysis is that there obviously exist simple modifications of BIPGREEDY which perform even better; these take into account greater neighborhood of v.

Computer simulation shows very little improvement of BIPGREEDY2 compared to BIPGREEDY, approximately by 0.1%.

Remark 4.21. One may ask the following questions related to our problem.

- For which graph parameters ζ and for which models of random graphs (represented by a sequences of probability spaces $\{\mathcal{G}_n\}_n$) there exists a randomized greedy algorithm, which is asymptotically tight (i.e. for every $\varepsilon > 0$ it holds, that $1 - \varepsilon < \frac{\zeta_{\mathrm{alg}}(G)}{\zeta(G)} < 1 + \varepsilon$ a.a.s. for $G \in_u \mathcal{G}_n$)? When the graph parameter represents size of a certain structure in the graph, one is usually interested not only in finding (close-to-exact) value of the parameter, but also in finding the structure itself (or something close to it). We would like our algorithm to find the structure also.

  The first such algorithm was an algorithm for finding a maximum matching in sparse random graphs devised by Karp and Sipser [KS81]. The analysis of the algorithm was revised by Aranson, Frieze and Pittel [AFP98] recently. They showed that Karp-Sipser algorithm works even much better then we demanded—for graphs $G \in_u \mathcal{G}_{c/n,n}, c < e$ (members $\mathcal{G}_{p,n}$ are defined as random graphs on n vertices where each edge is contained with probability p, mutually independently) the algorithm does find the maximum matching asymptotically almost surely and for $G \in_u \mathcal{G}_{c/n,n}, c > e$ the error of the algorithm (i.e. the difference between the size of the maximum matching and the matching found by the algorithm) is $O(n^{1/5})$.

- We investigated what happens when $G \in_u \mathcal{G}^*_{3,n}$ is given on the input of BIP-GREEDY. What happens if $G \in_u \mathcal{G}_{3/n,n}$? One can, of course, generalize the question to other randomized algorithms.

Chapter 5

Consequences for graph homomorphisms

In this section we apply the results we obtained about bi(G) to graph homomorphisms. Our motivation is the Pentagon Conjecture due to Nešetřil, which is stated as Question 5.3 here.

We say that mapping $h : V(G) \to V(H)$ is a *homomorphism of graph G to graph H* if for every edge $\{v_1, v_2\} \in E(G)$ it holds that $\{h(v_1), h(v_2)\} \in E(H)$. We say that *graph G is homomorphic to graph H* if there is a homomorphism of graph G to graph H and denote the fact by $G \to H$.

In this section we shall present only terminology and statements which is needed to apply the results about bipartite graphs in cubic graphs to weakened version of the Pentagon Conjecture. Book [HN04] gives comprehensive survey on the topic.

Observation 5.1 (Transitivity of homomorphisms)**.** *If G, H, K are graphs $G \to H$ and $H \to K$ then $G \to K$.* □

The following lemma is useful when one wants to show that G is not homomorphic to H. First we prepare some notation for it.

A graph H is said to be *vertex-transitive* if for any two vertices $u, v \in V(H)$ there is a automorphism h of H (an automorphism of H is a one-to-one homomorphism of H to H) such that $h(u) = v$. Let G, K be graphs; we denote by $n(G, K)$ the maximum number of vertices in an induced subgraph of G that is homomorphic to K.

Lemma 5.2 (Proposition 1.22 in [HN04])**.** *Suppose G, H, K are graphs, where H is vertex-transitive. If $G \to H$ then*

$$\frac{n(G,K)}{|V(G)|} \geq \frac{n(H,K)}{|V(H)|}$$

5.1 The Pentagon Conjecture

Question 5.3 ([Neš99])**.** *Is it true that any cubic graph G with sufficiently large girth is homomorphic to C_5?*

If C_5 in Question 5.3 is replaced with C_3, the is answer affirmative (by Brook's Theorem), whenever the girth is at least 4.

There was a series of results which showed that a random cubic graph a.a.s. is not homomorphic to cycle of length 11 ([KNS01]), 9 ([WW01]) and 7 ([Hat05]). Note that from transitivity it follows that if G is not homomorphic to C_{2l+1} it also is not homomorphic to C_{2k+1} for any $k \geq l$; and thus Hatami's result implies the one of Kostochka, Nešetřil, Smolíková and the one of Wanless, Wormald. Since, by Fact 3.3, a random cubic graph on n vertices has girth at least k with probability tending to $\prod_{i=2}^{k-1} \exp\left(-\frac{2^i}{i+1}\right) > 0$, it holds that for any girth k there is a cubic graph which is not homomorphic to C_7.

