Martin Nägele

The Lovász Local Lemma

Selected Variants and Applications

Bachelor Thesis in Mathematics

Advisers:
Prof. Dr. B. Sudakov,
Dr. J. Volec

July 15, 2015

Department of Mathematics
ETH Zürich

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Abstract

The aim of this bachelor thesis is to explore various aspects of the Lovász Local Lemma, a powerful tool from the area of probabilistic combinatorics.

We introduce and prove the Lovász Local Lemma and some of its variants in Chapter 1, preparing their application in the subsequent chapters. More precisely, besides the standard Local Lemma, we introduce the so-called lopsided Local Lemma and their respective symmetric versions. Moreover, we present a variation giving upper instead of lower bounds for the probability in question.

In Chapter 2, we study a result of Alon, Krivelevich and Sudakov on the chromatic number of graphs with sparse neighbourhoods [1]. Amongst others, the symmetric Lovász Local Lemma is applied iteratively to gradually split a given graph into well-behaved subgraphs.

Chapter 3 establishes a basis for applications of the lopsided Local Lemma in the probability spaces of random matchings in complete bipartite graphs and cliques, which is due to Lu and Székely [11, 12]. Using the developed machinery, we prove a sufficient condition for disjointly packing two hypergraphs into a complete hypergraph. Surprisingly, by combining the lopsided Local Lemma and the variation giving upper bounds, this framework can also be used for asymptotic counting, which we apply to the asymptotic enumeration of d-regular graphs.

The last chapter, Chapter 4, is devoted to algorithmic aspects of the Local Lemma. We study a rather general algorithmic proof of the Local Lemma by Moser and Tardos [17] before exploring a successful algorithmic approach to acyclic edge colourings using entropy compression by Esperet and Parreau [7].

Zusammenfassung

Das Ziel dieser Bachelorarbeit ist es, verschiedene Aspekte des Lovász Local Lemma zu beleuchten.

Das Lovász Local Lemma und einige Varianten davon werden in Kapitel 1 eingeführt, um die Anwendungen in späteren Kapiteln vorzubereiten. Neben dem gewöhnlichen Lemma betrachten wir das sogenannte lopsided Local Lemma sowie symmetrische Varianten davon. Zusätzlich wird eine Variation präsentiert, die eine obere Schranke für die gefragte Wahrscheinlichkeit angibt.

In Kapitel 2 betrachten wir ein Resultat von Alon, Krivelevich und Sudakov über die chromatische Zahl von Graphen, in denen jede Nachbarschaft nur wenige Kanten enthält [1]. Hier wird das Lovász Local Lemma unter anderem iterativ angewendet, um einen gegebenen Graphen Schritt für Schritt in Teilgraphen mit bestimmten Eigenschaften zu zerlegen.

Kapitel 3 bildet eine Basis für Anwendungen des lopsided Local Lemma in Wahrscheinlichkeitsräumen von Zufallsmatchings in vollständigen und vollständigen bipartiten Graphen, die von Lu und Székely [11, 12] entwickelt wurde. Darauf aufbauend werden zwei Resultate präsentiert: Einerseits hinreichende Bedingungen für die disjunkte Einbettung zweier Hypergraphen in einen vollständigen Hypergraphen, andererseits erlaubt die gegebene Konfiguration eine asymptotische Bestimmung der Anzahl d-regulärer Graphen.

Das letzte Kapitel, Kapitel 4, ist algorithmischen Aspekten des Lovász Local Lemma gewidmet. Neben einem relativ allgemeinen algorithmischen Beweis des Lemmas von Moser und Tardos [17] wird ein erfolgreicher algorithmischer Zugang zu azyklischen Kantenfärbungen in Graphen präsentiert, der von Esperet und Parreau [7] gefunden wurde.

Acknowledgement

I would like to express my appreciation to Prof. Dr. Benny Sudakov for introducing me to the topic and for making this bachelor thesis possible.

I am deeply indebted to Dr. Jan Volec for his help throughout my work on the thesis. Be it the numerous discussions of my sources, his proficient and honest advice while writing the thesis or the very careful proofreading – he always took the time and did not spare any effort to sort out all my questions. I am grateful for his guidance.

Thanks as well to Alexander Breuss for his suggestions on the layout of this thesis.

Last but not least, I want to thank my parents Renate and Matthias Nägele for their countless support throughout my studies, and Pia Lattner for her continuous love and encouragement that helped leading me through all the ups and downs.

Contents

Chapter 1

The Lovász Local Lemma and its Variants

In applications of the probabilistic method, one usually wants to prove that a random object in some probability space has a positive probability of having certain properties. If each of the properties appears with positive probability, then assuming independence of the properties directly gives a positive probability of having all the properties.

In the setting of the Lovász Local Lemma, the target object's properties are given as a list of events that we want to avoid, so the goal is to get a positive lower bound for the probability of avoiding all the events. As before, the case of stochastically independent properties does not require any effort while dependencies among the properties forbid the immediate argument.

Nonetheless, under certain assumptions on the dependencies, it is still possible to get a non-zero lower bound. This chapter presents various versions of the Lovász Local Lemma, each stating sufficient conditions for a result of the described type to hold.

1.1 The basic Local Lemma

Intuitively, a conclusion similar to the one stated in the introduction above should be possible if there are only few dependencies among the forbidden properties. Indeed, the following lemma concretises this idea.

Lemma 1.1 (Symmetric Lovász Local Lemma, [2]). *Let $A_1, \dots, A_n$ be events in an arbitrary probability space such that each event is mutually independent from all except at most d of the others. Moreover, assume that $\Pr[A_i] \leqslant p$ for some $p \in [0,1)$ and all $i \in [n]$. If $ep(d+1) \leqslant 1$, where e denotes the Euler number, then $\Pr\left[\bigcap_{i=1}^n \overline{A_i}\right] > 0$.*

We will derive this symmetric version from a more general version where the number of dependencies per event is not uniformly bounded. The deduction is given on page 4.

In order to collect information about mutual dependencies in a set of events, we use the model of a *dependency graph*.

Definition 1.2. *For events $A_1, \ldots, A_n$ in a probability space, a* dependency graph *of $A_1, \ldots, A_n$ is a graph $G = (V, E)$ with $V = [n]$ such that for all $j \in [n]$, the event A_j is mutually independent from the events in $\{A_i \mid \{i, j\} \notin E\}$.*

Using this definition, the condition that each event is mutually independent from all but at most d of the others in Lemma 1.1 can be translated to the condition that a dependency graph of $A_1, \ldots, A_n$ has maximum degree at most d. In the general Local Lemma, this uniform bound is replaced by different assumptions.

Note that the complete graph is always a dependency graph – but it does not contain any information about mutual dependencies. On the other hand, an empty dependency graph encodes that all events are mutually independent.

Lemma 1.3 (Lovász Local Lemma, [2]). *Let $A_1, \ldots, A_n$ be events in an arbitrary probability space with a dependency graph G. Assume that there are numbers $x_1, \ldots, x_n$ such that for all $i \in [n]$, we have $0 \leqslant x_i < 1$ and*

$$\Pr[A_i] \leqslant x_i \cdot \prod_{j \in N_G(i)} (1 - x_j). \tag{1.1}$$

Then

$$\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] \geqslant \prod_{i=1}^{n} (1 - x_i), \tag{1.2}$$

so in particular, with positive probability no event A_i occurs.

Proof. By iteratively applying the definition of conditional probability, we see that

$$\begin{aligned}\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] &= \Pr\left[\overline{A_n} \,\middle|\, \bigcap_{i=1}^{n-1} \overline{A_i}\right] \cdot \Pr\left[\bigcap_{i=1}^{n-1} \overline{A_i}\right] = \ldots = \\ &= \prod_{j=1}^{n} \Pr\left[\overline{A_j} \,\middle|\, \bigcap_{i=1}^{j-1} \overline{A_i}\right] = \prod_{j=1}^{n} \left(1 - \Pr\left[A_j \,\middle|\, \bigcap_{i=1}^{j-1} \overline{A_i}\right]\right). \end{aligned} \tag{1.3}$$

Consequently, it suffices to prove $\Pr\left[A_j \,\middle|\, \bigcap_{i=1}^{j-1} \overline{A_i}\right] \leqslant x_j$ in order to conclude (1.2) from (1.3). We prove a more general inequality, namely that for every $j \in [n]$ and every set $I \subseteq [n]$, we have

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] \leqslant x_j. \tag{1.4}$$

The proof is by fixing j and doing induction on $|I|$. If $|I| = 0$, we get the above directly from the assumption:

$$\Pr[A_j] \leqslant x_j \cdot \prod_{i \in N_G(j)} (1 - x_i) \leqslant x_j.$$

For the inductive step, assume we know the statement for all index sets of cardinality smaller than the cardinality of a set I. Let $I_0 \coloneqq I \setminus N_G(j)$ and $I_1 \coloneqq I \cap N_G(j)$ so that we can write

$$\Pr\left[A_j \,\middle|\, \bigcap_{i\in I} \overline{A_i}\right] = \Pr\left[A_j \,\middle|\, \bigcap_{i\in I_1} \overline{A_i} \cap \bigcap_{i\in I_0} \overline{A_i}\right]. \tag{1.5}$$

Note that if $j \in I$, then $\Pr\left[A_j \,\middle|\, \bigcap_{i\in I} \overline{A_i}\right] = 0$, so the inequality (1.4) trivially holds. Therefore, we may assume $j \notin I$ so that I_0 is a set of indices such that A_j is mutually independent from $\{A_i \mid i \in I_0\}$.

For arbitrary events X, Y and Z in a probability space, the definition of conditional probability gives

$$\begin{aligned} \Pr[X \cap Y \cap Z] = \Pr[X \cap Y \mid Z] \cdot \Pr[Z] \quad \text{and} \quad \Pr[Y \cap Z] = \Pr[Y \mid Z] \cdot \Pr[Z], \\ \implies \quad \Pr[X \mid Y \cap Z] = \frac{\Pr[X \cap Y \cap Z]}{\Pr[Y \cap Z]} = \frac{\Pr[X \cap Y \mid Z]}{\Pr[Y \mid Z]}. \end{aligned} \tag{1.6}$$

Applying this to (1.5) with $X = A_j$, $Y = \bigcap_{i\in I_1} \overline{A_i}$ and $Z = \bigcap_{i\in I_0} \overline{A_i}$, we get

$$\Pr\left[A_j \,\middle|\, \bigcap_{i\in I} \overline{A_i}\right] = \frac{\Pr\left[A_j \cap \bigcap_{i\in I_1} \overline{A_i} \,\middle|\, \bigcap_{i\in I_0} \overline{A_i}\right]}{\Pr\left[\bigcap_{i\in I_1} \overline{A_i} \,\middle|\, \bigcap_{i\in I_0} \overline{A_i}\right]}. \tag{1.7}$$

We treat numerator and denominator separately. For the numerator, note that we can first apply monotonicity of the probability measure and then use the fact that the events A_i with $i \in I_0$ are mutually independent from A_j. This gives

$$\Pr\left[A_j \cap \bigcap_{i\in I_1} \overline{A_i} \,\middle|\, \bigcap_{i\in I_0} \overline{A_i}\right] \leqslant \Pr\left[A_j \,\middle|\, \bigcap_{i\in I_0} \overline{A_i}\right] = \Pr[A_j] \leqslant x_j \cdot \prod_{i\in N_G(j)} (1 - x_i), \tag{1.8}$$

where we used the assumption (1.1) in the last step. For the denominator, we let $I_1 = \{i_1, \ldots, i_r\}$ and iteratively apply (1.6) to get

$$\begin{aligned} \Pr\left[\bigcap_{i\in I_1} \overline{A_i} \,\middle|\, \bigcap_{i_0\in I_0} \overline{A_{i_0}}\right] &= \Pr\left[\overline{A_{i_1}} \cap \ldots \cap \overline{A_{i_r}} \,\middle|\, \bigcap_{i_0\in I_0} \overline{A_{i_0}}\right] = \\ &= \Pr\left[\overline{A_{i_1}} \cap \ldots \cap \overline{A_{i_{r-1}}} \,\middle|\, \overline{A_{i_r}} \cap \bigcap_{i_0\in I_0} \overline{A_{i_0}}\right] \cdot \Pr\left[\overline{A_{i_r}} \,\middle|\, \bigcap_{i_0\in I_0} \overline{A_{i_0}}\right] = \ldots = \\ &= \prod_{\ell=1}^{r} \Pr\left[\overline{A_{i_\ell}} \,\middle|\, \bigcap_{k=\ell+1}^{r} \overline{A_{i_k}} \cap \bigcap_{i_0\in I_0} \overline{A_{i_0}}\right] = \prod_{\ell=1}^{r} \left(1 - \Pr\left[A_{i_\ell} \,\middle|\, \bigcap_{i\in I_{\ell+1}} \overline{A_i}\right]\right), \end{aligned}$$

where $I_{\ell+1} \coloneqq \{i_{\ell+1}, \ldots, i_r\} \cup I_0$ for each $\ell \in [r]$ (with $I_{r+1} = I_0$, of course). Now note that for each $\ell \in [r]$, $|I_{\ell+1}| < |I|$ because at least $i_1 \notin I_{\ell+1}$. Hence the inductive assumption can be applied, resulting in

$$\prod_{\ell=1}^{r} \left(1 - \Pr\left[A_{i_\ell} \,\middle|\, \bigcap_{i\in I_{\ell+1}} \overline{A_i}\right]\right) \geqslant \prod_{\ell=1}^{r} (1 - x_{i_\ell}) \geqslant \prod_{i\in N_G(j)} (1 - x_i)$$

because $\{i_1, \ldots, i_r\} = I_1 \subseteq N_G(j)$. Plugging the results for the numerator and denominator into (1.7), we get

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] \leqslant \frac{x_j \cdot \prod_{i \in N_G(j)}(1 - x_i)}{\prod_{i \in N_G(j)}(1 - x_i)} = x_j,$$

which is (1.4), so the inductive step and hence also the proof of the Lovász Local Lemma is complete. □

Having the general statement of the Lovász Local Lemma (Lemma 1.3) at hand, we can now deduce the symmetric version, Lemma 1.1.

Proof of Lemma 1.1. We show that the condition of Lemma 1.3 is satisfied for the choice $x_1 = \ldots = x_n = \frac{1}{d+1}$.

As already remarked, the condition that every event A_i is independent from all but at most d of the others translates to having the maximal degree in the dependency graph G for the events $A_1, \ldots, A_n$ bounded from above by d. In other words, we have $|N_G(i)| \leqslant d$ for all $i \in [n]$, which gives

$$x_i \cdot \prod_{j \in N_G(i)} (1 - x_j) = \frac{1}{d+1} \cdot \left(1 - \frac{1}{d+1}\right)^{|N_G(i)|} \geqslant \frac{1}{d+1} \cdot \left(1 - \frac{1}{d+1}\right)^{d}.$$

By the well-known inequality $1 + x \leqslant e^x$, we have $1 - \frac{1}{d+1} = \left(1 + \frac{1}{d}\right)^{-1} \geqslant e^{-1/d}$, so we get

$$\frac{1}{d+1} \cdot \left(1 - \frac{1}{d+1}\right)^{d} \geqslant \frac{1}{e(d+1)} \geqslant p \geqslant \Pr[A_i],$$

where the last two inequalities are assumptions of Lemma 1.1. Combining the above two chains of inequalities gives

$$\Pr[A_i] \leqslant x_i \cdot \prod_{j \in N_G(i)} (1 - x_j),$$

so the assumption of Lemma 1.3 is satisfied, giving the conclusion

$$\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] \geqslant \prod_{i=1}^{n} (1 - x_i) = \left(1 - \frac{1}{d+1}\right)^{n} > 0$$

as desired. □

1.2 The Lopsided Local Lemma

There is an extension of the Lovász Local Lemma introduced by Erdős and Spencer in [6] called the lopsided Local Lemma, which originates from a careful study of the proof.

In fact, the only step that uses the assumption that G contains the information of mutual dependencies among the events is the equality in (1.8). It would be enough to have the graph G such that for every $j \in [n]$ and every index set I_0 not containing and any neighbours of j in G, we have

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I_0} \overline{A_i}\right] \leqslant \Pr[A_j].$$

This gives rise to the following definition.

Definition 1.4 ([6]). *For events $A_1, \ldots, A_n$ in a probability space, a graph $G = (V, E)$ is called* negative dependency graph *for the events $A_1, \ldots, A_n$ if $V = [n]$ and E is such that for all $j \in V$ and all sets $I \subseteq V \setminus N_G(j)$, we have*

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] \leqslant \Pr[A_j]$$

whenever the conditional probability on the left-hand side is well-defined.

Sometimes, a negative dependency graph is also called *lopsided dependency graph*. As pointed out, having a negative dependency graph is enough to draw the conclusion of the Lovász Local Lemma.

Lemma 1.5 (Lopsided Local Lemma). *Let $A_1, \ldots, A_n$ be events in an arbitrary probability space with a negative dependency graph G. Assume that there are numbers $x_1, \ldots, x_n$ such that for all $i \in [n]$, we have $0 \leqslant x_i < 1$ and*

$$\Pr[A_i] \leqslant x_i \cdot \prod_{j \in N_G(i)} (1 - x_j).$$

Then

$$\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] \geqslant \prod_{i=1}^{n} (1 - x_i).$$

Proof. As already indicated, the proof of Lemma 1.5 is almost identical to the proof of Lemma 1.3, the only difference being that the equation sign in (1.8) is replaced by an inequality sign, justified by the assumption that G is a negative dependency graph for $A_1, \ldots, A_n$. Therefore, we do not repeat the arguments here. □

Similarly to the case of the "standard" Lovász Local Lemma, the lopsided version allows a specification to a symmetric version, given below.

Lemma 1.6 (Symmetric Lopsided Local Lemma, [6]). *Let $A_1, \ldots, A_n$ be events in an arbitrary probability space with a negative dependency graph G of maximum degree d. Moreover, assume that $\Pr[A_i] \leqslant p$ for some $p \in [0, 1)$ and all $i \in [n]$. If $ep(d + 1) \leqslant 1$, where e denotes the Euler number, then $\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] > 0$.*

Proof. The symmetric lopsided Lovász Local Lemma can be derived from the lopsided Local Lemma by following the very same steps as when we derived the symmetric Lovász Local Lemma from the standard Local Lemma on page 4, so we do not repeat the steps here. □

One thing that can easily be observed is that the lopsided Local Lemma indeed generalises the "standard" Local Lemma.

Proposition 1.7. *For events $A_1, \dots, A_n$ in an arbitrary probability space, a dependency graph is also a negative dependency graph.*

Proof. Let G be a dependency graph and fix $j \in [n]$ and $I \subseteq V \setminus N_G(i)$. If $j \in I$, then

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] = \frac{\Pr\left[A_j \cap \overline{A_j} \cap \bigcap_{i \in I \setminus \{j\}} \overline{A_i}\right]}{\Pr\left[\bigcap_{i \in I} \overline{A_i}\right]} = 0,$$

and the condition trivially holds. If $j \notin I$, then A_j is independent from all A_i with $i \in I$, hence

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] = \Pr[A_j],$$

so in this case the inequality constraints for G being a negative dependency graph are satisfied with equality. □

In fact, the above proposition affirms that every dependency graph can be seen as a negative dependency graph. However, the question of how to find non-trivial negative dependency graphs, i. e. negative dependency graphs that are not already dependency graphs, remains. In Chapter 3, we prove that if the probability space consists of all maximum cardinality matchings in a complete or complete bipartite graph, then certain so-called canonical events inherit a structure that allows defining a natural negative dependency graph.

1.3 Local Lemma-type Upper Bounds

In some applications, for example those of asymptotic counting presented in Chapter 3, it is also desirable to have upper bounds on the probability $\Pr\left[\bigcap_{i=1}^n \overline{A_i}\right]$. To this end, Lu and Székely ([12]) introduced a framework that allows to prove such bounds, starting with the following definition.

Definition 1.8 ([12]). *For events $A_1, \dots, A_n$ in an arbitrary probability space and an $\varepsilon \in (0,1)$, a graph $G = (V, E)$ with $V = [n]$ is an* ε-near positive dependency graph *for the events $A_1, \dots, A_n$ if*

(i) $\Pr[A_i \cap A_j] = 0$ *for all* $\{i, j\} \in E$.

(ii) For any $j \in [n]$ and any subset $I \subseteq V \setminus (N_G(j) \cup \{j\})$, we have

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] \geqslant (1 - \varepsilon) \Pr[A_j]$$

whenever the conditional probability on the left-hand side is well-defined.