First we apply McKay's result that a.a.s. $\alpha(G) \leq 0.4554n$ for a random cubic graph G. Substituting $K = K_1$ into Lemma 5.2 we get that a random cubic graph a.a.s. is not homomorphic to C_{13} since $0.4554 \leq \frac{n(G,K_1)}{|V(G)|} < \frac{n(C_{13},K_1)}{|V(C_{13})|} = \frac{6}{13} \doteq 0.4615$. Thus we were able to show that there is a cubic graph of arbitrary high girth which is not homomorphic to C_{13}. But we are also able to find limits of this technique. Due to Frieze and Suen's $\alpha(G) \geq 0.4327n$ and thus we will never be able (by applying Lemma 5.2 with $K = K_1$), even when sharpening McKay's bound on $\alpha(G)$, to prove that a random cubic graph is not homomorphic to C_7 ($\frac{n(C_7,K_1)}{|V(C_7)|} = \frac{3}{7} \doteq 0.4285$).

Now we use the same method for $K = K_2$. Then $n(G, K_2) = \mathrm{bi}(G)$. Due to Theorem 3.8 $\mathrm{bi}(G) \leq 0.9082n$ a.a.s. for a random cubic graph G. This means that G is a.a.s. not homomorphic to C_{11} since $\frac{n(C_{11},K_2)}{|V(C_{11})|} = \frac{10}{11} \doteq 0.9090 > 0.9082$. On the other hand, data obtained from computer simulation show that there is no hope disproving Pentagon Conjecture using this method (for a 5-cycle: $\frac{n(C_5,K_2)}{|V(C_5)|} = \frac{4}{5} < 0.81 \doteq \frac{\mathrm{bi}_{\mathrm{alg}}(G)}{n}$ a.a.s. for $G \in_u \mathcal{G}^*_{3,n}$). Theorem 4.16 itself is void in this case.

Bibliography

[AFP98] Jonathan Aronson, Alan Frieze, and Boris G. Pittel. Maximum matchings in sparse random graphs: Karp-Sipser revisited. *Random Structures Algorithms*, 12(2):111–177, 1998.

[AS00] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

[BC78] Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978.

[Bol01] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[Die05] Reinhard Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition, 2005.

[FRS95] Alan Frieze, A. J. Radcliffe, and Stephen Suen. Analysis of a simple greedy matching algorithm on random cubic graphs. *Combin. Probab. Comput.*, 4(1):47–66, 1995.

[FS94] Alan Frieze and Stephen Suen. On the independence number of random cubic graphs. *Random Structures & Algorithms*, 5(5):649–664, 1994.

[GJS76] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoret. Comput. Sci.*, 1(3):237–267, 1976.

[Hat05] Hamed Hatami. Random cubic graphs are not homomorphic to the cycle of size 7. *J. Combin. Theory Ser. B*, 93(2):319–325, 2005.

[HN04] Pavol Hell and Jaroslav Nešetřil. *Graphs and homomorphisms*, volume 28 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004.

[Jan02] Svente Janson. On concentration of probability. In *Contemporary combinatorics*, volume 10 of *Bolyai Soc. Math. Stud.*, pages 289–301. János Bolyai Math. Soc., Budapest, 2002.

[JŁR00] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[KNS01] Alexandr Kostochka, Jaroslav Nešetřil, and Petra Smolíková. Colorings and homomorphisms of degenerate and bounded degree graphs. *Discrete Math.*, 233(1-3):257–276, 2001. Graph theory (Prague, 1998).

[KS81] Richard M. Karp and Michael Sipser. Maximum matchings in sparse random graphs. In *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computing*, pages 364–375, 1981.

[McK82] Brendan D. McKay. Maximum bipartite subgraphs of regular graphs with large girth. In *Proc. Thirteenth Southeastern Conf. on Combinatorics, Graph Theory and Computing*, 1982.

[McK87] Brendan D. McKay. Independent sets in regular graphs of high girth. In *Proceedings of the Singapore conference on combinatorial mathematics and computing, 1986 (Singapore)*, volume 23, pages 179–185, 1987.

[MU05] Michael Mitzenmacher and Eli Upfal. *Probability and computing*. Cambridge University Press, Cambridge, 2005. Randomized algorithms and probabilistic analysis.

[Neš99] Jaroslav Nešetřil. Aspects of structural combinatorics (graph homomorphisms and their use). *Taiwanese J. Math.*, 3(4):381–423, 1999.

[Ros96] Sheldon M. Ross. *Stochastic processes*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1996.

[Ros00] Sheldon M. Ross. *Introduction to probability models*. Harcourt/Academic Press, Burlington, MA, seventh edition, 2000.

[Šám06] Robert Šámal. *On XY mappings*. PhD thesis, Charles University in Prague, 2006.

[Wor99a] Nick C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, pages 73–155. PWN, Warsaw, 1999.

[Wor99b] Nick C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.

[WW01] Ian M. Wanless and Nick C. Wormald. Regular graphs with no homomorphisms onto cycles. *J. Combin. Theory Ser. B*, 82(1):155–160, 2001.