Using this definition, the analogue of the Local Lemma giving upper bounds is the following:

Lemma 1.9 ([12]). *Let $A_1, \ldots, A_n$ be events in an arbitrary probability space with an ε-near positive dependency graph $G = (V, E)$ for some $\varepsilon \in (0, 1)$. Then we have*

$$\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] \leqslant \prod_{i=1}^{n} \Big(1 - (1 - \varepsilon) \Pr[A_i]\Big). \tag{1.9}$$

Proof. If $\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] = 0$, then the statement is true for sure as every factor on the right-hand side is non-negative. Otherwise, by iteratively applying the definition of conditional probability, we get

$$\begin{aligned}\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] &= \Pr\left[\overline{A_1} \,\middle|\, \bigcap_{i=2}^{n} \overline{A_i}\right] \cdot \Pr\left[\bigcap_{i=2}^{n} \overline{A_i}\right] = \ldots = \\ &= \prod_{j=1}^{n} \Pr\left[\overline{A_j} \,\middle|\, \bigcap_{i=j+1}^{n} \overline{A_i}\right] = \prod_{j=1}^{n} \left(1 - \Pr\left[A_j \,\middle|\, \bigcap_{i=j+1}^{n} \overline{A_i}\right]\right). \end{aligned} \tag{1.10}$$

In order to derive the bound in (1.9) from (1.10), it is enough to show that we have $\Pr\left[A_j \,\middle|\, \bigcap_{i=j+1}^{n} \overline{A_i}\right] \geqslant (1 - \varepsilon) \Pr[A_j]$. We prove a more general inequality, namely that for all $j \in [n]$ and $I \subseteq [n] \setminus j$,

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] \geqslant (1 - \varepsilon) \Pr[A_j]. \tag{1.11}$$

Fix j and I and set $I_0 \coloneqq I \setminus N_G(j)$ and $I_1 \coloneqq I \cap N_G(j)$. Applying the definition of conditional probability (in fact, we can use (1.6) with $X = A_j$, $Y = \bigcap_{i \in I_1} \overline{A_i}$ and $Z = \bigcap_{i \in I_0} \overline{A_i}$), we get

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] = \frac{\Pr\left[A_j \cap \bigcap_{i \in I_1} \overline{A_i} \,\middle|\, \bigcap_{i \in I_0} \overline{A_i}\right]}{\Pr\left[\bigcap_{i \in I_1} \overline{A_i} \,\middle|\, \bigcap_{i \in I_0} \overline{A_i}\right]}. \tag{1.12}$$

For $i \in I_1$, the definition of I_1 provides $\{i, j\} \in E(G)$, so by the first condition of an ε-near positive dependency graph,

$$\Pr\left[A_j \cap \overline{A_i}\right] = \Pr[A_j] - \Pr[A_j \cap A_i] = \Pr[A_j].$$

This extends to

$$\Pr\left[A_j \cap \bigcap_{i \in I_1} \overline{A_i}\right] = \Pr[A_j] - \Pr\left[A_j \cap \bigcup_{i \in I_1} A_i\right] = \Pr[A_j]$$

and obviously also holds if the events are conditioned on $\bigcap_{i \in I_0} \overline{A_i}$, so

$$\Pr\left[A_j \cap \bigcap_{i \in I_1} \overline{A_i} \,\middle|\, \bigcap_{i \in I_0} \overline{A_i}\right] = \Pr\left[A_j \,\middle|\, \bigcap_{i \in I_0} \overline{A_i}\right].$$

Plugging this into (1.12) and bounding the probability in the denominator from above by 1 results in

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] \geqslant \Pr\left[A_j \,\middle|\, \bigcap_{i \in I_0} \overline{A_i}\right] \geqslant (1 - \varepsilon) \Pr[A_j].$$

The last inequality follows from the second condition in the definition of an ε-near positive dependency graph because $I_0 \subseteq I \setminus N_G(j) \subseteq [n] \setminus (N_G(j) \cup \{j\})$. This proves (1.11) and hence finishes the proof of Lemma 1.9. □

The purpose of Lemma 1.9 may not be obvious as it stands here for itself – in fact, its strength comes from combination with the lower bound that we get in the standard Local Lemma: In certain settings, the lower and upper bounds match at least asymptotically, giving the option to actually count the number of objects avoiding all events A_j. This is treated more concretely in Chapter 3.

Chapter 2

Colouring Graphs with sparse Neighbourhoods

The *chromatic number* $\chi(G)$ of a graph G is the minimal number of colours that are necessary to colour all vertices of G without generating a monochromatic edge, i. e. an edge with both endpoints of the same colour.

Calculating the chromatic number for a general graph G is an NP-hard problem ([10]), but there are lots of results bounding the chromatic number. The easiest one is that if the graph G has maximum degree Δ, then $\chi(G) \leqslant \Delta + 1$. Improvements are possible if, for example, no neighbourhood of any vertex contains edges, i. e. if the graph is triangle-free: Johansson ([9]) proved that under this assumption $\chi(G) = \mathcal{O}(\Delta/\log \Delta)$, see Theorem 2.3.

Alon, Krivelevich and Sudakov ([1]) provide a result of similar spirit if the number of edges in every neighbourhood is not zero but bounded from above by Δ^2/f for some f with $2 \leqslant f \leqslant \Delta^2$: They prove $\chi(G) = \mathcal{O}(\Delta/\log f)$ in this case.

From a technical point of view, the key ingredient for their proof is the Lovász Local Lemma. In particular, they use an interesting iterative approach, progressing towards the final goal step by step. This chapter presents a proof of the mentioned result by Alon, Krivelevich and Sudakov, closely following their exposition in [1].

2.1 Theorem Statement and Proof Outline

Theorem 2.1. *There exists an absolute positive constant c such that the following holds: For any graph $G = (V, E)$ with maximum degree Δ and an f with $2 \leqslant f \leqslant \Delta^2$ such that the neighbourhood $N_G(v)$ of every vertex $v \in V$ spans at most Δ^2/f edges, the chromatic number satisfies*

$$\chi(G) \leqslant c \cdot \frac{\Delta}{\log f}.$$

As indicated in the introduction, the theorem may be read as the asymptotic statement $\chi(G) = \mathcal{O}(\Delta/\log f)$ for $|V| \to \infty$. All asymptotics in this section are with respect to the number of vertices going to infinity, so if there occur asymptotics without further specification, they are understood in this sense.

The proof of Theorem 2.1 is split into the following two cases for a fixed $\varepsilon > 0$:

Case 1: $f \geqslant \Delta^{4\varepsilon}$.

Having f large corresponds to having a small number of edges in each neighbourhood. Using the Lovász Local Lemma, we are going to split the graph into $\Theta(\Delta^{1-\varepsilon}/\varepsilon^2)$ vertex-disjoint triangle-free induced subgraphs with a bound of $\mathcal{O}(\Delta^{\varepsilon})$ on the maximum degree. For each of them, we can apply a result of Johansson (Theorem 2.3) to get a colouring with $\mathcal{O}(\Delta^{\varepsilon}/\log\Delta^{\varepsilon})$ many colours.

Colouring each of the subgraphs from its own colour palette, we get a proper colouring that asymptotically uses the desired number of colours.

This is elaborated in Section 2.2.

Case 2: $f < \Delta^{4\varepsilon}$.

This case corresponds to allowing many edges in each neighbourhood. As before, the idea is to split the graph into vertex-disjoint pieces with the goal of getting induced subgraphs that satisfy the assumption of Case 1.

By an application of the Lovász Local Lemma, we will be able to prove the existence of a splitting into two pieces such that each of them is "closer" to meeting the conditions of Case 1 than the whole graph we started with. Iteratively applying this procedure, we will eventually be left with a system of subgraphs all satisfying the conditions – hence colouring each of them from their own colour palette by the construction in Case 1 is possible.

Last but not least, we will see that the number of colours stays in the bound of $\mathcal{O}(\Delta/\log f)$ that we are aiming at.

This second case is covered by Section 2.3.

Besides just proving Theorem 2.1, it is also worth asking how tight the bound provided by the theorem is. One can prove that it is asymptotically tight, i. e. tight up to the constant c. Section 2.4 is devoted to proving this fact.

2.2 The first Case: $f \geqslant \Delta^{4\varepsilon}$

We prove the following result:

Theorem 2.2. *Let $\varepsilon > 0$ and let $G = (V, E)$ be a graph with maximum degree Δ in which the neighbourhood $N_G(v)$ of any vertex $v \in V$ spans at most $\Delta^{2-4\varepsilon}$ edges. Then the chromatic number of G satisfies*

$$\chi(G) = \mathcal{O}\left(\frac{\Delta}{\varepsilon^3 \log\Delta}\right).$$

Note that if we fix ε, then this is indeed a special case of Theorem 2.1 with the additional restriction that $f \geqslant \Delta^{4\varepsilon}$, resulting in the bound of $\Delta^2/f \leqslant \Delta^{2-4\varepsilon}$ for the number of edges in each neighbourhood.

As already outlined, the proof of Theorem 2.2 builds upon a result of Johansson, which is the following:

Theorem 2.3 (Johansson, [9]). *If G is a triangle-free graph on n vertices with maximum degree Δ, then*

$$\chi(G) = \mathcal{O}\left(\frac{\Delta}{\log \Delta}\right)$$

for $n \to \infty$.

Coincidentally, the proof of this theorem also uses the Lovász Local Lemma, but nevertheless, we do not present it here and only refer to [15, Thm. 13.1].

The subsequent lemma is the key to establishing a setting in which we can apply Theorem 2.3, it reveals a splitting of the graph into vertex-disjoint and triangle-free induced subgraphs.

Lemma 2.4. *Let $\varepsilon \in (0,1)$ and let $G = (V,E)$ be a graph with maximum degree at most Δ in which the neighbourhood $N_G(v)$ of any vertex $v \in V$ spans at most $\Delta^{2-4\varepsilon}$ edges. Then there exists a partition of the vertex set $V = V_1 \dot\cup \ldots \dot\cup V_k$ with $k = \Theta(\Delta^{1-\varepsilon}/\varepsilon^2)$ such that for any $i \in [k]$, the induced subgraph $G[V_i]$ is triangle-free and has maximum degree $\mathcal{O}(\Delta^{\varepsilon})$.*

Proof. Note that the given setting allows us to assume that Δ is large enough whenever we need it.

For a vertex $v \in V$, we call a neighbour $u \in N_G(v)$ a *bad neighbour* if u and v have at least $\Delta^{1-2\varepsilon}$ common neighbours, i. e. if $|N_G(u) \cap N_G(v)| \geqslant \Delta^{1-2\varepsilon}$. Otherwise, we call u a *good neighbour*.

We consider the probability space of all partitions of G into $\Delta^{1-\varepsilon}$ parts so that a random element in this space has each vertex assigned randomly and independently to one of the parts. For each $v \in V$, define the following three events in this probability space:

A_v: Vertex v has at least $2\Delta^{\varepsilon}$ neighbours inside its part of the partition.

B_v: Vertex v has more than $10/\varepsilon$ bad neighbours inside its part of the partition.

C_v: The set of good neighbours of v inside its part spans at least $100/\varepsilon^2$ edges.

We want to apply the symmetric Lovász Local Lemma, Lemma 1.1, to see that there is a partition avoiding all events $\{A_v, B_v, C_v \mid v \in V\}$.

To ensure that we satisfy the assumptions of Lemma 1.1, we first address mutual dependencies. Thereto, fix a vertex v and note that each of the events A_v, B_v and C_v is determined by v and the vertices of distance one from v, so these events are for sure mutually independent from all events corresponding to vertices of distance more than two. As the maximum degree in G is at most Δ, there are at most $\Delta^2 + \Delta$ vertices of distance at most two from v, so we may choose $d = 3(\Delta^2 + \Delta)$ in Lemma 1.1.

Moreover, we need to bound the probabilities of each of the events. To do so, fix a vertex v again.

For $\Pr[A_v]$, let X be the random variable counting the number of neighbours of vertex v inside its part. Each of the $|N_G(v)|$ neighbours of v belongs to the same part as v with probability $1/\Delta^{1-\varepsilon}$ independently from the others, so X is binomially distributed with parameters $\deg_G(v)$ and $1/\Delta^{1-\varepsilon}$.

We know that $\deg_G(v) \leqslant \Delta$, so $\mathbb{E}[X] = \deg_G(v) \cdot \frac{1}{\Delta^{1-\varepsilon}} \leqslant \Delta^{\varepsilon}$. Applying the Chernoff estimate (i) in Theorem A.1 with $a = np = \Delta^{\varepsilon}$, we get

$$\Pr[X > 2\Delta^{\varepsilon}] \leqslant e^{\Delta^{\varepsilon}(1-2\log(2))} < \frac{1}{\Delta^3}, \tag{2.1}$$

where the second inequality holds for Δ large enough because $1 - 2\log(2) < 0$.

To bound $\Pr[B_v]$, we first bound the total number of bad neighbours. If v has b bad neighbours, then there are at least $\frac{b}{2}\Delta^{1-2\varepsilon}$ edges inside $N_G(v)$, so we get $\frac{b}{2}\Delta^{1-2\varepsilon} \leqslant \Delta^{2-4\varepsilon}$ or equivalently $b \leqslant 2\Delta^{1-2\varepsilon}$.

For a given set of $10/\varepsilon$ bad neighbours of v, the probability that they all belong to the same part as v is $(1/\Delta^{1-\varepsilon})^{10/\varepsilon}$, so a union bound gives

$$\begin{aligned}\Pr[B_v] &\leqslant \binom{2\Delta^{1-2\varepsilon}}{10/\varepsilon} \cdot \left(\frac{1}{\Delta^{1-\varepsilon}}\right)^{\frac{10}{\varepsilon}} < \frac{\left(2\Delta^{1-2\varepsilon}\right)^{\frac{10}{\varepsilon}}}{\left(\frac{10}{\varepsilon}\right)!} \cdot \frac{1}{\left(\Delta^{1-\varepsilon}\right)^{\frac{10}{\varepsilon}}} = \\ &= \frac{2^{10/\varepsilon}}{\left(\frac{10}{\varepsilon}\right)!} \cdot \Delta^{-10} < \frac{1}{\Delta^3},\end{aligned} \tag{2.2}$$

where the last inequality again holds for Δ large enough.

Last but not least, we bound $\Pr[C_v]$. Let V_0 be the set of good neighbours of v inside its part. If $G[V_0]$ spans at least $100/\varepsilon^2$ edges, which is the event that we are considering, then we claim that at least one of the following two situations occurs: Either, there is a vertex of degree at least $9/\varepsilon$ or there is a matching of size at least $9/\varepsilon$.

To see this, assume the maximum degree in $G[V_0]$ is smaller than $9/\varepsilon$, i. e. there is no vertex of degree at least $9/\varepsilon$. Then Vizing's theorem (see [4, Theorem 5.3.2]) provides that $\chi'(G[V_0]) < \frac{9}{\varepsilon} + 1$, so in an optimal proper edge-colouring of $G[V_0]$, there is a colour class of size at least

$$\frac{100}{\varepsilon^2}\left(\frac{9}{\varepsilon} + 1\right)^{-1} = \frac{100}{9+\varepsilon} \cdot \frac{\varepsilon}{\varepsilon^2} \geqslant \frac{9}{\varepsilon}.$$

The inequality holds because $\varepsilon \leqslant 1$. A colour class in a proper edge-colouring forms a matching, so the claim follows.

Let $C_v^{(1)}$ and $C_v^{(2)}$ be the events that the graph $G[V_0]$ contains a vertex of degree at least $9/\varepsilon$ and that the graph $G[V_0]$ contains a matching of size at least $9/\varepsilon$, respectively. By the above claim, $C_v \subseteq C_v^{(1)} \cup C_v^{(2)}$, hence a union bound gives

$$\Pr[C_v] \leqslant \Pr\left[C_v^{(1)}\right] + \Pr\left[C_v^{(2)}\right],$$

so we aim for bounds on the two probabilities on the right-hand side.

For any choice of a good neighbour u of v and $9/\varepsilon$ common neighbours of u and v, the probability that both u and the common neighbours belong to the same part as v is $\left(1/\Delta^{1-\varepsilon}\right)^{9/\varepsilon+1}$. Each event in $C_v^{(1)}$ corresponds to at least one such choice of u and the common neighbours of u and v, so a union bound gives

$$\Pr\left[C_v^{(1)}\right] \leqslant \Delta \cdot \binom{\Delta^{1-2\varepsilon}}{\frac{9}{\varepsilon}} \cdot \left(\frac{1}{\Delta^{1-\varepsilon}}\right)^{\frac{9}{\varepsilon}+1} \tag{2.3}$$

because there are at most Δ choices for u as a good neighbour of v, and $|N_G(u) \cap N_G(v)| < \Delta^{1-2\varepsilon}$, so there are less than $\binom{\Delta^{1-2\varepsilon}}{9/\varepsilon}$ choices for the $9/\varepsilon$ common neighbours.

By bounding the binomial coefficient in (2.3), we get

$$\Pr\left[C_v^{(1)}\right] < \Delta \cdot \frac{\Delta^{(1-2\varepsilon)\frac{9}{\varepsilon}}}{\left(\frac{9}{\varepsilon}\right)!} \cdot \Delta^{(\varepsilon-1)(\frac{9}{\varepsilon}+1)} = \frac{\Delta^{-9+\varepsilon}}{\left(\frac{9}{\varepsilon}\right)!} < \frac{1}{2}\Delta^{-3},$$

where the last inequality holds for Δ large enough.

A similar argument works for $\Pr\left[C_v^{(2)}\right]$: Each event in $C_v^{(2)}$ corresponds to at least one choice of $9/\varepsilon$ disjoint edges spanned by the good neighbours of v and assigning them to the part v belongs to. The probability for getting the desired assignment under a random partition is $\left(1/\Delta^{1-\varepsilon}\right)^{18/\varepsilon}$ because a matching of $9/\varepsilon$ edges has a support of $18/\varepsilon$ vertices. Hence another union bound gives

$$\Pr\left[C_v^{(2)}\right] \leqslant \binom{\Delta^{2-4\varepsilon}}{\frac{9}{\varepsilon}} \cdot \left(\frac{1}{\Delta^{1-\varepsilon}}\right)^{\frac{18}{\varepsilon}}, \tag{2.4}$$

where the binomial coefficient bounds the number of possible matchings of size $9/\varepsilon$ in the graph induced by the good neighbours of v because by assumption, the number of edges in $N_G(v)$ is at most $\Delta^{2-4\varepsilon}$. Bounding the binomial coefficient in (2.4) gives

$$\Pr\left[C_v^{(2)}\right] < \frac{\Delta^{(2-4\varepsilon)\frac{9}{\varepsilon}}}{\left(\frac{9}{\varepsilon}\right)!} \cdot \Delta^{(\varepsilon-1)\frac{18}{\varepsilon}} = \frac{\Delta^{-18}}{\left(\frac{9}{\varepsilon}\right)!} < \frac{1}{2}\Delta^{-3},$$

where the last inequality again holds for Δ large enough.

Altogether, we get

$$\Pr[C_v] \leqslant \Pr\left[C_v^{(1)}\right] + \Pr\left[C_v^{(2)}\right] < \frac{1}{\Delta^3}. \tag{2.5}$$

The inequalities in (2.1), (2.2) and (2.5) show that all events in $\{A_v, B_v, C_v \mid v \in V\}$ have probabilities smaller than Δ^{-3}, so in order to apply Lemma 1.1, we may choose $p = \Delta^{-3}$. The condition of Lemma 1.1 is $ep(d+1) \leqslant 1$, which translates to

$$e \cdot \frac{1}{\Delta^3} \cdot (3\Delta^2 + 3\Delta + 1) \leqslant 1$$

in our setting. It is satisfied if Δ is large enough, hence the symmetric Lovász Local Lemma proves the existence of a partition of the vertices of G avoiding all the events $\{A_v, B_v, C_v \mid v \in V\}$. In other words, there exists a partition $V = U_1 \dot\cup \ldots \dot\cup U_\ell$ with $\ell = \Delta^{1-\varepsilon}$ such that we have the following:

(i) For all $i \in [\ell]$, $G[U_i]$ has maximum degree at most $2d^\varepsilon$.

(ii) For every $i \in [\ell]$ and $v \in U_i$, there exists a set $S_{v,i} \subseteq N_{G[U_i]}(v)$ of at most $110/\varepsilon^2$ vertices such that the vertices in $N_{G[U_i]}(v) \setminus S_{v,i}$ span no edges.

The first point is simply a reformulation of avoiding all events A_v. For the second, fix $i \in [l]$ and a vertex $v \in U_i$. Avoiding the event B_v means that v has at most $10/\varepsilon$ bad neighbours in U_i while avoiding C_v means that the set of good neighbours of v spans at most $100/\varepsilon^2$ edges. So if we let $S_{v,i}$ contain all bad neighbours of v and one vertex of each edge spanned by the good neighbours of v, then $|S_{v,i}| \leqslant 10/\varepsilon + 100/\varepsilon^2 < 110/\varepsilon^2$ and $N_G(v) \setminus S_{v,i}$ spans no edges, as desired.

Now fix a part U_i of the partition. To colour U_i, we construct an auxiliary digraph $D_i = (U_i, A)$ with arcs given by $A \coloneqq \{(v,u) \mid v \in U_i, u \in S_{v,i}\}$. Item (ii) above guarantees that for each v, $|S_{v,i}| \leqslant 110/\varepsilon^2$, so the outdegree of every vertex in the digraph D_i is at most $110/\varepsilon^2$.

We claim that the digraph D_i is $(220/\varepsilon^2)$-degenerate in total degrees: Indeed, if there was an induced subgraph H with minimum total degree larger than $220/\varepsilon^2$, counting edges in H would give

$$|E(H)| = \frac{1}{2}\sum_{v\in U_i} \deg(v) > \sum_{v\in U_i} \frac{110}{\varepsilon^2} \geqslant \sum_{v\in U_i} \deg^+(v) = |E(H)|,$$

a contradiction. It is well-known that d-degenerate graphs can be properly coloured using $d+1$ colours, so by the above, D_i is $(220/\varepsilon^2 + 1)$-colourable.

Fix such a proper colouring of the vertices in U_i with $220/\varepsilon^2 + 1$ colours and note that using this colouring, there cannot be any monochromatic triangle in $G[U_i]$: If $v \in U_i$ and $u_1, u_2 \in N_{G[U_i]}(v)$ span a triangle, then one of the u_i belongs to $S_{v,i}$ by item (ii) above and hence has not the same colour as v, so the triangle is not monochromatic.

Altogether, we can colour every one of the $\Delta^{1-\varepsilon}$ parts from its own colour palette with at most $220/\varepsilon^2 + 1$ colours, i. e. using $\left(\frac{220}{\varepsilon^2} + 1\right)\Delta^{1-\varepsilon} = \Theta\left(\frac{\Delta^{1-\varepsilon}}{\varepsilon^2}\right)$ colours in total, such that each colour class is triangle-free. Moreover, item (i) guarantees that the maximum degree inside each colour class is at most $2\Delta^\varepsilon = \mathcal{O}(\Delta^\varepsilon)$. So for the conclusion of the lemma, we may choose the sets V_i to be the constructed colour classes. □

With this lemma at hand, the conclusion of Theorem 2.2 is almost immediate:

Proof of Theorem 2.2. Lemma 2.4 partitions the vertices of the graph G into $k = \Theta(\Delta^{1-\varepsilon}/\varepsilon^2)$ parts $(V_i)_{i\in[k]}$, inducing triangle-free subgraphs of maximum degree $\mathcal{O}(\Delta^\varepsilon)$. Applying Theorem 2.3, we see that for each $i \in [k]$, we have

$$\chi(G[V_i]) = \mathcal{O}\left(\frac{2\Delta^\varepsilon}{\log(2\Delta^\varepsilon)}\right),$$

so colouring each set V_i with its own colours, we get the desired asymptotic bound

$$\chi(G) = \Theta\left(\frac{\Delta^{1-\varepsilon}}{\varepsilon^2}\right) \cdot \mathcal{O}\left(\frac{2\Delta^\varepsilon}{\log(2\Delta^\varepsilon)}\right) = \mathcal{O}\left(\frac{\Delta}{\varepsilon^3 \log \Delta}\right). \qquad \square$$

2.3 The second Case: $f < \Delta^{4\varepsilon}$

In the case of $f < \Delta^{4\varepsilon}$, we do not have the number of edges in a neighbourhood bounded away from $\Theta(\Delta^2)$ as before, which makes an argument as in Lemma 2.4 impossible. In particular, the bounds on $\Pr[A_v]$, $\Pr[B_v]$ and $\Pr[C_v]$ rely heavily on the gap introduced by the assumption $f \geqslant \Delta^{4\varepsilon}$.

Nonetheless, the ideas of the first case are useful for the second one in the sense that we also aim for a splitting into vertex-disjoint induced subgraphs that are easier to colour. In fact, the following lemma states that a reduction to the first case is possible.

Lemma 2.5. *Let $\varepsilon \in (0, \frac{1}{6})$ and let $G = (V, E)$ be a graph with maximum degree Δ in which the neighbourhood $N_G(v)$ of any vertex $v \in V$ spans at most Δ^2/f edges for some $f < \Delta^{4\varepsilon}$. There exists an absolute constant f_0 such that the following holds whenever $f \geqslant f_0$: Let j be the smallest positive integer such that*

$$f \geqslant \left(\frac{8\Delta}{2^j}\right)^{\varepsilon}.$$

Then G can be split into at most 2^j vertex-disjoint induced subgraphs such that in each of them, either the maximum degree is at most one, or the maximum degree is at most $\frac{8\Delta}{2^j}$ and the neighbourhood of each vertex spans at most $\frac{8\Delta}{2^j}/f$ vertices.

Indeed, one can observe that Theorem 2.2 applies to each of the subgraphs provided by the above lemma. In Theorem 2.7, we will see that colouring each of the subgraphs from their own colour palette completes the proof of the second case.

In order to prove Lemma 2.5, we iteratively split the graph into smaller pieces, at each step coming closer to the desired bounds on the degrees and the number of edges in neighbourhoods. Lemma 2.6 provides the details for a single step.

Lemma 2.6. *Let $G = (V, E)$ be a graph on n vertices with maximum degree $d \geqslant 2$ in which the neighbourhood $N_G(v)$ of any vertex $v \in V$ spans at most s edges. Then there exists a partition $V = V_1 \dot\cup V_2$ such that for $i \in \{1, 2\}$, the induced subgraph $G[V_i]$ has maximum degree at most d' and the neighbourhood $N_{G[V_i]}(v)$ of each vertex $v \in V_i$ spans at most s' edges, where*

$$d' = \frac{d}{2} + 2\sqrt{d \log d} \quad \text{and} \quad s' = \frac{s}{4} + 4d^{3/2}\sqrt{\log d}.$$

Proof. We prove that a random partition of V into two sets V_1, V_2 has the desired properties with positive probability by an application of the symmetric Lovász Local Lemma (Lemma 1.1). To generate a random partition, assign each vertex independently and with uniform probability of $\frac{1}{2}$ to one of the two sets.

For each vertex $v \in V$, define the following two events:

A_v: The degree of v inside its subgraph $G[V_i]$ is larger than d'.

B_v: The neighbourhood $N_{G[V_i]}(v)$ spans more than s' edges.

The lemma claims the existence of a partition of the graph such that all events in $\{A_v, B_v \mid v \in V\}$ are avoided, which is exactly what the Lovász Local Lemma provides, so we check its conditions.

Note that for each $v \in V$, both A_v and B_v depend only on events corresponding to vertices of distance one or two from v. There are at most $d^2 + d$ such vertices, hence each event is mutually independent from all but at most $2(d^2 + d)$ others and we may use $2d^2 + 2d$ as an upper bound for the dependencies in Lemma 1.1.

Moreover, we need to bound the probabilities of the events A_v and B_v:

For $\Pr[A_v]$, note that by construction of the random partition, the random variable X counting neighbours of $v \in V_i$ in $G[V_i]$ is binomially distributed with parameters $\deg_G(v)$ and $\frac{1}{2}$.

By assumption, $\deg_G(v) \leqslant d$ and hence $\mathbb{E}[X] = \deg_G(v) \cdot \frac{1}{2} \leqslant \frac{d}{2}$. Applying the Chernoff bound (ii) in Theorem A.1 with $a = 2\sqrt{d \log d}$ and $n = d$, we get

$$\Pr[A_v] = \Pr\bigl[X > d'\bigr] \leqslant e^{-8 \log d} = \frac{1}{d^8}. \tag{2.6}$$

To bound $\Pr[B_v]$, we use Azuma's inequality, which is derived in Appendix A.2, applied to vertex exposure martingales. Thereto, let i be such that $v \in V_i$ and let f_v be the graph theoretic function counting the number of edges in $N_{G[V_i]}(v)$. Enumerate the vertices in G such that all neighbours of v are among the the first d vertices and let $X_0, \dots, X_n$ be the vertex exposure martingale for f_v using this ordering.

Note that $X_0 = \mathbb{E}[f_v|\mathcal{G}_0] = \mathbb{E}[f_v] \leqslant \frac{s}{4}$ and $X_d = \mathbb{E}[f_v|\mathcal{G}_d] = f(G[V_i])$ because all neighbours of v are among the first d vertices that are exposed. Moreover, we have $|X_{j+1} - X_j| \leqslant d$ for all $j \in [d]$ because all vertices have maximum degree d and hence exposing a new vertex reveals at most d new edges that are present for sure, so the conditional expectation changes by at most d.

Define $X'_0 = 0$, and $X'_j \coloneqq \frac{1}{d}(X_j - X_0)$ for all $j \in [d]$, so that $|X'_{j+1} - X'_j| \leqslant 1$ for all $j \in [d]$. This means that the sequence $(X'_j)_{j=0,\dots,d}$ satisfies the assumption of Azuma's inequality (Theorem A.4) and we can conclude

$$\begin{aligned}\Pr[B_v] = \Pr\bigl[f_v(G[V_i]) > s'\bigr] &= \Pr\left[X_d > \frac{s}{4} + 4d^{3/2}\sqrt{\log d}\right] \leqslant \\ &\leqslant \Pr\left[\frac{1}{d}(X_d - X_0) > 4\sqrt{d \log d}\right] = \Pr\left[X'_d > \lambda\sqrt{d}\right] \leqslant \\ &\leqslant e^{-\frac{\lambda^2}{2}} = e^{-\frac{(4\sqrt{\log d})^2}{2}} = e^{-8 \log d} = \frac{1}{d^8},\end{aligned} \tag{2.7}$$

where we applied Theorem A.4 with $\lambda = 4\sqrt{\log d}$.

So in the application of Lemma 1.1, we can set $p = d^{-8}$ as an upper bound for the probabilities, justified by (2.6) and (2.7). The condition $ep(d+1) \leqslant 1$ then translates to

$$e \cdot \frac{1}{d^8} \cdot \left(2d^2 + 2d + 1\right) \leqslant 1.$$

Note that the left-hand side is decreasing in d and the inequality holds for $d = 2$, so it holds for all $d \geqslant 2$ and hence Lemma 1.1 applies, giving the conclusion. $\square$

Based on Lemma 2.6, we can now prove Lemma 2.5.

Proof of Lemma 2.5. Let $d_0 \coloneqq \Delta$ and $s_0 \coloneqq \Delta^2/f$ and define sequences $(d_i)_{i=0}^j$ and $(f_i)_{i=0}^j$ by

$$d_i = \frac{d_{i-1}}{2} + 2\sqrt{d_{i-1}\log d_{i-1}} \quad \text{and} \quad s_i = \frac{s_{i-1}}{4} + 4d_{i-1}^{3/2}\sqrt{\log d_{i-1}}$$

for all $t \in [j]$. If $d_0 = \Delta \geqslant 2$, applying Lemma 2.6 with $d = d_0$ and $s = s_0$ to G, we can split G into two vertex-disjoint subgraphs with maximum degree at most d_1 and such that each neighbourhood in these two subgraphs spans at most s_1 edges.

Another application of Lemma 2.6 to each of the two subgraphs (provided that they have maximum degree at least 2) results in four (or less) subgraphs of maximum degree at most d_2 in which every neighbourhood spans at most s_2 edges.

Iterating this procedure j times, at each step applying Lemma 2.6 to all subgraphs with maximum degree at least 2, we in the end get at most 2^j subgraphs. Each of them either had maximum degree at most one at some step, or it was created in the j^{th} iteration. Note that in the last case, d_j is a bound on the maximum degree and s_j is a bound on the number of edges spanned by neighbourhoods.

So in order to prove Lemma 2.5, it remains to check

$$d_j \leqslant \frac{8\Delta}{2^j} \quad \text{and} \quad s_j \leqslant \frac{8\Delta}{2^j f}. \tag{2.8}$$

By definition of $(d_i)_{i=0}^j$ and j, we have

$$d_{i-1} \geqslant d_0/2^{i-1} > \Delta/2^{j-1} \geqslant f^{1/\varepsilon}/8 \geqslant f_0^{1/\varepsilon}/8 \tag{2.9}$$

for all $i \in [j]$, hence by choosing the constant f_0 large enough, we get that d_{i-1} is large enough to satisfy $2\sqrt{d_{i-1}\log d_{i-1}} \leqslant \frac{3}{2}d_{i-1}^{2/3}$, implying

$$d_i = \frac{d_{i-1}}{2} + 2\sqrt{d_{i-1}\log d_{i-1}} \leqslant \frac{d_{i-1}}{2} + \frac{3d_{i-1}^{2/3}}{2} \leqslant \frac{\left(\sqrt[3]{d_{i-1}} + 1\right)^3}{2},$$

which can – after taking cube roots and subtracting $\frac{1}{2^{1/3}-1}$ on both sides – be rewritten in the form

$$d_i^{1/3} - \frac{1}{2^{1/3}-1} \leqslant \frac{1}{2^{1/3}}\left(d_{i-1}^{1/3} - \frac{1}{2^{1/3}-1}\right).$$

Inductively applying this inequality gives

$$d_i^{1/3} - \frac{1}{2^{1/3}-1} < \frac{1}{2^{i/3}}\left(d_0^{1/3} - \frac{1}{2^{1/3}-1}\right) \tag{2.10}$$

for all $i \in [j]$. In particular, for $i = j$ we obtain

$$d_j^{1/3} \leqslant \frac{d_0^{1/3}}{2^{j/3}} + \frac{1}{2^{1/3}-1} \leqslant \frac{d_0^{1/3}}{2^{j/3}} + 4 \leqslant \frac{2d_0^{1/3}}{2^{j/3}},$$

where the last inequality holds if d_0 is large enough, which can again be guaranteed due to (2.9) by choosing f_0 large enough. Taking third powers and plugging in $d_0 = \Delta$ gives $d_j \leqslant \frac{8\Delta}{2^j}$, which is the first inequality in (2.8).

To derive the second one, note that from (2.10), we can conclude by the same arguments as above that for any $i \in [j]$, we have $d_i \leqslant 8d_0/2^i$. Consequently, the definition of s_i and j imply that we have

$$s_{i-1} \geqslant \frac{s_0}{4^{i-1}} = \frac{d_0^2}{4^{i-1}f} = \frac{\left(\frac{8d_0}{2^{i-1}}\right)^2}{64f} \geqslant \frac{1}{64}\left(\frac{8d_0}{2^{i-1}}\right)^{2-\varepsilon} \geqslant \frac{1}{64}d_{i-1}^{2-\varepsilon} \tag{2.11}$$

for all $i \in [j]$. Combining (2.9) and (2.11), we see that by choosing f_0 large enough, we can make all s_{i-1} larger than any constant that we wish.

We claim that $s_i \leqslant s_{i-1}/4 + 3s_{i-1}^{5/6}$ for all $i \in [j]$. To see this, note that (2.11) can be rewritten to $d_{i-1} \leqslant (64s_{i-1})^{1/(2-\varepsilon)}$, and plugging this into the definition of s_i gives

$$s_i = \frac{s_{i-1}}{4} + 4d_{i-1}^{3/2} \cdot \sqrt{\log d_{i-1}} \leqslant \frac{s_{i-1}}{4} + 4 \cdot (64s_{i-1})^{\frac{3}{2(2-\varepsilon)}} \cdot \sqrt{\log\left((64s_{i-1})^{\frac{1}{2-\varepsilon}}\right)},$$

so it suffices to prove

$$4\,(64s_{i-1})^{\frac{3}{2(2-\varepsilon)}} \cdot \sqrt{\log\left((64s_{i-1})^{\frac{1}{2-\varepsilon}}\right)} \leqslant 3s_{i-1}^{\frac{5}{6}}$$

$$\iff \quad \log\left(64s_{i-1}\right) \leqslant c \cdot s_{i-1}^{\frac{5}{6}-\frac{3}{4-2\varepsilon}}$$

with the positive constant $c = \frac{9}{16} \cdot 64^{\frac{-3}{2-\varepsilon}} \cdot (2-\varepsilon)$. By monotonicity of the exponent and the assumption $\varepsilon \in (0, \varepsilon_0)$ with $\varepsilon_0 \coloneqq \frac{1}{6}$, it is enough to have

$$\log\left(64s_{i-1}\right) \leqslant c \cdot s_{i-1}^{\frac{5}{6}-\frac{3}{4-2\varepsilon_0}} = c \cdot s_{i-1}^{\frac{1}{66}},$$

which can be guaranteed to be true by choosing f_0 large enough.

Using the claim, we get

$$s_i \leqslant s_{i-1}/4 + 3s_{i-1}^{5/6} \leqslant \frac{1}{4}\left(\sqrt[6]{s_{i-1}} + 2\right)^6$$

for all $i \in [j]$, which can – after taking sixth roots and subtracting $\frac{2}{4^{1/6}-1}$ on both sides – be rewritten in the form

$$s_i^{1/6} - \frac{2}{4^{1/6}-1} \leqslant \frac{1}{4^{1/6}}\left(s_{i-1}^{1/6} - \frac{2}{4^{1/6}-1}\right).$$

Iteratively applying this inequality gives

$$s_j^{1/6} - \frac{2}{4^{1/6-1}} \leqslant \frac{1}{4^{j/6}}\left(s_0^{1/6} - \frac{2}{4^{1/6}-1}\right),$$

and using $s_0 = d_0^2/f$ and $4^{1/6} - 1 > 1/4$, we get

$$s_j^{1/6} < \frac{s_0^{1/6}}{4^{j/6}} + \frac{2}{4^{1/6}-1} < \frac{d_0^{2/6}}{4^{j/6}f^{1/6}} + 8 \leqslant \frac{2d_0^{2/6}}{4^{j/6}f^{1/6}},$$

where the last step can be made true by choosing f_0 such that d_0 is large enough to satisfy the inequality. Taking sixth powers and plugging in $d_0 = \Delta$ gives $s_j \leqslant \frac{8\Delta}{2^j}/f$, which is the second inequality in (2.8). □

Having the graph split into components satisfying the assumptions of Theorem 2.2, the only thing left to do is to make sure that colouring each part from its own colour palette by that theorem does not require too many colours so that we stay in the desired asymptotics. This is covered by the following statement.

Theorem 2.7. *Let $\varepsilon \in (0, \frac{1}{6})$ and let $G = (V, E)$ be a graph with maximum degree Δ in which the neighbourhood $N_G(v)$ of any vertex $v \in V$ spans at most Δ^2/f edges for some $f < \Delta^{4\varepsilon}$. Then the chromatic number of G satisfies*

$$\chi(G) = \mathcal{O}\Big(\frac{\Delta}{\varepsilon^2 \log f}\Big).$$

Proof. We intend to apply Lemma 2.5 to the graph G. As we are aiming for a proof of the statement $\chi(G) = \mathcal{O}(\Delta/\log f)$ for $|V| \to \infty$, issues for small f can be hidden in the constant suppressed by the $\mathcal{O}$-notation and we may assume that f is larger than some fixed constant, in particular the constant f_0 provided by Lemma 2.5.

Consequently, we can apply Lemma 2.5 to get at most 2^j vertex-disjoint induced subgraphs such that each of them has maximum degree at most one or the maximum degree is at most $\frac{8\Delta}{2^j}$ and the neighbourhood of each vertex spans at most $\frac{8\Delta}{2^j}/f$ vertices.

We colour the subgraphs of maximum degree at most one using two colours, and as all other subgraphs satisfy the assumptions of Theorem 2.2 with Δ replaced by $\frac{8\Delta}{2^j}$ and ε replaced by $\varepsilon/4$, we know that each of them can be coloured using

$$\mathcal{O}\left(\frac{\frac{8\Delta}{2^j}}{\left(\frac{\varepsilon}{4}\right)^3 \cdot \log\left(\frac{8\Delta}{2^j}\right)}\right) = \mathcal{O}\Big(\frac{\Delta}{2^j \varepsilon^2 \log f}\Big)$$

colours, so colouring all subgraphs from their own colour palette gives the bound

$$\chi(G) \leqslant 2^j \cdot \mathcal{O}\Big(\frac{\Delta}{2^j \varepsilon^2 \log f}\Big) = \mathcal{O}\Big(\frac{\Delta}{\varepsilon^2 \log f}\Big),$$

which is the what we wanted to prove. □

2.4 Conclusion and Asymptotic Tightness

For completeness, we state the proof of the main theorem of this section, Theorem 2.1, which pieces the last two sections together.

Proof of Theorem 2.1. Let the graph G, its maximum degree Δ and the number f be as in the theorem statement. We can now fix some $\varepsilon \in (0, \frac{1}{6})$ and distinguish cases: If $f \geqslant \Delta^{4\varepsilon}$, we apply Theorem 2.2. This gives the bound

$$\chi(G) = \mathcal{O}\Big(\frac{\Delta}{\varepsilon^3 \log \Delta}\Big) \leqslant \mathcal{O}\left(\frac{\Delta}{\frac{1}{2}\varepsilon^3 \log f}\right) = \mathcal{O}\Big(\frac{\Delta}{\log f}\Big),$$

where we used the assumption $f \leqslant \Delta^2$ to get $\log \Delta \geqslant \frac{1}{2} \log f$ and absorbed the factor $\frac{1}{2}\varepsilon^3$ into the $\mathcal{O}$-notation because it is constant.

In the case $f < \Delta^{4\varepsilon}$, we can apply Theorem 2.7, which gives

$$\chi(G) = \mathcal{O}\Big(\frac{\Delta}{\varepsilon^2 \log f}\Big) = \mathcal{O}\Big(\frac{\Delta}{\log f}\Big),$$

where we again absorbed the constant factor ε^2 in the $\mathcal{O}(.)$-notation. This concludes the proof of Theorem 2.1. □

Last but not least, it is of general interest to know how good the bound $\mathcal{O}(\Delta/\log f)$ on the chromatic number provided by Theorem 2.1 is. It turns out that the chromatic number is not just upper bounded by $c \cdot \Delta/\log f$, but that there are graphs with chromatic number lower bounded by $b \cdot \Delta/\log f$ for some constant b.

Theorem 2.8. *There exists an absolute positive constant b such that the following holds: For every positive integer Δ and every real number f satisfying $2 \leqslant f \leqslant \Delta^2$, there is a graph $G = (V, E)$ with maximum degree at most Δ in which the neighbourhood $N_G(v)$ of every vertex $v \in V$ spans at most Δ^2/f edges and*

$$\chi(G) \geqslant b \cdot \frac{\Delta}{\log f}.$$

As an auxiliary statement, we first prove the following lemma on existence of graphs satisfying certain bounds for maximum degree, girth and independence number.

Lemma 2.9. *There exists a constant C such that for every integer $\Delta \geqslant 2$ there exists a triangle-free graph G on n vertices with maximum degree Δ and no independent set of size at least $Cn \log \Delta/\Delta$.*

Proof. Note that if we choose $C \geqslant 7/\log 7$, then for $\Delta \leqslant 7$, the statement of the above lemma is that there exists a non-empty triangle-free graph of maximum degree at least Δ, which is trivially true. So we assume that $\Delta \geqslant 8$ for the rest of the proof.

Let $n_0 := 2\Delta^3$ and $p := \frac{\Delta}{2n_0}$. We prove that a random graph $G_0 \in \mathcal{G}(n_0, p)$ is already close to having the desired properties in the sense that

$$\Pr\Big[\#\{\text{triangles in } G_0\} \geqslant \frac{n_0}{4}\Big] \leqslant \frac{1}{4}, \tag{2.12}$$

$$\Pr\Big[\#\{\text{vertices of degree larger than } \Delta \text{ in } G_0\} \geqslant \frac{n_0}{4}\Big] \leqslant \frac{1}{4}, \tag{2.13}$$

$$\text{and} \quad \Pr\Big[\#\Big\{\text{indep. sets of size at least } \frac{Cn_0 \log \Delta}{2\Delta} \text{ in } G_0\Big\} \geqslant 1\Big] \leqslant \frac{1}{4}. \tag{2.14}$$

For (2.12), let X be the random variable counting triangles in G_0 and note that linearity of expectation gives

$$\mathbb{E}[X] \leqslant \binom{n_0}{3} \cdot p^3 < n_0^3 \cdot p^3 = \frac{\Delta^3}{8} = \frac{n_0}{16},$$

so by Markov's inequality,

$$\Pr\Big[X \geqslant \frac{n_0}{4}\Big] \leqslant \frac{\mathbb{E}[X]}{\frac{n_0}{4}} \leqslant \frac{1}{4},$$

which is (2.12). For (2.13), let Y be the random variable counting vertices of degree larger than Δ in G_0. Note that the degree of a fixed vertex v in G_0 is binomially distributed with parameters $n_0 - 1$ and p, so $\mathbb{E}[Y] = (n_0 - 1)p < n_0 p = \frac{\Delta}{2}$. Hence linearity of expectation and the Chernoff estimate (i) in Theorem A.1 with $a = \frac{\Delta}{2}$ give

$$\mathbb{E}[Y] \leqslant n_0 \cdot \Pr[\deg v > \Delta] \leqslant n_0 \cdot e^{\Delta(1-2\log 2)} < \frac{n_0}{16},$$

where the last inequality holds for $\Delta \geqslant 8$. As above, we can now deduce (2.13) from Markov's inequality:

$$\Pr\left[Y \geqslant \frac{n_0}{4}\right] \leqslant \frac{\mathbb{E}[X]}{\frac{n_0}{4}} \leqslant \frac{1}{4}.$$

For (2.14), let the random variable Z count the number of independent sets of size at least $a := \frac{Cn_0 \log \Delta}{2\Delta}$. Linearity of expectation gives

$$\mathbb{E}[Z] \leqslant \binom{n_0}{a} \cdot (1-p)^{\binom{a}{2}} \leqslant \left(\frac{n_0 e}{a}\right)^a \cdot e^{-p\frac{a^2}{2}} = e^{a \cdot \left(\log n_0 + 1 - \log a - \frac{a\Delta}{2n_0}\right)}.$$

We want to have $\mathbb{E}[Z] \leqslant \frac{1}{4}$ so that Markov's inequality yields $\Pr[Z \geqslant 1] \leqslant \frac{1}{4}$, which is the desired inequality (2.14). Equivalently, we need

$$\begin{aligned} & a \cdot \left(\log n_0 + 1 - \log a - \frac{a\Delta}{2n_0}\right) \leqslant \log \frac{1}{4} \\ \iff \quad & \left(\frac{C}{4} - 1\right) \cdot \log \Delta + \log \frac{C}{2} + \log\log \Delta - 1 \geqslant \frac{2\Delta \log 4}{Cn_0 \log \Delta}, \end{aligned}$$

where we divided by $-a$ and plugged in the value of a. Note that we have $\log \frac{C}{2} + \log\log \Delta - 1 \geqslant 0$, so it suffices to prove

$$\left(\frac{C}{4} - 1\right) \cdot \log \Delta \geqslant \frac{2\Delta \log 4}{Cn_0 \log \Delta} \quad \iff \quad C(C-4) \geqslant \frac{8 \log 4}{(\Delta \log \Delta)^2},$$

which holds for all $\Delta \geqslant 2$ if C is chosen large enough (in fact $C = 6$, which is also larger than $7/\log 7$, is already sufficient).

By (2.12), (2.13), (2.14) and a union bound, with a probability of at least $\frac{1}{4}$, the random graph G_0 has less than $\frac{n_0}{4}$ triangles, less than $\frac{n_0}{4}$ vertices of degree larger than Δ and no independent sets of size at least a.

By deleting $\frac{n_0}{2}$ vertices from G_0, we can hence make sure to get a graph G on $n = \frac{n_0}{2}$ vertices with no triangles and no vertices of degree larger than Δ. Moreover, as deleting vertices does not increase the size of any independent set, G has no independent set of size at least $a = \frac{Cn \log \Delta}{\Delta}$. □

Proof of Theorem 2.8. Fix Δ and f as in Theorem 2.8. Assume first that $f > \frac{\Delta - 2}{4}$ and let G be a triangle-free graph on n vertices of maximum degree Δ and with no independent set of size at least $a := \frac{Cn \log \Delta}{\Delta}$ as provided by Lemma 2.9. Then, as desired,

$$\chi(G) \geqslant \frac{n}{a} \geqslant \frac{\Delta}{C \log \Delta} = \frac{\Delta}{\log(4f) + 2} = \Omega\left(\frac{\Delta}{\log f}\right).$$

If $f \leqslant \frac{\Delta-2}{4}$, then let H be a triangle-free graph with maximum degree $2f$ on n vertices with no independent set of size $a := \frac{Cn\log(2f)}{2f}$ as provided by Lemma 2.9. Let G be the graph that we obtain from H by replacing every vertex by a clique of $\left\lfloor \frac{\Delta}{2f+1} \right\rfloor$ vertices. Then the maximum degree of G is at most Δ because $\frac{\Delta}{2f+1}+2f \leqslant \Delta$. Moreover, the maximum number of edges in a neighbourhood is at most

$$2f \cdot \binom{\frac{\Delta}{2f+1}}{2} \leqslant 2f \cdot \frac{\Delta^2}{2(2f)^2} < \frac{\Delta^2}{f}$$

and when constructing G from H, the size of the largest independent set remained at most a, so

$$\chi(G) \geqslant \frac{|E(G)|}{a} = \frac{n \cdot \left\lfloor \frac{\Delta}{2f+1} \right\rfloor}{\frac{Cn\log(2f)}{2f}} = \Omega\left(\frac{\Delta}{\log f}\right). \qquad \square$$

Chapter 3

Applications of the Lopsided Lemma

When we introduced the lopsided Local Lemma in Section 1.2, we already remarked that in order to exploit it, we need to find non-trivial negative dependency graphs.

This chapter is devoted to the investigation of two similar settings, both developed by Lu and Székely in [11, 12], where we are able to describe negative dependency graphs. More concretely, we work in the probability space of uniformly chosen random matchings in complete and complete bipartite graphs. The key is to restrict to a family of so-called canonical events in these spaces that interact naturally, providing the structure that we need.

Interestingly, the configuration circumscribed above provides negative dependency graphs that are – at least under certain circumstances – at the same time ε-near positive dependency graphs for some $\varepsilon > 0$. So besides lower bounding the probabilities in question by the lopsided Local Lemma, we can also upper bound them using the tools developed in Section 1.3. Under certain conditions, these bounds match asymptotically, which gives a method for asymptotically counting the number of objects in the probability space avoiding the events in question.

The results shown in this chapter are all due to Lu and Székely, and the presentation given here arises from their papers [11] and [12]. Section 3.2 develops negative dependency graphs for the case of matchings in complete bipartite graphs and shows an application to hypergraph packings, while Section 3.3 covers matchings in complete graphs, using a somewhat different approach that can be modified to apply to the first case as well. This chapter culminates in the presentation of the technique of asymptotic counting (Subsection 3.3.3) and its application to the asymptotic enumeration of d-regular graphs in Subsection 3.3.4.

3.1 The Space of Random Matchings

To start with, we repeat and introduce some terminology related to matchings.

Definition 3.1. *Let $G = (V, E)$ be a graph.*

(i) A matching *in G is a set $M \subseteq E$ such that no two different edges in M share a vertex.*

(ii) *The* support *of a matching M in G is the set* $\operatorname{supp} M \coloneqq \bigcup_{e \in E} e$*.*

(iii) *A matching M in G is a* maximum cardinality matching *if*

$$|\operatorname{supp} M| = \max\{|\operatorname{supp} M'| \mid M' \text{ is a matching in } G\}.$$

(iv) *A matching M in G is a* perfect matching *if* $\operatorname{supp} M = V$*.*

(v) *Two matchings M_1 and M_2 in G are* conflicting *if $M_1 \cup M_2$ is not a matching.*

The subsequent definition introduces the space of random matchings and defines the crucial *canonical events*, among which we will be able to identify a family of negative dependency graphs if G is a complete or complete bipartite graph.

Definition 3.2. *Let G be a graph.*

(i) *The probability space Ω^G is the set of all maximum cardinality matchings in G equipped with a uniform probability measure.*

(ii) *For a matching M in G, we define the* canonical event A_M^G *in Ω^G associated to M by*

$$A_M^G \coloneqq \left\{ M' \in \Omega^G \,\middle|\, M \subseteq M' \right\}.$$

(iii) *Two canonical events $A_{M_1}^G$ and $A_{M_2}^G$ are conflicting if the matchings M_1 and M_2 are conflicting.*

If the underlying graph G is clear from the context, we may denote the canonical event associated to M by A_M only. Note that A_M is precisely the set of all maximum cardinality matchings in G extending M.

The following proposition indicates for a first time that specializing the graph G to a complete or complete bipartite graph provides enormously helpful regularity structures.

Proposition 3.3 ([12]). *Let $G = (V, E)$ be a complete or complete bipartite graph and let M_1 and M_2 be two matchings in G.*

(i) *The matchings M_1 and M_2 are conflicting if and only if $A_{M_1} \cap A_{M_2} = \emptyset$.*

(ii) *If M_1 and M_2 are not conflicting, then*

$$\overline{A_{M_2 \setminus M_1}} \subseteq \overline{A_{M_2}} \quad \text{and} \quad \overline{A_{M_2}} \cap A_{M_1} = \overline{A_{M_2 \setminus M_1}} \cap A_{M_1}.$$

Proof. (i) If M_1 and M_2 are conflicting, then there are two edges $e \in M_1$ and $f \in M_2$ and a vertex $v \in V$ such that $e \cap f = \{v\}$. Assume for contradiction that there is a matching M extending both M_1 and M_2. In particular, M needs to contain e and f – but then, v is covered by two edges of M, so M cannot be a matching, contradicting the assumption. Hence $A_{M_1} \cap A_{M_2} = \emptyset$.

If on the other hand M_1 and M_2 are not conflicting, then $M \coloneqq M_1 \cup M_2$ is a matching. We claim that this partial matching can be extended to a maximum cardinality matching, which proves that $A_{M_1} \cap A_{M_2} \neq \emptyset$.

The claim is obvious if G is a complete graph: As all edges are present, simply pairing the remaining vertices in $V \setminus \mathrm{supp}(M)$ extends M to a maximum cardinality matching because eventually, there is only at most one unpaired vertex left.

If G is bipartite, the same procedure extends M to a maximum cardinality matching if we always pair vertices from different sets of the bipartition. Here, one of the parts will eventually not contain any unpaired vertices any longer, so the extension is of maximum cardinality as well.

(ii) To prove $\overline{A_{M_2 \setminus M_1}} \subseteq \overline{A_{M_2}}$, take complements go get the equivalent inclusion $A_{M_2} \subseteq A_{M_2 \setminus M_1}$, which is obvious: Any matching extending M_2 also extends $M_2 \setminus M_1$.

To see $\overline{A_{M_2}} \cap A_{M_1} = \overline{A_{M_2 \setminus M_1}} \cap A_{M_1}$, note that the inclusion "$\supseteq$" is obviously true. For the other direction, take $M \in \overline{A_{M_2}} \cap A_{M_1}$, i. e. a matching conflicting M_2 and extending M_1. As M_1 and M_2 are not conflicting, their edges are either equal or disjoint, so any conflict between M and M_2 must come from a conflict with $M_2 \setminus M_1$. This proves that $M \in \overline{A_{M_2 \setminus M_2}}$ and hence establishes the inclusion "$\subseteq$" as well. □

The crucial graphs associated to a set of events in Ω^G the are the so-called *conflict graphs*.

Definition 3.4 ([13]). *Let G be a graph and let $A_1, \dots, A_m$ be canonical events in Ω^G. The* conflict graph *of $A_1, \dots, A_m$ is the graph $H = (V, E)$ defined by*

$$V = [m] \quad \textit{and} \quad E = \{\{i, j\} \mid A_i \textit{ and } A_j \textit{ are conflicting}\}.$$

Doing a slight abuse of notation, we sometimes use the conflict graph with the events themselves as the vertex set and edges defined similarly.

The following two chapters prove that conflict graphs are negative dependency graphs if the underlying graph G is either a complete or a complete bipartite graph. Under certain restrictions, we will also see that the conflict graph is an ε-near positive dependency graph.

3.2 Random Matchings in K_{n_1,n_2}

For the first family of negative dependency graphs, we explore matchings in a complete bipartite graph $G = K_{n_1,n_2}$. For simplicity, we write Ω^{n_1,n_2} for the probability space of maximum cardinality matchings in K_{n_1,n_2}.

It turns out that in this special case, we can profit from not only keeping the record of the edges of a matching M but also including the information about the bipartition: Every edge contains exactly one vertex from each set of the bipartition.

Definition 3.5. *Let $G = (V, E)$ be a bipartite graph with parts V_1 and V_2. An f-*representation *of a matching M in G is a triple (S, T, f) with $S \subseteq V_1$, $T \subseteq V_2$ and a bijection $f : S \to T$ such that $M = \{(s, f(s)) \mid s \in S\}$.*

It is easy to see that matchings and f-representations are in one-to-one correspondence once V_1 and V_2 are fixed, justifying that in the setting of a bipartite graph G, we will identify a matching M with the corresponding f-representation (S, T, f) and hence also A_M with $A_{(S,T,f)}$.

Let V_1 and V_2 be the parts of the bipartition of K_{n_1,n_2} of sizes n_1 and n_2, respectively. Assuming $n_1 \leqslant n_2$, all maximum cardinality matchings have size n_1, so their f-representations are of the form $(V_1, f(V_1), f)$ for an injection $f : V_1 \to V_2$. More generally, f-representations allow the interpretation of partial matchings as partial injections from V_1 to V_2.

Every permutation σ of V_2 acts naturally on matchings in K_{n_1,n_2} via the map π_σ defined as follows: For any matching M with f-representation (S, T, f), $\pi_\sigma(M)$ is the matching M' corresponding to (S', T', f') with

$$S' = S, \quad T' = \sigma(T) \quad \text{and} \quad f' = \sigma \circ f.$$

Extending π_σ to a set function by element-wise application, we get $\pi_\sigma(A_M) = A_{M'}$. One important fact is that because we use a uniform probability measure, permutations of the form π_σ are measure-preserving in Ω^G.

3.2.1 A Negative Dependency Graph for Canonical Events

We now prove that for a set of canonical events, the natural relation of conflicting or not, which is captured by the conflict graph, gives a negative dependency graph.

Theorem 3.6 ([11]). *Let n_1, n_2 be positive integers and let $A_1, \ldots, A_n$ be canonical events in Ω^{n_1,n_2}. Then the conflict graph $G = (V, E)$ for the events $A_1, \ldots, A_n$ is a negative dependency graph.*

Before turning to the theorem, we prove a lemma.

Lemma 3.7. *Consider canonical events $A_1, \ldots, A_n$ in Ω^{n_1,n_2} generated by matchings $(S_k, T_k, f_k)_{k\in[n]}$ in K_{n_1,n_2}. Fix $j \in [n]$, a set T and a bijection f such that (S_j, T, f) is another matching. Then $\Pr[A_j] = \Pr\big[A_{(S_j,T,f)}\big]$. Moreover, if G is the conflict graph for the events $A_1, \ldots, A_n$ and $I \subseteq [n] \setminus N_G(j)$, then*

$$\Pr\left[\bigcap_{i\in I} \overline{A_i} \cap A_{(S_j,T,f)}\right] \geqslant \Pr\left[\bigcap_{i\in I} \overline{A_i} \cap A_j\right]. \tag{3.1}$$

Proof. As before, let V_1 and V_2 denote the sets of the bipartition of the underlying graph K_{n_1,n_2} such that for all $k \in [n]$, $S_k \subseteq V_1$ and $T_k, T \subseteq V_2$. For the first statement, $\Pr[A_j] = \Pr\big[A_{(S_j,T,f)}\big]$, our goal is to construct a permutation σ of V_2 such that $\pi_\sigma\big(A_{(S_j,T,f)}\big) = A_j$. As π_σ is measure-preserving in Ω^{n_1,n_2}, this proves the equality.

We claim that we get the desired by defining $\sigma : V_2 \to V_2$, such that

$$\sigma|_{V_2\setminus(T\cup T_j)} = \mathrm{id}, \quad \sigma|_T = f_j \circ f^{-1} \quad \text{and} \quad \sigma|_{T_j\setminus T} = \sigma_0, \tag{3.2}$$

where $\sigma_0 : T_j \setminus T \to T \setminus T_j$ is an arbitrary bijection. Such σ_0 exists for sure because $|T_j \setminus T| = |T| - |T_j \cap T| = |T_j| - |T_j \cap T| = |T \setminus T_j|$, but in general, it is not unique.

Fixing some σ_0 of the described form, it can be directly seen that σ is well-defined by (3.2) and bijective. In particular, $\sigma|_T = f_j \circ f^{-1}$ gives

$$\sigma\big((S_j, T, f)\big) = \big(S_j, \sigma(T), \sigma \circ f\big) = (S_j, T_j, f_j), \tag{3.3}$$

as desired, also see Figure 3.1.

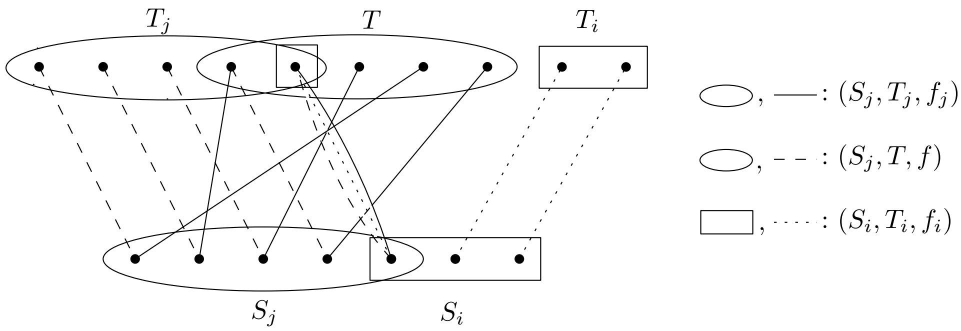

Figure 3.1: Matchings (S_j, T_j, f_j), (S_j, T, f) and (S_i, T_i, f_i) with $i \in I_0$.

This approach works for the second part as well. To start with, partition the set I such that $I_1 := \{i \mid A_i \text{ and } A_{(S_j,T,f)} \text{ are conflicting}\}$ and $I_0 := I \setminus I_1$. Note that by definition of G and because $I \subseteq [n] \setminus N_G(j)$, A_j and A_i are not conflicting for any $i \in I$.

By (i) in Proposition 3.3, we get that for all $i \in I_1$, $A_{(S_j,T,f)} \subseteq \overline{A_i}$, hence

$$\bigcap_{i \in I} \overline{A_i} \cap A_{(S_j,T,f)} = \bigcap_{i \in I_0} \overline{A_i} \cap A_{(S_j,T,f)}. \tag{3.4}$$

Applying π_σ with σ as in (3.2) to the right-hand side of (3.4), we on the one hand see that $\sigma\big((S_j, T, f)\big) = (S_j, T_j, f_j)$ as in (3.3). On the other hand, we claim that $\sigma(A_i) = A_i$ for all $i \in I_0$: For $t \in T_i \cap T$, the fact that $A_{(S_j,T,f)}$ and A_i as well as A_i and A_j are not conflicting gives

$$f^{-1}(t) = f_i^{-1}(t) = f_j^{-1}(t) \quad \Longrightarrow \quad \sigma(t) = f_j \circ f^{-1}(t) = t.$$

Secondly, the set $(T_i \cap T_j) \setminus T$ is empty: If $t \in (T_i \cap T_j) \setminus T$, then because A_j and A_i are not conflicting, there exists $s \in S_j \cap S_i$ such that $f_j(s) = t = f_i(s)$. In particular, $s \in S_j$, so because A_i and $A_{(S_j,T,f)}$ are not conflicting, we have $t = f_i(s) = f(s) \in T$, a contradiction.

Finally, $T_i \setminus (T \cup T_j) \subseteq V_2 \setminus (T \cup T_j)$, hence $\sigma(t) = t$ for all $t \in T_i \setminus (T \cup T_j)$ by definition. Altogether, we see that

$$\pi_\sigma \left(\bigcap_{i \in I_0} \overline{A_i} \cap A_{(S_j,T,f)} \right) = \bigcap_{i \in I_0} \overline{A_i} \cap A_{(S_j,T_j,f_j)}. \tag{3.5}$$

Using that π_σ is measure-preserving, (3.4), (3.5) and monotonicity of the probability measure imply

$$\Pr\left[\bigcap_{i\in I}\overline{A_i}\cap A_{(S_j,T,f)}\right] = \Pr\left[\bigcap_{i\in I_0}\overline{A_i}\cap A_{(S_j,T,f)}\right] = \Pr\left[\pi_\sigma\left(\bigcap_{i\in I_0}\overline{A_i}\cap A_{(S_j,T,f)}\right)\right] =$$
$$= \Pr\left[\bigcap_{i\in I_0}\overline{A_i}\cap A_{(S_j,T_j,f_j)}\right] \geqslant \Pr\left[\bigcap_{i\in I}\overline{A_i}\cap A_{(S_j,T_j,f_j)}\right],$$

which is (3.1) and hence proves the lemma. □

Now we are ready to prove Theorem 3.6.

Proof of Theorem 3.6. Let $A_1,\dots,A_n$ and G be as in the theorem statement. To see that G is a negative dependency graph, we need to prove that for all $j\in[n]$ and $I\subseteq N_G(j)$, we have

$$\Pr\left[A_j \,\middle|\, \bigcap_{i\in I}\overline{A_i}\right] \leqslant \Pr[A_j].$$

By definition of conditional probability, we have

$$\Pr\left[A_j \,\middle|\, \bigcap_{i\in I}\overline{A_i}\right] = \frac{\Pr\left[A_j\cap\bigcap_{i\in I}\overline{A_i}\right]}{\Pr\left[\bigcap_{i\in I}\overline{A_i}\right]} = \frac{\Pr\left[\bigcap_{i\in I}\overline{A_i}\,\middle|\, A_j\right]\cdot\Pr[A_j]}{\Pr\left[\bigcap_{i\in I}\overline{A_i}\right]},$$

so it suffices to prove

$$\Pr\left[\bigcap_{i\in I}\overline{A_i}\,\middle|\, A_j\right] \leqslant \Pr\left[\bigcap_{i\in I}\overline{A_i}\right].$$

Note that for fixed j, the family $\mathcal{A} := \{A_{(S_j,T,f)} \mid (S_j,T,f) \text{ is a matching}\}$ of events is a partition of Ω^{n_1,n_2}. This fact and Lemma 3.7 give

$$\Pr\left[\bigcap_{i\in I}\overline{A_i}\right] = \sum_{(S_j,T,f)\in\mathcal{A}}\Pr\left[\bigcap_{i\in I}\overline{A_i}\cap A_{(S_j,T,f)}\right] \geqslant$$
$$\geqslant \sum_{(S_j,T,f)\in\mathcal{A}}\Pr\left[\bigcap_{i\in I}\overline{A_i}\cap A_j\right] =$$
$$= \sum_{(S_j,T,f)\in\mathcal{A}}\Pr\left[\bigcap_{i\in I}\overline{A_i}\,\middle|\, A_j\right]\cdot\Pr[A_j] =$$
$$= \sum_{(S_j,T,f)\in\mathcal{A}}\Pr\left[\bigcap_{i\in I}\overline{A_i}\,\middle|\, A_j\right]\cdot\Pr\left[A_{(S_j,T,f)}\right] = \Pr\left[\bigcap_{i\in I}\overline{A_i}\,\middle|\, A_j\right],$$

which is the desired inequality. □

3.2.2 An Application to Packing Problems

We present one application of the above results on negative dependency graphs from [11] here, namely results on hypergraph packing problems. For clarity, we repeat definitions concerning hypergraphs.

Definition 3.8. *(i) A* hypergraph $H = (V, E)$ *is a finite set of vertices* V *together with a set* $E \subseteq 2^V$ *of edges.*

(ii) A subhypergraph *of a hypergraph* $H = (V, E)$ *is a hypergraph* $H' = (V', E')$ *with* $V' \subseteq V$ *and* $E' \subseteq E$.

(iii) An r-uniform hypergraph *is a hypergraph* $H = (V, E)$ *with* $|e| = r$ *for all* $e \in E$.

(iv) A complete r-uniform hypergraph *on* n *vertices is a hypergraph* $H = (V, E)$ *with* $|V| = n$ *and* $E = \binom{V}{r}$, *denoted by* $K_n^{(r)}$.

(v) Two subhypergraphs $H_1 = (V_1, E_1)$ *and* $H_2 = (V_2, E_2)$ *of* $K_n^{(r)}$ *are* edge-disjoint *if* $E_1 \cap E_2 = \emptyset$.

The *hypergraph packing problem* is the following ([11]): Given two r-uniform hypergraphs $H_1 = (V_1, E_1)$ and $H_2 = (V_2, E_2)$ and an integer n such that $n \geqslant \max\{|V_1|, |V_2|\}$, do there exist embeddings φ_1 and φ_2 of H_1 and H_2, respectively, into $K_n^{(r)}$ such that $\varphi_1(H_1)$ and $\varphi_2(H_2)$ are edge-disjoint?

The following theorem states sufficient conditions under which the hypergraph packing problem can be answered positively.

Theorem 3.9 ([11]). *Let* $H_1 = (V_1, E_1)$ *and* $H_2 = (V_2, E_2)$ *be two* r*-uniform hypergraphs with* $|E_1| = m_1$ *and* $|E_2| = m_2$. *Let* d_1 *and* d_2 *be integers such that for both* $i \in \{1, 2\}$, *every edge in* E_i *intersects at most* d_i *other edges in* E_i. *If* n *is an integer such that*

$$(d_1 + 1) \cdot m_2 + (d_2 + 1) \cdot m_1 \leqslant \frac{1}{e}\binom{n}{r}, \tag{3.6}$$

then there exist injections $\varphi_i : V_i \to [n]$ *such that the natural images* $\varphi_1(H_1)$ *and* $\varphi_2(H_2)$ *are edge-disjoint.*

Proof. Let H_1, H_2 and integers m_1, m_2, d_1, d_2, n be as in the theorem such that (3.6) holds.

Without loss of generality, assume that H_2 is given as a subhypergraph of $K_n^{(r)}$ on the vertex set $[n]$. We want to embed V_1 into $[n]$ via an injection φ_1 such that no edge of $\varphi_1(H_1)$ coincides with an edge of H_2.

First of all, the interpretation of injections as matchings allows translating the injection $\varphi_1 : V_1 \to [n]$ to a maximum cardinality matching with f-representation $(V_1, \varphi_1(V_1), \varphi_1)$ in the complete bipartite graph with parts V_1 and $[n]$. Obviously, this correspondence is one to one.

Moreover, given two edges $e \in E_1$ and $f \in E_2$, the maximum cardinality matchings realising the bad event that e is mapped to f are exactly the ones in $\bigcup_\varphi A_{(e,f,\varphi)}$, where the union is taken over all bijections $\varphi : e \to f$.

Our goal is to show that a random maximum cardinality matching in the complete bipartite graph with parts V_1 and $[n]$ has a positive probability of avoiding all events of the form $A_{(e,f,\varphi)}$ with e, f and φ as above. To do so, we apply the symmetric lopsided Local Lemma, Lemma 1.6.

The conflict graph G of all the events

$$\mathcal{A} := \{A_{(e,f,\varphi)} \mid e \in E_1, f \in E_2, \varphi : e \to f \text{ bijective}\}$$

is a negative dependency graph by Theorem 3.6. To bound the maximum degree d in G, observe the following: Two events $A_{(e,f,\varphi)}$ and $A_{(e',f',\varphi')}$ in $\mathcal{A}$ are conflicting if and only if one of the following two points is satisfied:

(i) The edges e and e' have empty intersection while their images f and f' have non-empty intersection.

(ii) The edges e and e' have non-empty intersection but φ and φ' are defined differently at a vertex $v \in e \cap e'$.

Fix an event $A_{(e,f,\varphi)}$. To determine the degree d of the corresponding vertex in G, we count the number of events in $\mathcal{A}$ conflicting $A_{(e,f,\varphi)}$.

By assumption, there are m_1 edges in E_1 and at most d_2+1 edges in E_2 intersecting f (including f), so there are at most $m_1(d_2+1)$ choices for the edges e' and f' of an event $A_{(e',f',\varphi')}$ generating a conflict of type (i) with $A_{(e,f,\varphi)}$. For each such choice of edges, there are $r!$ choices of a bijection $e \to f$, giving a total bound of $r! \cdot m_1(d_2+1)$ conflicts of type (i).

Similarly, there are at most d_1+1 edges in E_1 intersecting e and there are m_2 edges in E_2, hence at most $m_2(d_1+1)$ choices of edges e' and f' of an event $A_{(e',f',\varphi')}$ generating a conflict of type (ii) with $A_{(e,f,\varphi)}$. Analogously to the above, this gives a total bound of $r! \cdot m_2(d_1+1)$ conflicts of type (ii).

Note that the above counts the event $A_{(e,f,\varphi)}$ itself in both cases, so together with the assumption (3.6), this gives

$$d+1 < r! \cdot \big(m_1(d_2+1) + m_2(d_1+1)\big) \leqslant \frac{r!}{e}\binom{n}{r} = \frac{n!}{e \cdot (n-r)!}. \tag{3.7}$$

The probability of an event $A_{(e,f,\varphi)} \in \mathcal{A}$ is

$$p := \Pr\Big[A_{(e,f,\varphi)}\Big] = \frac{1}{r! \cdot \binom{n}{r}} = \frac{(n-r)!}{n!} \tag{3.8}$$

because there are $\binom{n}{r}$ choices for the image of e and $r!$ choices for the bijection mapping e to this image; exactly one of these choices corresponds to f and φ.

To apply Lemma 1.6, we need $e(d+1)p \leqslant 1$, which is a direct consequence of (3.7) and (3.8). The lemma implies that for a random matching in the complete bipartite graph with parts V_1 and $[n]$, the probability of avoiding all events in $\mathcal{A}$, which is $\Pr[\bigcap_{A\in\mathcal{A}} \overline{A}]$, is positive, hence there exists a matching avoiding all events in $\mathcal{A}$.

This matching corresponds to the injection of $V_1 \to [n]$ inducing the desired inclusion of H_1 into $K_n^{(r)}$ which is edge-disjoint from H_2. □

Note that one can construct examples that show tightness of Theorem 3.9 up to the constant $\frac{1}{e}$. Lu and Székely provide one in [11] that shows that $\frac{1}{e}$ cannot be replaced by 2.

3.3 Random Matchings in K_n

This section proves results in the probability space of maximum cardinality matchings in complete graphs K_n with even n, analogously to what we did in Section 3.2 for matchings in complete bipartite graphs.

Note that maximum cardinality matchings in an even complete graph are always perfect matchings, so Ω^n is the space of perfect matchings. We prove that conflict graphs for canonical events in this space are negative dependency graphs. Moreover, we will find conditions under which the same graphs are also near-positive dependency graphs.

One can observe that the proof of Theorem 3.6 does not immediately generalise to graphs other than complete bipartite graphs because it relies on the structure provided by the bipartition. In contrast, the methods presented in this section for complete graphs are general in the sense that they can be easily adjusted to complete bipartite graphs. We stay with complete graphs here and point to [12] for the alterations.

We always assume that for positive integers n, the complete graph K_n has $[n]$ as the set of its vertices. For easier notation, we write Ω^n instead of Ω^{K_n} for the set of maximum cardinality matchings in K_n. Similarly, for a matching M in K_n, we use A^n_M instead of $A^{K_n}_M$ for the associated canonical event.

Under these assumptions, for $n, s \in \mathbb{Z}^+$, every matching M in K_n can be seen as a matching in K_{n+s} via the natural embedding of $[n]$ into $[n+s]$. Note that a matching M in K_n can generate different events A^n_M and A^{n+s}_M in Ω^n and Ω^{n+s}, respectively.

Every permutation σ of $[n]$ acts on matchings in K_n by permuting the underlying set of vertices $[n]$, namely via the map τ_σ defined as follows:

$$\{u, v\} \in M \qquad \Longleftrightarrow \qquad \{\sigma(u), \sigma(v)\} \in \tau_\sigma(M).$$

By element-wise application, τ_σ extends to a set function, in particular if $\tau_\sigma(M) = M'$, then $\tau_\sigma(A_M) = A_{M'}$. Moreover, τ_σ permutes Ω^n, hence as we use a uniform probability measure, τ_σ is measure-preserving.

3.3.1 Negative Dependency Graphs for Canonical Events

The following theorem is the analogue of Theorem 3.6, applying to matchings in complete graphs on an even number of vertices here.

Theorem 3.10 ([12]). *Let $n \in \mathbb{Z}^+$ be even and let matchings $M_1, \ldots, M_m$ in K_n generate canonical events $A_{M_1}, \ldots, A_{M_m}$ in Ω^n, where m is a positive integer. Then, the conflict graph $G = (V, E)$ for the events $A_{M_1}, \ldots, A_{M_m}$ is a negative dependency graph.*

The proof of the above theorem is partly outsourced into the following lemma.

Lemma 3.11 ([12]). *Let n be a positive integer. For any collection $\mathcal{M}$ of matchings in K_n, we have*

$$\Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^n}\right] \leqslant \Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n+2}}\right].$$

Proof. For $i \in [n+1]$, let E_i denote the partial matching in K_{n+2} consisting of the single edge $\{i, n+2\}$. Note that the events $\{A_{E_i} \mid i \in [n+1]\}$ partition the space Ω^{n+2}: Perfect matchings cover all $n+2$ vertices, hence in each such matching, exactly one of the edges $\{i, n+2\}$ appears.

Additionally, let $\mathcal{M}_i := \{M \in \mathcal{M} \mid M \text{ and } E_i \text{ do not conflict}\}$ for each $i \in [n+1]$. For $M \notin \mathcal{M}_i$, Proposition 3.3 (i) gives $A_M^{n+2} \cap A_{E_i} = \emptyset$, hence $A_{E_i} \subseteq \overline{A_M^{n+2}}$ or equivalently, $\overline{A_M^{n+2}} \cap A_{E_i} = A_{E_i}$. Combining these two observations, we get

$$\begin{aligned}\Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n+2}}\right] &= \sum_{i=1}^{n+1} \Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n+2}} \cap A_{E_i}\right] = \\ &= \sum_{i=1}^{n+1} \Pr\left[\bigcap_{M\in\mathcal{M}_i} \overline{A_M^{n+2}} \cap A_{E_i}\right]. \end{aligned} \tag{3.9}$$

For $i \in [n+1]$, let σ_i be the permutation of $[n+2]$ transposing i and $n+1$ and fixing all other elements. As each $M \in \mathcal{M}_i$ does not cover i, we have $\tau_{\sigma_i}\left(A_M^{n+2}\right) = A_M^{n+2}$. Moreover, $\tau_{\sigma_i}(A_{E_i}) = A_{E_{n+1}}$ and τ_{σ_i} is measure-preserving, so (3.9) extends to

$$\begin{aligned}\Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n+2}}\right] &= \sum_{i=1}^{n+1} \Pr\left[\tau_{\sigma_i}\left(\bigcap_{M\in\mathcal{M}_i} \overline{A_M^{n+2}} \cap A_{E_i}\right)\right] = \\ &= \sum_{i=1}^{n+1} \Pr\left[\bigcap_{M\in\mathcal{M}_i} \overline{A_M^{n+2}} \cap A_{E_{n+1}}\right] = \\ &= \sum_{i=1}^{n+1} \Pr\left[\bigcap_{M\in\mathcal{M}_i} \overline{A_M^{n+2}} \,\middle|\, A_{E_{n+1}}\right] \cdot \Pr[A_{E_{n+1}}] = \\ &= \sum_{i=1}^{n+1} \Pr\left[\bigcap_{M\in\mathcal{M}_i} \overline{A_M^{n}}\right] \cdot \Pr[A_{E_{n+1}}] \geqslant \\ &\geqslant \Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n}}\right] \cdot \sum_{i=1}^{n+1} \Pr[\tau_{\sigma_i}(A_{E_{n+1}})] = \Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n}}\right], \end{aligned} \tag{3.10}$$

which is the desired. The inequality in (3.10) holds by monotonicity of the probability measure because $\mathcal{M}_i \subseteq \mathcal{M}$. Moreover, we used

$$\Pr[A_{E_{n+1}}] = \Pr[\tau_{\sigma_i}(A_{E_{n+1}})] \quad \text{and} \quad \sum_{i=1}^{n+1} \Pr[\tau_{\sigma_i}(A_{E_{n+1}})] = \sum_{i=1}^{n+1} \Pr[A_{E_i}] = 1,$$

which follows from the facts that τ_{σ_i} is measure preserving and that the events A_{E_i} partition the Ω^{n+2}. ☐

We can now apply the above lemma to get a proof of Theorem 3.10.

Proof of Theorem 3.10. Let $(M_i)_{i\in[m]}$ be matchings generating canonical events $(A_{M_i})_{i\in[m]}$ with conflict graph G. Fix $j \in [m]$ and $I \subseteq [m] \setminus N_G(j)$. We need to prove

$$\Pr\left[A_{M_j} \,\middle|\, \bigcap_{i\in I} \overline{A_{M_i}}\right] \leqslant \Pr\left[A_{M_j}\right]. \tag{3.11}$$

If $\Pr\left[\bigcap_{i\in I} \overline{A_{M_i}}\right] = 0$, there is nothing to prove, so assume the opposite. Note that $\Pr[A_{M_j}] \neq 0$, so by definition of conditional probability, the above is equivalent to

$$\frac{\Pr\left[\bigcap_{i\in I} \overline{A_{M_i}} \cap A_{M_j}\right]}{\Pr\left[A_{M_j}\right]} \leqslant \Pr\left[\bigcap_{i\in I} \overline{A_{M_i}}\right].$$

By definition of I, M_i and M_j are not conflicting for any $i \in I$, so by (ii) in Proposition 3.3, we get

$$\frac{\Pr\left[\bigcap_{i\in I} \overline{A_{M_i}} \cap A_{M_j}\right]}{\Pr\left[A_{M_j}\right]} = \frac{\Pr\left[\bigcap_{i\in I} \overline{A_{M_i\setminus M_j}} \cap A_{M_j}\right]}{\Pr\left[A_{M_j}\right]} = \Pr\left[\bigcap_{i\in I} \overline{A_{M_i\setminus M_j}} \,\middle|\, A_{M_j}\right], \tag{3.12}$$

and combining (3.11) and (3.12), we see that it suffices to prove

$$\Pr\left[\bigcap_{i\in I} \overline{A_{M_i\setminus M_j}} \,\middle|\, A_{M_j}\right] \leqslant \Pr\left[\bigcap_{i\in I} \overline{A_{M_i}}\right]. \tag{3.13}$$

If $M_i \setminus M_j = \emptyset$ for some i, the left-hand side in the above is equal to zero and there is nothing to show. Otherwise, we transform it to be able to apply Lemma 3.11. Let $s = |\operatorname{supp}(M_j)|$ and define σ to be a permutation of $[n]$, the vertex set of the underlying K_n, such that

$$\sigma(\operatorname{supp}(M_j)) = \{n - s + 1, \dots, n\}.$$

Because for all $i \in I$, M_i and M_j are not conflicting, this gives $\operatorname{supp}(\sigma(M_i \setminus M_j)) \subseteq [n - s]$. Define the set $\mathcal{M}^I := \{\sigma(M_i \setminus M_j) \mid i \in I\}$ of matchings in K_{n-s}. Applying the measure-preserving τ_σ in the left-hand side of (3.13), we get

$$\begin{aligned}
\Pr\left[\bigcap_{i\in I} \overline{A^n_{M_i\setminus M_j}} \,\middle|\, A^n_{M_j}\right] &= \Pr\left[\tau_\sigma\left(\bigcap_{i\in I} \overline{A^n_{M_i\setminus M_j}}\right) \,\middle|\, \tau_\sigma\left(A^n_{M_j}\right)\right] = \\
&= \Pr\left[\bigcap_{i\in I} \overline{A^n_{\sigma(M_i\setminus M_j)}} \,\middle|\, A^n_{\sigma(M_j)}\right] = \Pr\left[\bigcap_{M\in\mathcal{M}^I} \overline{A^n_M} \,\middle|\, A_{\sigma(M_j)}\right] = \\
&= \Pr\left[\bigcap_{M\in\mathcal{M}^I} \overline{A^{n-s}_M}\right],
\end{aligned} \tag{3.14}$$

where the last equality follows because conditioning on $A_{\sigma(M_j)}$ fixes all edges in $\sigma(M_j)$, which cover $\{n - s + 1, \dots, n\}$.

Note that s is the cardinality of the support of a matching, hence it is even, so by applying Lemma 3.11 exactly $s/2$ times to the collection $\mathcal{M}^I$, we get

$$\Pr\left[\bigcap_{M\in\mathcal{M}^I} \overline{A_M^{n-s}}\right] \leqslant \Pr\left[\bigcap_{M\in\mathcal{M}^I} \overline{A_M^{n-s+2}}\right] \leqslant \ldots \leqslant \Pr\left[\bigcap_{M\in\mathcal{M}^I} \overline{A_M^{n}}\right].$$

Plugging back in the definition of $\mathcal{M}^I$ and using monotonicity of the probability measure, the above extends to

$$\Pr\left[\bigcap_{M\in\mathcal{M}^I} \overline{A_M^{n-s}}\right] \leqslant \Pr\left[\bigcap_{M\in\mathcal{M}^I} \overline{A_M^{n}}\right] = \Pr\left[\bigcap_{i\in I} \overline{A_{M_i\setminus M_j}}\right] \leqslant \Pr\left[\bigcap_{i\in I} \overline{A_{M_i}}\right]. \tag{3.15}$$

Putting together (3.14) and (3.15) gives (3.13), which proves the theorem. □

3.3.2 Near-Positive Dependency Graphs for Canonical Events

This chapter proves that under certain assumptions, the conflict graphs are not only negative dependency graphs, but at the same time near-positive dependency graphs.

In order to state the assumptions, we need the definitions of *δ-sparsity* and *r-boundedness.*

Definition 3.12 ([12]). *Let n be an even integer and let $\mathcal{M} = \{M_1, \ldots, M_m\}$ be a collection of matchings in K_n and let $\{A_{M_1}, \ldots, A_{M_m}\}$ be the events with conflict graph G they generate. Let $s_j = |\operatorname{supp}(M_j)|$ for all $j \in [m]$.*

(i) For all $M \in \mathcal{M}$, define

$$\mathcal{M}_M := \left\{ M' \setminus M \;\middle|\; \begin{array}{c} M' \in \mathcal{M},\ M' \neq M,\ M' \cap M \neq \emptyset, \\ M' \text{ and } M \text{ are not conflicting} \end{array} \right\}.$$

*(ii) The collection $\mathcal{M}$ is said to be δ-*sparse *if*

- *No matching from $\mathcal{M}$ is subset of another matching in $\mathcal{M}$.*
- *For all $j \in [m]$, we have*

$$\Pr\Big[A_{M_j}\Big] < \delta \quad \text{and} \quad \sum_{i\in N_G(j)} \Pr[A_{M_i}] + 2\Pr[A_{M_i}]^2 < \delta. \tag{3.16}$$

- *For all edges $\{i,j\}$ of K_n, we have*

$$\sum_{M\in\mathcal{M}:\,\{i,j\}\in M} \Pr[A_M] + 2\Pr[A_M]^2 < \delta. \tag{3.17}$$

- *For any $j \in [m]$ and a permutation σ_j mapping* $\operatorname{supp}(M_j)$ *to the set $\{n - s_j + 1, \ldots, n\}$, we have*

$$\sum_{M\in\mathcal{M}_{M_j}} \Pr\Big[A^{n-s_j}_{\sigma_j(M)}\Big] + \Pr\Big[A^{n-s_j}_{\sigma_j(M)}\Big]^2 < \delta. \tag{3.18}$$

*(iii) For a positive integer r, we say that $\mathcal{M}$ is r-*bounded *if for all $j \in [m]$, we have $|M_j| \leqslant r$.*

Note that the condition (3.18) does not depend on the particular choice of σ_j. The point of permuting the vertices via σ_j is that it allows interpreting a matching $M \setminus M_j$ as a matching in the complete graph on the vertices $[n - s_j]$. Alternatively, one could just think of the matching $M \setminus M_j$ as a matching in the complete graph on the vertices $[n] \setminus \text{supp}(M_j)$.

For technical reasons, we also need the following proposition.

Proposition 3.13. *For each $\gamma \in (0, 1/4)$, the equation*

$$1 = ye^{-\gamma y}$$

has a unique solution $y(\gamma) \in [1, 2]$ and defines a function $y : (0, 1/4) \to [1, 2]$.

This proposition is just of analytical nature, so we take it for granted for now and defer its proof to Appendix B (Proposition B.1). Having the above at hand, we can now state the main theorem of this section, giving near-positive dependency graphs.

Theorem 3.14 ([12]). *For a positive integer m and an even positive integer n, let $A_1, \dots, A_m$ be canonical events in Ω^n generated by the matchings in the family $\mathcal{M} = \{M_1, \dots, M_m\}$. If $\mathcal{M}$ is δ-sparse and r-bounded for some $\delta \in (0, 1/8)$ and a positive integer r, then the conflict graph G for the events $A_1, \dots, A_m$ is an ε-near positive dependency graph with*

$$\varepsilon = 1 - e^{-\delta y(2\delta) - \delta^2 y(2\delta)^2} y(2\delta)^{-2r}.$$

The proof of this theorem uses parts of the lopsided Local Lemma. In the following lemma, we adjust the lopsided Local Lemma to the setting that we have here in such a way that it provides precisely the conclusions that we need.

Lemma 3.15 ([12]). *Let $A_1, \dots, A_m$ be events in an arbitrary probability space with negative dependency graph G and let $\varepsilon \in (0, 1/4)$ such that for all $j \in [m]$, we have*

$$\Pr[A_j] < \varepsilon \quad \textit{and} \quad \sum_{i \in N_G(j)} \Pr[A_i] + 2 \Pr[A_i]^2 < \varepsilon. \tag{3.19}$$

Then, we have the following:

(i) For any $S, T \subseteq [m]$ with $S \cap T = \emptyset$, we have

$$\Pr\left[\bigcap_{i \in S} \overline{A_i} \,\middle|\, \bigcap_{j \in T} \overline{A_j} \right] \geqslant \prod_{i \in S} \Big(1 - \Pr[A_i] y(\varepsilon) \Big).$$

(ii) We have

$$\Pr\left[\bigcap_{i=1}^{m} \overline{A_i} \right] \geqslant \exp\left(- \sum_{i=1}^{m} \Big(\Pr[A_i] y(\varepsilon) + \Pr[A_i]^2 y(\varepsilon)^2 \Big) \right).$$

Proof. The idea is to use parts of the proof of Lemma 1.5, the lopsided Local Lemma, so for all $i \in [n]$, we need $x_i \in [0, 1)$ such that

$$\Pr[A_i] \leqslant x_i \cdot \prod_{j \in N_G(i)} (1 - x_j).$$

We claim that setting $x_i = \Pr[A_i]y(\varepsilon)$ is a good choice: On the one hand, $x_i \in (0, 1/2)$ as $0 \leqslant \Pr[A_i] < \varepsilon \leqslant 1/4$ and $y(\varepsilon) \in [1, 2]$. On the other hand, the definition of $y(\varepsilon)$, the assumption (3.19), $y(\varepsilon) \leqslant 2$ and the inequality $e^{-x-x^2} \leqslant 1 - x$ give

$$\Pr[A_i] = \frac{x_i}{y(\varepsilon)} = x_i e^{-\varepsilon y(\varepsilon)} \leqslant x_i \exp\left(-\sum_{j\in N_G(i)}\left(x_j + x_j^2\right)\right) \leqslant x_i \cdot \prod_{j\in N_G(i)} (1 - x_j),$$

where the fact that $e^{-x-x^2} \leqslant 1 - x$ holds for $x \in (0, 1/2)$ can be proved by taking logarithms and comparing derivatives. So the assumptions for Lemma 1.5 hold.

To conclude (i), we use an intermediate result from the proof of the Local Lemma, namely (1.4): For every $j \in [m]$ and every set $I \subseteq [m]$, we have

$$\Pr\left[A_j \,\middle|\, \bigcap_{i\in I} \overline{A_i}\right] \leqslant x_j \tag{3.20}$$

Let $S, T \subseteq [m]$ such that $S \cap T = \emptyset$ and assume $S = \{m_1, \dots, m_s\}$. By repeated application of the definition of conditional probability,

$$\begin{aligned}
\Pr\left[\bigcap_{i\in S} \overline{A_i} \,\middle|\, \bigcap_{j\in T} \overline{A_j}\right] &= \frac{\Pr\left[\bigcap_{i=1}^{s} \overline{A_{m_i}} \cap \bigcap_{j\in T} \overline{A_j}\right]}{\Pr\left[\bigcap_{j\in T} \overline{A_j}\right]} = \\
&= \frac{\Pr\left[\overline{A_{m_1}} \,\middle|\, \bigcap_{i=2}^{s} \overline{A_{m_i}} \cap \bigcap_{j\in T} \overline{A_j}\right] \cdot \Pr\left[\bigcap_{i=2}^{s} \overline{A_{m_i}} \cap \bigcap_{j\in T} \overline{A_j}\right]}{\Pr\left[\bigcap_{j\in T} \overline{A_j}\right]} = \dots = \\
&= \prod_{k=1}^{s} \Pr\left[\overline{A_{m_k}} \,\middle|\, \bigcap_{i=k+1}^{s} \overline{A_{m_i}} \cap \bigcap_{j\in T} \overline{A_j}\right] = \prod_{k=1}^{s}\left(1 - \Pr\left[A_{m_k} \,\middle|\, \bigcap_{i\in I_k} \overline{A_i}\right]\right) \geqslant \\
&\geqslant \prod_{k=1}^{s} (1 - x_{m_k}) = \prod_{i\in S}\Big(1 - \Pr[A_i]y(\varepsilon)\Big),
\end{aligned}$$

where $I_k := T \cup \{m_{k+1}, \dots, m_s\}$ and the inequality follows from (3.20). This proves (i), and to conclude (ii), set $S = [m]$, $T = \emptyset$ and lower-bound the product using $1 - x \geqslant e^{-x-x^2}$ as we did when checking the conditions Lemma 1.5 above. ☐

The proof of Theorem 3.14 is in spirit similar to the one of Theorem 3.10, and the following lemma corresponds to Lemma 3.11.

Lemma 3.16 ([12]). *For a positive integer m, let $A_1, \dots, A_m$ be canonical events in Ω^n generated by matchings in $\mathcal{M} = \{M_1, \dots, M_m\}$, let G be the conflict graph of the events $A_1, \dots, A_m$ and let $\varepsilon \in (0, 1/4)$ be such that we have*

$$\forall j \in [m] : \quad \Pr[A_j] < \varepsilon \quad \textit{and} \quad \sum_{i\in N_G(j)} \Pr[A_i] + 2\Pr[A_i]^2 < \varepsilon, \tag{3.21}$$

$$\textit{and} \quad \forall \{i, j\} \in \binom{[n]}{2} : \quad \sum_{M\in\mathcal{M}:\{i,j\}\in M} \Pr[A_M] + 2\Pr[A_M]^2 < \varepsilon. \tag{3.22}$$

Then we have

$$\Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n+2}}\right] \leqslant y(\varepsilon)^2 \cdot \Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n}}\right].$$

Proof. We proceed as in the proof of Lemma 3.11: For $i \in [n+1]$, introduce the partial matchings E_i in K_{n+2} consisting of the single edge $\{i, n+2\}$ and note that the events A_{E_i} partition Ω^{n+2}. Moreover, let $\mathcal{M}_i$ denote the set of matchings in $\mathcal{M}$ not covering vertex i.

In parts of (3.10), we proved that in this setting,

$$\Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^{n+2}}\right] = \sum_{i=1}^{n+1} \Pr\left[\bigcap_{M\in\mathcal{M}_i} \overline{A_M^{n}}\right] \cdot \Pr[A_{E_i}].$$

As $\sum_{i=1}^{n+1} \Pr[A_{E_i}] = 1$, it suffices to show that for every $i \in [n+1]$, we have

$$\frac{\Pr\left[\bigcap_{M\in\mathcal{M}} \overline{A_M^n}\right]}{\Pr\left[\bigcap_{M\in\mathcal{M}_i} \overline{A_M^n}\right]} = \Pr\left[\bigcap_{M\in\mathcal{M}\setminus\mathcal{M}_i} \overline{A_M^n} \,\middle|\, \bigcap_{M\in\mathcal{M}_i} \overline{A_M^n}\right] \geqslant y(\varepsilon)^{-2}. \tag{3.23}$$

By the assumption (3.21), we may apply Lemma 3.15. With $S = \mathcal{M}\setminus\mathcal{M}_i$ and $T = \mathcal{M}_i$, part (i) gives

$$\Pr\left[\bigcap_{M\in\mathcal{M}\setminus\mathcal{M}_i} \overline{A_M^n} \,\middle|\, \bigcap_{M\in\mathcal{M}_i} \overline{A_M^n}\right] \geqslant \prod_{M\in\mathcal{M}\setminus\mathcal{M}_i} \Big(1 - \Pr[A_M]y(\varepsilon)\Big). \tag{3.24}$$

Note that $\mathcal{M}\setminus\mathcal{M}_i$ is the set of matchings in $\mathcal{M}$ covering vertex i. We distinguish three cases:

- There are no matchings covering i. Then, the product in (3.24) is empty, so the right-hand side is one and in particular larger than $y(\varepsilon)^{-2}$ because $y(\varepsilon) \in [1,2]$.
- There is a vertex $j \neq i$ such that for all $M \in \mathcal{M}$, $i \in M$ implies $\{i,j\} \in M$. Then

  $$\prod_{M\in\mathcal{M}\setminus\mathcal{M}_i} \Big(1 - \Pr[A_M]y(\varepsilon)\Big) \geqslant \exp\left(-\sum_{\substack{M\in\mathcal{M}:\\ \{i,j\}\in M}} \Big(\Pr[A_M]y(\varepsilon) + \Pr[A_M]^2 y(\varepsilon)^2\Big)\right)$$

  by the inequality $1 - x \geqslant e^{-x-x^2}$, which holds for $x \in (0, 1/2)$ as we have it here. Using $y(\varepsilon) \leqslant 2$ and the assumption (3.22), the right-hand side can be lower bounded by $\exp(-\varepsilon y(\varepsilon)) = y(\varepsilon)^{-1} \geqslant y(\varepsilon)^{-2}$, which is what we need.
- If the previous cases do not apply, then there are matchings M_k and M_ℓ covering vertex i with different edges. Hence if a matching covers vertex i, it has to conflict either M_k or M_ℓ, so

  $$\prod_{M\in\mathcal{M}\setminus\mathcal{M}_i} \Big(1 - \Pr[A_M]y(\varepsilon)\Big) \geqslant \prod_{i\in N_G(k)} \Big(1 - \Pr[A_i]y(\varepsilon)\Big) \cdot \prod_{i\in N_G(\ell)} \Big(1 - \Pr[A_i]y(\varepsilon)\Big)$$

  For each of the products on the right-hand side, we can proceed as in the previous case: Use the inequality $1 - x \geqslant e^{-x-x^2}$, $y(\varepsilon) \leqslant 2$ and the assumption (3.21) to get a lower bound of $\exp(-\varepsilon y(\varepsilon)) = y(\varepsilon)^{-1}$ for each product, together giving the bound $y(\varepsilon)^2$ that were aiming for.

So in each of the three cases, we could bound the product in (3.24) by $y(\varepsilon)^{-2}$, which gives the inequality in (3.23). □

Now, we are ready to combine the two lemmas for a proof of Theorem 3.14.

Proof of Theorem 3.14. Recall that the conflict graph $G = (V, E)$ for the events $A_1, \ldots, A_m$ is defined by

$$V \coloneqq [m] \quad \text{and} \quad E \coloneqq \{\{i, j\} \mid M_i \text{ and } M_j \text{ are conflicting}\}$$

and that it is a negative dependency graph by Theorem 3.10.

To have an ε-near positive dependency graph according to Definition 1.8, we on the one hand need that $\Pr[A_i \cap A_j] = 0$ for all $\{i, j\} \in E$. But by definition, $\{i, j\} \in E$ is equivalent to A_i and A_j conflicting, which is in turn equivalent to $A_i \cap A_j = \emptyset$ by Proposition 3.3 (i), so $\Pr[A_i \cap A_j] = 0$ follows.

On the other hand, we need that for all $j \in [n]$ and $I \subseteq [n] \setminus (N_G(j) \cup \{j\})$, we have

$$\Pr\left[A_j \,\middle|\, \bigcap_{i \in I} \overline{A_i}\right] \geqslant (1 - \varepsilon) \Pr[A_j] \tag{3.25}$$

whenever the conditional probability is well-defined. Switching to the notation explicitly indicating the matchings, using the definition of conditional probability and Proposition 3.3 (ii) (which applies because by definition of I, M_j and M_i are not conflicting for $i \in I$), we get

$$\Pr\left[A_{M_j} \,\middle|\, \bigcap_{i \in I} \overline{A_{M_i}}\right] = \frac{\Pr\left[\bigcap_{i \in I} \overline{A_{M_i}} \cap A_{M_j}\right]}{\Pr\left[\bigcap_{i \in I} \overline{A_{M_i}}\right]} = \frac{\Pr\left[\bigcap_{i \in I} \overline{A_{M_i \setminus M_j}} \cap A_{M_j}\right]}{\Pr\left[\bigcap_{i \in I} \overline{A_{M_i}}\right]} =$$

$$= \frac{\Pr\left[\bigcap_{i \in I} \overline{A_{M_i \setminus M_j}} \,\middle|\, A_{M_j}\right]}{\Pr\left[\bigcap_{i \in I} \overline{A_{M_i}}\right]} \cdot \Pr\left[A_{M_j}\right].$$

Recall that $1 - \varepsilon = e^{-\delta y(2\delta) - \delta^2 y(2\delta)^2} \cdot y(2\delta)^{-2r}$, so after plugging this in and using the above, (3.25) is equivalent to

$$\Pr\left[\bigcap_{i \in I} \overline{A_{M_i \setminus M_j}} \,\middle|\, A_{M_j}\right] \geqslant e^{-\delta y(\delta) - \delta^2 y(\delta)^2} \cdot y(\delta)^{-2r} \cdot \Pr\left[\bigcap_{i \in I} \overline{A_{M_i}}\right]. \tag{3.26}$$

We now transform the left-hand side so that we can apply Lemma 3.16. Let $s = |\operatorname{supp}(M_j)|$ and let σ be a permutation of $[n]$, the underlying vertex set of K_n, such that

$$\sigma(\operatorname{supp}(M_j)) = \{n - s + 1, \ldots, n\}.$$

This gives $\operatorname{supp}(\sigma(M_i \setminus M_j)) \subseteq [n - s]$ for all $i \in I$ because M_i and M_j are not conflicting. Define the set $\mathcal{M}^I_{M_j} \coloneqq \{M_i \setminus M_j \mid i \in I\}$ of matchings. Applying the measure-preserving τ_σ in the left-hand side of (3.26), we get

$$\begin{aligned}\Pr\left[\bigcap_{i\in I}\overline{A^n_{M_i\setminus M_j}}\;\middle|\;A^n_{M_j}\right] &= \Pr\left[\tau_\sigma\left(\bigcap_{i\in I}\overline{A^n_{M_i\setminus M_j}}\right)\;\middle|\;\tau_\sigma\left(A^n_{M_j}\right)\right] = \\ &= \Pr\left[\bigcap_{i\in I}\overline{A^n_{\sigma(M_i\setminus M_j)}}\;\middle|\;A^n_{\sigma(M_j)}\right] = \Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^n_{\sigma(M)}}\;\middle|\;A_{\sigma(M_j)}\right] = \\ &= \Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^{n-s}_{\sigma(M)}}\right]. \end{aligned} \tag{3.27}$$

At this point, we can apply Lemma 3.16: The assumptions (3.18) and (3.17) in δ-sparsity imply that the matchings in $\sigma(\mathcal{M}^I_{M_j})$, which is a subset of $\sigma(\mathcal{M}_{M_j})$, satisfy the requirements in (3.21) and (3.22), respectively, with 2δ instead of ε. Iteratively applying Lemma 3.16 exactly $\frac{s}{2}$ times, we get

$$\begin{aligned}\Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^{n-s}_{\sigma(M)}}\right] &\geqslant \Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^{n-s+2}_{\sigma(M)}}\right]\cdot y(2\delta)^{-2} \geqslant \ldots \geqslant \\ &\geqslant \Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^{n}_{\sigma(M)}}\right]\cdot y(2\delta)^{-s} \geqslant \Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^{n}_{M}}\right]\cdot y(2\delta)^{-2r}, \end{aligned} \tag{3.28}$$

where we used that $y(\varepsilon) \geqslant 1$, that r-boundedness implies $s \leqslant 2r$, and that the permutation σ does not affect the probabilities.

Combining (3.27) and (3.28), we see that in order to conclude (3.26), it is enough to have

$$\Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^n_M}\right] \geqslant e^{-\delta y(2\delta)-\delta^2 y(2\delta)^2}\cdot\Pr\left[\bigcap_{i\in I}\overline{A^n_{M_i}}\right]. \tag{3.29}$$

Unwrapping the definitions and applying Proposition 3.3 (ii), we see that

$$\begin{aligned}\Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^n_M}\right] &= \Pr\left[\bigcap_{i\in I}\overline{A^n_{M_i\setminus M_j}}\right] = \Pr\left[\bigcap_{i\in I}\overline{A^n_{M_i\setminus M_j}}\right] = \\ &= \Pr\left[\bigcap_{i\in I}\overline{A^n_{M_i\setminus M_j}}\cap\overline{A^n_{M_i}}\right] = \Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^n_M}\cap\bigcap_{i\in I}\overline{A^n_{M_i}}\right], \end{aligned}$$

hence (3.29) is – after division by the probability on the right-hand side – equivalent to

$$\Pr\left[\bigcap_{M\in\mathcal{M}^I_{M_j}}\overline{A^n_M}\;\middle|\;\bigcap_{i\in I}\overline{A_{M_i}}\right] \geqslant e^{-\delta y(2\delta)-\delta^2 y(2\delta)^2},$$

This can be proved by an application of Lemma 3.15 (i) to the sets $\mathcal{M}^I_{M_j}$ and $\{M_i | i\in I\}$: The canonical events generated by the matchings in $\mathcal{M}^I_{M_j}\cup\{M_i|i\in I\}$ satisfy the assumption (3.19) with ε being equal to 2δ by the assumptions (3.16) and

(3.18) in the definition of δ-sparsity. Moreover, $\mathcal{M}^I_{M_j}$ and $\{M_i | i \in I\}$ are disjoint by the first assumption in δ-sparsity, so Lemma 3.15 gives

$$\Pr\left[\bigcap_{M \in \mathcal{M}^I_{M_j}} \overline{A^n_M} \;\middle|\; \bigcap_{i \in I} \overline{A^n_{M_i}}\right] \geqslant \prod_{M \in \mathcal{M}^I_{M_j}} \Big(1 - \Pr[A_M] y(2\delta)\Big).$$

Using $1 - x \geqslant e^{-x-x^2}$ and the assumption (3.18) together with $M^I_{M_j} \subseteq \mathcal{M}_{M_j}$, the above product can be lower bounded by

$$\exp\left(-\sum_{M \in \mathcal{M}^I_{M_j}} \Big(\Pr[A_M] y(2\delta) + \Pr[A_M]^2 y(2\delta)^2\Big)\right) \geqslant e^{-\delta y(2\delta) - \delta^2 y(2\delta)^2},$$

and combining the last two inequalities completes the proof of Theorem 3.14. □

3.3.3 Asymptotic Counting using Local Lemma

Looking at the last two sections, we established the basis of applying both the lopsided Local Lemma (Lemma 1.5) and the variation of the Local Lemma giving upper bounds (Lemma 1.9) in the setting of random matchings in complete graphs on an even number of vertices.

The idea for asymptotic counting using the Local Lemma is to define a sequence of probability spaces, in our case Ω^n, and events $A_1(n), \dots, A_{m(n)}(n)$ such that the lower and upper bounds on $\Pr\Big[\bigcap_{i=1}^{m(n)} \overline{A_i(n)}\Big]$ match asymptotically, i. e. are the same for $n \to \infty$.

The following theorem collects a set of sufficient conditions for this plan to work out.

Theorem 3.17 ([12]). *For a strictly increasing sequence of even positive integers n, let r_n be a positive integer and $\varepsilon_n \in (0, 1/16)$. Let $\mathcal{M}^{(n)}$ be a collection of matchings in K_n such that none of these matchings is subset of another. Let $\mu_n := \sum_{M \in \mathcal{M}^{(n)}} \Pr[A_M]$. Suppose that for all n, $\mathcal{M}^{(n)}$ satisfies*

(i) $|M| \leqslant r_n$ *for all* $M \in \mathcal{M}^{(n)}$.

(ii) $\Pr[A^n_M] < \varepsilon_n$ *for all* $M \in \mathcal{M}^{(n)}$.

(iii) $\sum_{M' \colon A^n_{M'} \cap A^n_M = \emptyset} \Pr[A^n_{M'}] < \varepsilon_n$ *for all* $M \in \mathcal{M}^{(n)}$.

(iv) $\sum_{M \colon \{i,j\} \in M} \Pr[A^n_M] < \varepsilon_n$ *for all edges* $\{i, j\}$ *in* K_n.

(v) $\sum_{M' \in \mathcal{M}^{(n)}_M} \Pr\Big[A^{n-s}_{\sigma(M')}\Big] < \varepsilon_n$ *for each* $M \in \mathcal{M}^{(n)}$*, where* $s = |\operatorname{supp}(M)|$ *and* σ *is a permutation of* $[n]$ *mapping* $\operatorname{supp}(M)$ *to* $\{n - s + 1, \dots, n\}$.

Then we have

$$\Pr\left[\bigcap_{M \in \mathcal{M}^{(n)}} \overline{A^n_M}\right] = e^{-\mu_n + \mathcal{O}(r_n \varepsilon_n \mu_n)} \quad \textit{for } n \to \infty,$$

and if $r_n \varepsilon_n \mu_n = \mathrm{o}(1)$ *for* $n \to \infty$*, we have*

$$\Pr\left[\bigcap_{M \in \mathcal{M}^{(n)}} \overline{A^n_M}\right] = \Big(1 + \mathcal{O}(r_n \varepsilon_n \mu_n)\Big) \cdot e^{-\mu_n} \quad \textit{for } n \to \infty. \tag{3.30}$$

In order to derive the asymptotics stated in the above theorem, we need bounds on the auxiliary function y introduced in Proposition 3.13.

Lemma 3.18. *Let* $y : (0, 1/4) \to [1, 2]$ *be as in Proposition 3.13. Then, for* $\gamma \in (0, 1/4)$*, we have*

$$y(\varepsilon) \leqslant 1 + 6\varepsilon.$$

As with Proposition 3.13, the proof of this lemma is of analytical nature, so we defer it to Appendix B (Lemma B.2).

Proof of Theorem 3.17. Let $G^{(n)}$ be the conflict graph for the events $\mathcal{M}^{(n)}$.

By Theorem 3.10, G is a negative dependency graph. Note that the assumptions (ii) and (iii) in Theorem 3.17 imply that the assumptions (3.19) of Lemma 3.15 hold for the events generated by the matchings in $\mathcal{M}^{(n)}$ with ε being equal to $2\varepsilon_n$, so by part (ii) of this lemma,

$$\begin{aligned}
\Pr\left[\bigcap_{M\in\mathcal{M}^{(n)}} \overline{A_M^n}\right] &\geqslant \exp\left(-\sum_{M\in\mathcal{M}^{(n)}}\left(\Pr[A_M^n]y(2\varepsilon_n) + \Pr[A_M^n]^2 y(2\varepsilon_n)^2\right)\right) > \\
&> \exp\left(-\sum_{M\in\mathcal{M}^{(n)}}\Pr[A_M^n]y(2\varepsilon_n) - \sum_{M\in\mathcal{M}^{(n)}}\Pr[A_M^n]\varepsilon_n y(2\varepsilon_n)^2\right) = \\
&= \exp\left(-\mu_n\left(y(2\varepsilon_n) + \varepsilon_n y(2\varepsilon_n)^2\right)\right) \geqslant \exp\left(-\mu_n - 16\varepsilon_n\mu_n\right). \qquad (3.31)
\end{aligned}$$

Here, the second inequality follows from $\Pr[A_M^n] < \varepsilon_n$ and the last is a consequence of Lemma 3.18: $y(\gamma) \leqslant 1 + 6\gamma$ together with $\varepsilon_n < \frac{1}{16}$ implies

$$y(2\varepsilon_n) + \varepsilon_n y(2\varepsilon_n)^2 \leqslant 1 + 12\varepsilon_n + \varepsilon_n + 24\varepsilon_n^2 + 144\varepsilon_n^3 < 1 + 16\varepsilon_n.$$

For upper bounds, we first observe that Theorem 3.14 yields that $G^{(n)}$ is a near-positive dependency graph for the events in $\mathcal{M}^{(n)}$: Assumption (i) in Theorem 3.17 precisely gives r_n-boundedness of $\mathcal{M}^{(n)}$. Moreover, $\mathcal{M}^{(n)}$ is $2\varepsilon_n$-sparse: No matching is subset of another by assumption, condition (3.16) is implied by assumptions (ii) and (iii), condition (3.17) is implied by (iv) and conditon (3.18) is implied by (v).

Having the near-positive dependency graph from Theorem 3.14 at hand, an application of Lemma 1.9 gives

$$\begin{aligned}
\Pr\left[\bigcap_{M\in\mathcal{M}^{(n)}} \overline{A_M^n}\right] &\leqslant \prod_{M\in\mathcal{M}^{(n)}}\left(1 - e^{-2\varepsilon_n y(4\varepsilon_n) - 4\varepsilon_n^2 y(4\varepsilon_n)^2} y(4\varepsilon_n)^{-2r_n}\cdot\Pr[A_M^n]\right) \leqslant \\
&\leqslant \exp\left(-\sum_{M\in\mathcal{M}^{(n)}} e^{-2\varepsilon_n y(4\varepsilon_n) - 4\varepsilon_n^2 y(4\varepsilon_n)^2} y(4\varepsilon_n)^{-2r_n}\cdot\Pr[A_M^n]\right) = \\
&= \exp\left(-\mu_n e^{-2\varepsilon_n y(4\varepsilon_n) - 4\varepsilon_n^2 y(4\varepsilon_n)^2 - 8\varepsilon_n r_n y(4\varepsilon_n)}\right) < \\
&< \exp\left(-\mu_n\left(1 - \left(2y(4\varepsilon_n) + 4\varepsilon_n y(4\varepsilon_n)^2 + 8r_n y(4\varepsilon_n)\right)\varepsilon_n\right)\right) < \\
&< \exp\left(-\mu_n + (7 + 20r_n)\varepsilon_n\mu_n\right). \qquad (3.32)
\end{aligned}$$

Here, the second and third inequality use $1-x \leqslant e^{-x}$, the equality in between follows from the definition of y and μ_n, and the last inequality follows from Lemma 3.18: $y(\gamma) \leqslant 1+6\gamma$ gives

$$2y(4\varepsilon_n) + 4\varepsilon_n y(4\varepsilon_n)^2 + 8r_n y(4\varepsilon_n) \leqslant \\ \leqslant 2 + 48\varepsilon_n + 4\varepsilon_n + 4\cdot 48\varepsilon_n^2 + 4\cdot 576\varepsilon_n^3 + 8r_n + 192 r_n\varepsilon_n < 7 + 20r_n$$

because $\varepsilon_n < \frac{1}{16}$.

From the lower and upper bounds in (3.31) and (3.32), we directly see that

$$\Pr\left[\bigcap_{M\in\mathcal{M}^{(n)}} \overline{A_M^n}\right] = e^{-\mu_n + \mathcal{O}(\mu_n r_n \varepsilon_n)},$$

proving the first part of the theorem.

If $r_n\varepsilon_n\mu_n = \mathrm{o}(1)$ for $n\to\infty$, then for $n\to\infty$, both

$$e^{-16\varepsilon_n\mu_n} = 1 + \mathcal{O}(r_n\varepsilon_n\mu_n) \quad \text{and} \quad e^{(8r_n+7)\varepsilon_n\mu_n} = 1 + \mathcal{O}(r_n\varepsilon_n\mu_n),$$

so the bounds give

$$\Pr\left[\bigcap_{M\in\mathcal{M}^{(n)}} \overline{A_M^n}\right] = \Big(1 + \mathcal{O}(\mu_n r_n\varepsilon_n)\Big)\cdot e^{-\mu_n},$$

which concludes the proof of the theorem. □

3.3.4 Enumeration of d-regular Graphs

We are going to apply the asymptotic counting method derived in the previous section by showing one of the applications due to Lu and Székely ([12]): Deriving the asymptotic number of labelled d-regular graphs on n vertices with a given lower bound g on the girth for $n\to\infty$.

It was Bollobás in [3] who already provided the precise asymptotics for $g=3$ using a different method. He introduced the so-called *configuration model*, which shows to be of advantage in our approach as well.

Definition 3.19. *The* degree sequence *of a labelled n-vertex graph G with vertices $v_1,\dots,v_n$ is the sequence* $(d_1,\dots,d_n) \coloneqq (\deg_G(v_1),\dots,\deg_G(v_n))$.

One can observe that if a sequence $(d_1,\dots,d_n)$ of non-negative integers is a degree sequence, then $\sum_{i=1}^n d_i$ is even.

Definition 3.20 ([3]). *Let $(d_1,\dots,d_n)\in\mathbb{Z}_{\geqslant 0}^n$ be a degree sequence of length n. Then the* configuration model *of a random multigraph with given degree sequence $(d_1,\dots,d_n)$ is defined as follows:*

- *Let $W \coloneqq \bigcup_{i=1}^n W_i$ be a fixed set of $2m = \sum_{i=1}^n d_i$ labelled vertices, where $|W_i| = d_i$. We call those $2m$ vertices* mini-vertices.
- *A* configuration *F of W is a matching of cardinality m on the vertices in W.*

– *The multigraph associated to a configuration F of W is the graph on vertices $v_1, \dots, v_n$ where for all $k, \ell \in [n]$, the number of edges $\{v_k, v_\ell\}$ is equal to the number of edges in F connecting a mini-vertex in W_k to a mini-vertex in W_ℓ. This identification is called the* projection *of F.*

Choosing a random configuration F, this construction gives a random multigraph with degree sequence $(d_1, \dots, d_n)$.

Using this model and Theorem 3.17, we prove the following theorem:

Theorem 3.21 ([12]). *Let $d = d(n) \geqslant 3$ and $g = g(n)$ be integers such that*

$$g^5(d-1)^{2g-3} = \mathrm{o}(n)$$

for $n \to \infty$. Then, the probability that a random d-regular multigraph as generated by the configuration model has girth at least $g \geqslant 1$ is

$$\Big(1 + \mathrm{o}(1)\Big) \cdot \exp\left(-\sum_{i=1}^{g-1} \frac{(d-1)^i}{2i}\right),$$

and therefore, the number of d-regular graphs on n vertices with girth at least $g \geqslant 3$ is

$$\Big(1 + \mathrm{o}(1)\Big) \cdot \exp\left(-\sum_{i=1}^{g-1} \frac{(d-1)^i}{2i}\right) \frac{(dn-1)!!}{(d!)^n}.$$

As indicated in the statement, the proof of the theorem allows both d and g to slowly go to infinity. However, we suppress the dependence on n in the sequel and only remark here that all steps can be done in the same way even if one takes the dependencies on n in to account.

Proof. We work with the configuration model of a random multigraph on n vertices with constant degree sequence $(d, \dots, d)$, i. e. we consider matchings in the complete graph K_{nd} on the nd mini-vertices occurring in the configuration model. Let W denote the set of those nd labelled mini-vertices and let $W_1, \dots, W_n$ be the partition of W where each W_i corresponds to one vertex v_i of the multigraph via projection. To estimate the probability in question, we apply Theorem 3.17.

For $i \in [g-1]$, let $\mathcal{M}_i^{(n)}$ be the set of partial matchings in W whose projections precisely give a cycle of length i. Let $\mathcal{M}^{(n)} := \bigcup_{i=1}^{g-1} \mathcal{M}_i^{(n)}$. Then, the events that we want to avoid are precisely the canonical events generated by the matchings in $\mathcal{M}^{(n)}$.

It is immediately seen that no matching in $\mathcal{M}^{(n)}$ is subset of another because no cycle is subset of another cycle.

We have

$$\left|\mathcal{M}_i^{(n)}\right| = \binom{n}{i} \frac{i!}{2i} \cdot d^i (d-1)^i.$$

This can be seen as follows: There are $\binom{n}{i}$ choices for the vertices appearing in the cycle after projection and $\frac{i!}{2i}$ options for choosing the order in which they appear (note that reversing an ordering does not change the cycle). Moreover, if we orient the cycle, then for each vertex, there are d choices for the mini-vertex incident to the incoming edge and $d-1$ choices for the one incident to the outgoing edge.

For each matching $M \in \mathcal{M}_i^{(n)}$, we have

$$\Pr[A_M] = \frac{1}{(nd-1)(nd-3)\cdot \ldots \cdot (nd-2i+1)}, \tag{3.33}$$

so summing up gives

$$\begin{aligned}\mu_n = \sum_{M\in\mathcal{M}} \Pr[A_M] &= \sum_{i=1}^{g-1} \binom{n}{i} \frac{i!}{2i} \frac{d^i(d-1)^i}{(nd-1)(nd-3)\cdot \ldots \cdot (nd-2i+1)} = \\ &= \sum_{i=1}^{g-1} \frac{(d-1)^i}{2i} \prod_{\ell=1}^{i} \frac{d(n-\ell+1)}{nd-2\ell+1} \leqslant \sum_{i=1}^{g-1} \frac{(d-1)^i}{2i} \left(1+\mathcal{O}\left(\frac{i^2}{n}\right)\right) \leqslant \\ &\leqslant \left(1+\mathcal{O}\left(\frac{g^2}{n}\right)\right) \cdot \sum_{i=1}^{g-1} \frac{(d-1)^i}{2i},\end{aligned}$$

where the asymptotics can be seen by writing the product in the form $\prod(1-x_\ell)$ and using the inequality $\prod(1-x_\ell) \geqslant 1 - \sum x_\ell$, where $x_\ell = \frac{d\ell-2\ell-d+1}{dn-2\ell+1}$.

We now check the conditions in Theorem 3.17 one after another for $r_n = g-1$ and $\varepsilon_n = \frac{K' g^3 (d-1)^{g-2}}{n}$ with some large constant K' for all $n \in \mathbb{Z}^+$. Obviously, for n large enough, $\varepsilon \in (0, 1/16)$ for sure.

Condition (i): *$\mathcal{M}^{(n)}$ is r_n-bounded for all $n \in \mathbb{Z}^+$.*

Matchings in $\mathcal{M}^{(n)}$ correspond to cycles of lengths at most $g-1$, so they are of cardinality at most $g-1$. Hence $\mathcal{M}^{(n)}$ is r_n-bounded and condition (i) holds.

Condition (ii): *For all $M \in \mathcal{M}^{(n)}$, we have* $\Pr[A_M] < \varepsilon_n$.

This condition is clearly true by choice of ε and (3.33).

Condition (iii): *For all $M \in \mathcal{M}^{(n)}$, we have* $\sum_{M' \colon A_{M'} \cap A_M = \emptyset} \Pr[A_{M'}] < \varepsilon_n$.

We distinguish cases according to the size of M and M'. Remember that $A_{M'} \cap A_M = \emptyset$ is equivalent to M and M' conflicting.

Case 1: $M \in \mathcal{M}_1^{(n)}$.

This implies that M consists of a single edge $e = \{x, y\}$ inside one of the parts, say W_ℓ, corresponding to a vertex v_ℓ via projection.

Consider some $M' \in \mathcal{M}_1^{(n)}$ conflicting M, then M' consists of one edge f in the same part W_ℓ sharing exactly one mini-vertex with e. There are $2(d-2)$ such f, and the probability of the a corresponding event is $\Pr[A_{\{f\}}] = \frac{1}{nd-1}$.

Now take a matching $M' \in \mathcal{M}_i^{(n)}$ conflicting M for some $i \geqslant 2$, then M' covers x or y and corresponds to an i-cycle incident to the vertex v_ℓ after projection. The number of ways to choose such a matching M' can be derived as follows: At first, there are $\binom{n-1}{i-1}$ ways to choose the vertices different from v_ℓ on the cycle, which can then be arranged in an oriented cycle together with v_ℓ in $(i-1)!$ ways. Let this oriented cycle contain the vertices $v_\ell, v_{j_1}, \dots, v_{j_{i-1}}$ in this order, corresponding to sets $W_\ell, W_{j_1}, \dots, W_{j_{i-1}}$ of mini-vertices.

Choose whether M' shall be incident to x or y, say we choose x. We claim that there are less than $d^{i-1}(d-1)^i$ ways to choose the edges of M': There are d ways to choose a mini-vertex w_{j_1} from W_{j_1} forming the edge $\{x, w_{j_1}\}$. Then there are $(d-1)d$ ways to choose the next edge $\{w'_{j_1}, w_{j_2}\}$ with $w_{j_1} \neq w'_{j_1} \in W_{j_1}$ and $w_{j_2} \in W_{j_2}$, etc. For the last edge $\{w'_{j_{i-1}}, w_\ell\}$, there are only $(d-1)^2$ choices because we must not choose $w_\ell = x$. This procedure double-counts the matchings M' incident to both x and y, so the result $2d \cdot (d(d-1))^{i-1} \cdot (d-1)^2 = 2d^{i-1}(d-1)^i$ is an upper bound.

This gives a total of less than $2\binom{n-1}{i-1}(i-1)! \cdot d^{i-1}(d-1)^i$ ways to choose a matching M' conflicting the edge e with a corresponding canonical event of probability $\Pr[A_{M'}] = \frac{1}{(nd-1)(nd-3)\cdot\ldots\cdot(nd-2i+1)}$. Altogether, we obtain

$$\begin{aligned}\sum_{\substack{M' \in \mathcal{M}_1^{(n)}:\\ A_{M'} \cap A_M = \emptyset}} \Pr[A_{M'}] &= \frac{2(d-2)}{nd-1} + \sum_{i=2}^{g-1} \sum_{\substack{M' \in \mathcal{M}_i^{(n)}:\\ A_{M'} \cap A_M = \emptyset}} \Pr[A_{M'}] < \\ &< \frac{2(d-2)}{nd-1} + \sum_{i=2}^{g-1} \frac{2\binom{n-1}{i-1}(i-1)! \cdot d^{i-1}(d-1)^i}{(nd-1)(nd-3)\cdot\ldots\cdot(nd-2i+1)} \leqslant \\ &\leqslant \frac{2(d-2)}{nd-1} + \sum_{i=2}^{g-1} \frac{2(d-1)^i}{nd-1} < \varepsilon_n \end{aligned} \tag{3.34}$$

for K' large enough: $K' > d-1$ is sufficient, for example. The second but last inequality follows from writing

$$\frac{2\binom{n-1}{i-1}(i-1)! \cdot d^{i-1}(d-1)^i}{(nd-1)(nd-3)\cdot\ldots\cdot(nd-2i+1)} = \frac{2(d-1)^i}{nd-1} \cdot \prod_{\ell=1}^{i} \frac{(n-\ell)d}{nd-\ell} \tag{3.35}$$

and bounding each factor in the product by 1.

Case 2: $M \in \mathcal{M}_k^{(n)}$ for some $k \geqslant 2$.

This implies that the edges of M connect k classes $W_{j_1}, \dots, W_{j_k}$ to form a cycle $v_{j_1}, \dots, v_{j_k}$ after projection.

First consider a matching $M' \in \mathcal{M}_1^{(n)}$ conflicting M, then M' is an edge inside one of the classes $W_{j_1}, \dots, W_{j_k}$ incident to an edge in M. In each of these classes, there are $2d-3$ edges incident to one of the two mini-vertices covered by M, giving $k(2d-3)$ in total. The corresponding probabilities of the canonical events associated to one of these single edges f are $\Pr[A_{\{f\}}] = \frac{1}{nd-1}$ as in Case 1.

Now take a matching $M' \in \mathcal{M}_i^{(n)}$ conflicting M for some $i \geqslant 2$, then one of the k classes $W_{j_1}, \dots, W_{j_k}$ contains two vertices x and y covered by M such that at least one of them is also covered by M'. Fixing one of the k classes, the arguments from Case 1 imply that there are less than $2\binom{n-1}{i-1}(i-1)! \cdot d^{i-1}(d-1)^i$ ways to choose such a matching M', so in total, there are less than k times as many choices for M'. Each of the associated canonical events $A_{M'}$ has probability $\Pr[A_{M'}] = \frac{1}{(nd-1)(nd-3)\cdot\ldots\cdot(nd-2i+1)}$ again, so we obtain

$$\begin{aligned}\sum_{\substack{M'\in\mathcal{M}_k^{(n)}:\\ A_{M'}\cap A_M=\emptyset}} \Pr[A_{M'}] &= \frac{k(2d-3)}{nd-1} + \sum_{i=2}^{g-1} \sum_{\substack{M'\in\mathcal{M}_i^{(n)}:\\ A_{M'}\cap A_M=\emptyset}} \Pr[A_{M'}] < \\ &< \frac{k(2d-3)}{nd-1} + k\sum_{i=2}^{g-1} \frac{2\binom{n-1}{i-1}(i-1)!\cdot d^{i-1}(d-1)^i}{(nd-1)(nd-3)\cdot\ldots\cdot(nd-2i+1)} \leqslant \\ &\leqslant \frac{(g-1)(2d-3)}{nd-1} + (g-1)\sum_{i=2}^{g-1}\frac{2(d-1)^i}{nd-1} < \varepsilon_n, \end{aligned} \tag{3.36}$$

where the second but last inequality follows as before and the last inequality again holds for K' large enough ($K' = d-1$ is sufficient, for example). Together, (3.34) and (3.36) affirm that condition (iii) holds.

Condition (iv): *All edges $\{i,j\}$ in K_{nd} satisfy $\sum_{M\colon\{i,j\}\in M} \Pr\left[A_M^{nd}\right] < \varepsilon_n$.*

Fix an edge $\{v_k, v_\ell\}$ in the projection arising from an edge $\{w_k, w_\ell\}$ connecting two mini-vertices from different sets W_k and W_ℓ. Note that $\{w_k, w_\ell\}$ cannot appear in any of the matchings in $\mathcal{M}_1^{(n)}$ because those consist of edges inside one class of mini-vertices.

We now bound the number of matchings $M \in \mathcal{M}_i^{(n)}$ with $i \geqslant 2$ containing the edge $\{w_k, w_\ell\}$. In addition to the classes W_k and W_ℓ, there are $\binom{n-2}{i-2}$ ways to choose the other $i-2$ classes, and there are $(i-2)!$ ways to arrange them and W_k and W_ℓ in a cycle with W_k and W_ℓ appearing consecutively. Similarly to what we did when checking the third condition, there are $(d-1)^i d^{i-2}$ ways to choose the mini-vertices forming the edges of M, which gives a total of $\binom{n-2}{i-2}(i-2)!\cdot(d-1)^i d^{i-2}$ ways to choose M. Each of the canonical events associated to a matching $M \in \mathcal{M}_i^{(n)}$ has probability $\Pr[A_M] = \frac{1}{(nd-1)(nd-3)\cdot\ldots\cdot(nd-2i+1)}$, which gives

$$\begin{aligned}\sum_{M\colon\{w_k,w_\ell\}\in M} \Pr[A_M] &= \sum_{i=2}^{g-1} \frac{\binom{n-2}{i-2}(i-2)!\cdot(d-1)^i d^{i-2}}{(nd-1)(nd-3)\cdot\ldots\cdot(nd-2i+1)} \leqslant \\ &\leqslant \frac{1}{d(n-1)}\sum_{i=2}^{g-1}\frac{(d-1)^i}{nd-1} < \varepsilon_n,\end{aligned}$$

where the first inequality follows from (3.35) and the second one holds by definition of ε_n if $K' \geqslant 1$, for example. This proves condition (iv).

Condition (v): *For all $M \in \mathcal{M}^{(n)}$ with $s = |\operatorname{supp}(M)|$ and a permutation σ of $[nd]$ mapping $\operatorname{supp}(M)$ to $\{nd-s+1,\ldots,nd\}$, we have $\sum_{M'\in\mathcal{M}_M^{(n)}} \Pr\left[A_{\sigma(M')}^{nd-s}\right] < \varepsilon_n$.*

Remember that

$$\mathcal{M}_M := \left\{ M'\setminus M \;\middle|\; \begin{array}{c} M'\in\mathcal{M},\ M'\neq M,\ M'\cap M\neq\emptyset, \\ M' \text{ and } M \text{ are not conflicting} \end{array} \right\}.$$

If $M \in \mathcal{M}_1^{(n)}$, then there is nothing to do because no other matching generating a cycle contains a loop after projection, which is what such an M corresponds to. So assume that $M \in \mathcal{M}_i^{(n)}$ and its projection corresponds to a cycle C_i. Take $M' = M_0 \setminus M \in \mathcal{M}_M^{(n)}$ with some $M_0 \in \bigcup_{s<d} \mathcal{M}_s^{(n)}$ corresponding to a cycle C_s after projection. We have $i, s \leqslant g-1$.

The components of $C_i \cap C_s$ having at least one edge are paths $P_1, \dots, P_t$ with $t \geqslant 1$. There are $2\binom{i}{2t}$ possible sets of t paths of length at least one in C_i: Choose the $2t$ vertices partitioning the cycle into $2t$ paths and take every other part to form a path, choosing from two options.

Let us now fix some paths and count how many cycles C_s generate exactly these paths. To the vertices in $P_1, \dots, P_t$, we add another ℓ vertices and edges joining them with the paths. This gives that the number of possible cycles C_s generating $P_1, \dots, P_t$ is upper bounded by

$$\sum_{\ell=1}^{g-1-2t} \binom{n}{l} \frac{(l+t-1)!}{2} \cdot 2^t,$$

where $\binom{n}{\ell}$ upper bounds the number of ways to choose the ℓ extra vertices, $\frac{(l+t-1)!}{2}$ accounts for the number of ways to arrange the paths and extra vertices in a cycle and 2^t counts the number of possible orientations of the t paths on the cycle.

We now count the number of matchings M_0 giving one particular cycle C_s with the mini-vertices in $M \cap M_0$ fixed. In the sets of mini-vertices corresponding to the ℓ extra vertices, we have $d(d-1)$ choices for the mini-vertices, while in the sets corresponding to ends of the t paths, we only need to choose one other mini-vertex, having $d-1$ choices – hence there are $d^\ell (d-1)^{\ell+2t}$ options for choosing M_0.

Now fix M_0 such that $M' = M \setminus M_0 \in \mathcal{M}_M^{(n)}$ and let $s' = \frac{1}{2}|\operatorname{supp}(M')|$. Let s, σ, ℓ and t be as above, then

$$\Pr\left[A_{\sigma(M')}^{nd-s}\right] = \frac{1}{(nd-s-1)(nd-s-3)\cdot \ldots \cdot (nd-s-2s'+1)} \leqslant \frac{1}{(nd-3g)^{\ell+t}},$$

where the inequality follows from $3g \geqslant s + 2s' - 1$ and the fact that M' has $s' = \ell + t$ edges. Putting the above together, we get

$$\sum_{M' \in \mathcal{M}_M^{(n)}} \Pr\left[A_{\sigma(M')}^{nd-s}\right] \leqslant \sum_{t=1}^{\lfloor i/2 \rfloor} 2\binom{i}{2t} \sum_{\ell=1}^{g-1-2t} \binom{n}{\ell} \frac{(\ell+t-1)!}{2} \cdot 2^t \cdot \frac{d^\ell (d-1)^{\ell+2t}}{(nd-3g)^{\ell+t}}. \quad (3.37)$$

We now use the falling factorial notation $(m)_k = \frac{m!}{(m-k)!}$. By the condition in the inner sum and $t \geqslant 1$, we have $\ell + t - 1 \leqslant g - 3$, and hence also $(\ell + t - 1)! \leqslant \ell!(\ell + t - 1)_{t-1} \leqslant \ell!(g-3)^{t-1}$. Moreover, $\binom{n}{\ell}\ell! = (n)_\ell$ and there is an absolute upper bound $K > \frac{(n)_\ell d^\ell}{(N-3g)^\ell}$. Using these observations, we can upper bound the right-hand side of (3.37) by

$$\sum_{t=1}^{\lfloor i/2\rfloor}\binom{i}{2t}\sum_{\ell=1}^{g-1-2t}K(d-1)^{g-1}\frac{2^t(g-3)^{t-1}}{(nd-3g)^t}\leqslant$$
$$\leqslant K(d-1)^{g-1}\sum_{t=1}^{\lfloor i/2\rfloor}\binom{i}{2t}\left(\frac{2(g-3)}{nd-3g}\right)^t\leqslant\frac{Kg^3(d-1)^{g-1}}{nd-3g}<\varepsilon_n,$$

where the estimate of the sum relies on the fact that the largest term occurs at $t=1$ for large enough n and there less than g terms. Moreover, we used that $\binom{i}{2}<\frac{i^2}{2}\leqslant\frac{g^2}{2}$, and ε_n satisfies the last inequality for large enough K'. This concludes the assertion of condition (v).

In order to conclude using (3.30), we need $r_n\varepsilon_n\mu_n=\mathrm{o}(1)$ for $n\to\infty$. We have

$$r_n\varepsilon_n\mu_n=(g-1)\frac{K'g^3(d-1)^{g-2}}{n}\left(1+\mathcal{O}\left(\frac{g^2}{n}\right)\right)\cdot\sum_{i=1}^{g-1}\frac{(d-1)^i}{2i}\leqslant$$
$$\leqslant\frac{K'g^5(d-1)^{2g-3}}{n}\cdot\left(1+\mathcal{O}\left(\frac{g^2}{n}\right)\right)=\mathrm{o}(1)$$

for $n\to\infty$, affirming the desired. So Theorem 3.17, more precisely (3.30), gives the desired asymptotics for the probability in question, namely

$$\Pr\left[\bigcap_{M\in\mathcal{M}^{(n)}}\overline{A_M}\right]=\Big(1+\mathcal{O}(r_n\varepsilon_n\mu_n)\Big)\cdot e^{-\mu_n}=\Big(1+\mathrm{o}(1)\Big)\cdot\exp\left(-\sum_{i=1}^{g-1}\frac{(d-1)^i}{2i}\right),$$

where the error term in μ_n can be neglected for $n\to\infty$ by the assumption. In total there are $(nd-1)!!$ matchings of the nd mini-vertices – but as permuting one class W_i of them does not change the projection, each multigraph is generated by $(d!)^n$ matchings. Hence the number of d-regular graphs with girth at least g is

$$\Big(1+\mathrm{o}(1)\Big)\cdot\exp\left(-\sum_{i=1}^{g-1}\frac{(d-1)^i}{2i}\right)\frac{(nd-1)!!}{(d!)^n}.$$

This concludes the proof of Theorem 3.21. □

Chapter 4

Algorithmic Aspects of the Local Lemma

As we explored in the previous chapters, the Lovász Local Lemma is a tool to prove the existence of an object in a certain probability space that avoids a family of events. From an algorithmic point of view, it is also of interest if one can design an algorithm that finds one of the desired objects efficiently.

In this last chapter, we present two constructive approaches to the problem. The first one is a rather general constructive proof of the Lovász Local Lemma due to Moser and Tardos ([17]). It provides a randomised algorithm that generates an instance avoiding all events under some mild assumptions on the underlying probability space and the events considered.

For a more concrete algorithmic approach, we address the so-called acyclic edge-colouring problem, which can be attacked using the non-constructive Lovász Local Lemma. In [7], Esperet and Parreau came up with an efficient randomised algorithm generating an acyclic edge-colouring of a given graph. Their result builds upon the so-called "entropy compression" method, which was introduced by Moser in an earlier, but less general attempt for an algorithmic proof of the Local Lemma ([16]).

4.1 A Constructive Proof of the Local Lemma

The problem that we are facing is the following: We are given a probability space Ω and a set $\mathcal{A}$ of events and the goal is to find an element $\omega \in \Omega$ such that no event in $\mathcal{A}$ occurs.

To make the above problem algorithmically accessible, it is necessary to impose some assumptions on the probability space and the events considered. One possible specification is to consider a finite collection $\mathcal{P}$ of mutually independent random variables on a fixed probability space Ω and only events that are determined by the values of some of these random variables.

These assumptions are satisfied in many applications: When working with random graphs, as we did in the previous chapters, for example, every edge could be modelled by a random variable having value 1 if the edge is present and value 0 otherwise.

In the subsequent two sections, we show that the above assumptions are sufficient for our purposes. The exposition closely follows the original paper [17] by Moser and Tardos.

4.1.1 The Resampling Algorithm

In order to state the actual algorithm, we introduce some terminology.

Definition 4.1. *Let $\mathcal{P}$ be a finite collection of mutually independent random variables in some probability space Ω.*

(i) An event A is determined *by a subset $S \subseteq \mathcal{P}$ if the values of the random variables in S are sufficient to decide whether or not A occurs.*

(ii) If A is determined by some subset $S \subseteq \mathcal{P}$, we say that an evaluation of the variables in S violates *A if it makes A happen.*

(iii) The minimal subset $S \subseteq \mathcal{P}$ that determines A is denoted by $\mathrm{vbl}(A)$.

The algorithm proposed by Moser and Tardos starts at some random point in Ω. As long as there is a violated event A, resample the variables in $\mathrm{vbl}(A)$. If eventually, no event is violated, then we are done. This idea is formalised in Algorithm 4.1 below.

Algorithm 4.1: The resampling algorithm ([17])

```
function sequential_lll (𝒫, 𝒜)
    for all P ∈ 𝒫 do
        v_P ← a random evaluation of P;
    end
    while ∃A ∈ 𝒜: A is violated by (v_P)_{P∈𝒫} do
        pick an arbitrary violated event A ∈ 𝒜;
        for all P ∈ vbl(A) do
            v_P ← a random evaluation of P;
        end
    end
    return (v_P)_{P∈𝒫};
```

The critical point is to argue whether or not this algorithm is efficient or even terminates. It is also a question how to measure efficiency: We assume that resampling variables and checking whether an event is violated can be done efficiently, so that the complexity of the algorithm is dominated by the number of resampling steps that need to be done.

Assuming the conditions of the Lovász Local Lemma, we prove an upper bound on the expected number of resampling steps done by the algorithm. This is the content of Theorem 4.3.

The Local Lemma needs a dependency graph for the events in question. In the setting that we introduced, we can define a dependency graph that we call the *intersection graph*.

Definition 4.2. *For a finite family $\mathcal{A}$ of events determined by a set of finitely many mutually independent random variables $\mathcal{P}$, the* intersection graph $G = (V, E)$ *is the graph defined by*

$$V = \mathcal{A} \quad \textit{and} \quad E = \{\{A, B\} \mid A, B \in \mathcal{A},\ \mathrm{vbl}(A) \cap \mathrm{vbl}(B) \neq \emptyset\}.$$

It can be easily seen that in the above setting, the intersection graph G is indeed a dependency graph for the events in $\mathcal{A}$.

Theorem 4.3 (Algorithmic Local Lemma, [17]). *Let $\mathcal{P}$ be a finite set of mutually independent random variables in a probability space, let $\mathcal{A}$ be a finite set of events determined by these variables and let G be their intersection. If there exist real numbers $(x_A)_{A\in\mathcal{A}} \in (0,1)^{|\mathcal{A}|}$ such that for all $A \in \mathcal{A}$, we have*

$$\Pr[A] \leqslant x_A \cdot \prod_{B \in N_G(A)} (1 - x_B), \tag{4.1}$$

then the following holds true:

(i) There exists an assignment of values to the variables in $\mathcal{P}$ not violating any of the events in $\mathcal{A}$.

(ii) The randomised Algorithm 4.1 resamples an event $A \in \mathcal{A}$ at most an expected $\frac{x_A}{1-x_A}$ times before it finds such an evaluation, and hence the expected number of resampling steps is at most $\sum_{a\in\mathcal{A}} \frac{x_A}{1-x_A}$.

The proof of this theorem is given in the next section.

4.1.2 Execution Logs and Witness Trees

To get our hands on what the algorithm actually does, we need to keep a record of the steps it performs.

In each iteration of the while-loop, the resampling algorithm, Algorithm 4.1, selects some violated event. The selection itself can be made deterministic by implementing any concrete rule, so that the choice of which variables to resample depends only on the current values $(v_P)_{P\in\mathcal{P}}$ of the random variables. For a formal bookkeeping of the choices, we define the so-called *execution log.*

Definition 4.4 ([17]). *The* execution log *of one execution of the resampling algorithm is the – possibly partial – map $C : \mathbb{N} \to \mathcal{A}$ such that for all n, we have the following: In the n-th iteration of the while-loop, the event $C(n)$ is picked and and its variables are resampled.*

Note that C may be only partial because the algorithm may terminate. If it terminates after k steps, we have an execution log of the form $C\colon [k] \to \mathcal{A}$. Also observe that C may be seen as a random variable depending on the random evaluations done in course of the resampling algorithm.

In order to proof Theorem 4.3, we bound the expected number of times an event occurs in the execution log. To this end, we introduce *witness trees.* Recall that $N_G^+(v) \coloneqq N_G(v) \cup \{v\}$ is the inclusive neighbourhood of a vertex v in G.

Definition 4.5 ([17]). *Let $\mathcal{A}$ be a finite family of events determined by mutually independent random variables in $\mathcal{P}$ and let G be their intersection graph.*

(i) *A* witness tree *$\tau = (T, \sigma_T)$ is a finite rooted tree $T = (V, E)$ with a labelling $\sigma : V \to \mathcal{A}$ of its vertices such that the children of any vertex $u \in V$ have labels in $N_G^+(u)$ only.*

(ii) *A witness tree is* proper *if the children of each vertex have pairwise distinct labels.*

(iii) *Let C be an execution log of length at least t. The witness tree $\tau_C(t)$ associated to step t is the witness tree constructed as follows:*

– *Let $\tau_C^{(t)}(t)$ be an isolated root vertex labelled $C(t)$.*

– *For each $i = t-1, \ldots, 1$, do the following: If there is no vertex $v \in \tau_C^{(i+1)}(t)$ such that $C(i) \in N_G^+(v)$, then let $\tau_C^{(i)}(t) = \tau_C^{(i+1)}(t)$.*

If there is such a vertex v, then choose one with maximum distance from the root in $\tau_C^{(i+1)}(t)$ and attach a child labelled $C(i)$ to it, constructing the witness tree $\tau_C^{(i)}(t)$.

– *Let $\tau_C(t) = \tau_C^{(1)}(t)$.*

(iv) *A witness tree τ* occurs *in an execution log C if there exists $t \in \mathbb{N}$ such that $\tau = \tau_C(t)$.*

If $\tau = (T, \sigma_T)$, we write $V(\tau)$ for the set of vertices of T and $[v] = \sigma_T(v)$ for the label of a vertex $v \in V(\tau)$ in order to keep notation short.

The connection from witness trees to the number of times an event appears in an execution log is the following: If $A = C(t)$ for some execution log C and a step $t \in \mathbb{N}$, then the witness tree $\tau_C(t)$ has its root labelled A. Moreover, we will observe that any two witness trees occurring in an execution log with identically labelled roots are different. These facts together with the following two lemmas are the basis of the proof of Theorem 4.3.

Lemma 4.6 ([17]). *Let $\mathcal{A}$ be a finite family of events determined by mutually independent random variables in $\mathcal{P}$. Let τ be a fixed witness tree and let C be a random execution log produced by the algorithm. Then, we have the following:*

(i) *If τ occurs in C, then τ is proper.*

(ii) *The probability that τ appears in C is at most $\prod_{v \in V(\tau)} \Pr[[v]]$.*

Proof. Assume that τ occurs in the execution log C so that $\tau = \tau_C(t)$ for some t. For a vertex $v \in V(\tau)$, let $d(v)$ denote the distance of vertex v from the root of τ and let $q(v)$ be the index of the step at which v is attached to $\tau_C(t)$, i. e. the largest index q such that v is contained in $\tau_C^{(q)}(t)$.

For part (i), assume for contradiction that there are two vertices u and v of equal depth in $\tau_C(t)$ such that $\mathrm{vbl}([u]) \cap \mathrm{vbl}([v]) \neq \emptyset$, which in particular is the case if they are labelled identically. Without loss of generality, assume $q(v) < q(u)$ and let w be the parent node of v.

At step $q(v)$, when we add vertex v to the tree $\tau_C^{(q(v)+1)}(t)$, we choose a vertex with label in $N_G^+([v])$ of maximum depth and attach v to it. Observe that u is already present at step $q(v)$, $[u] \in N_G^+([v])$ and $d(u) = d(v) = d(w) + 1 > d(w)$, so v is for sure not attached to w, a contradiction. This proves (i).

For (ii), Moser and Tardos introduce a routine called *τ-check* ([17]): For a witness tree τ, consider the vertices vertex $v \in V(\tau)$ one after another following an order of decreasing depth, take a random evaluation of the variables in $\mathrm{vbl}([v])$ and check whether or not this evaluation violates $[v]$. We say that the τ-check *passes* if all events are violated when checked.

For establishing a relation between the τ-check and the resampling algorithm, we need to fix a *random source*, i. e. a random list $P^{(0)}$, $P^{(1)}, \dots$ of values for each variable $P \in \mathcal{P}$. Every time we request a new evaluation of a variable P, we take the first unused element of the corresponding list in the random source.

We claim that if a witness tree τ occurs in the execution log generated by the algorithm using a fixed random source, then the τ-check using the same random source passes.

Assuming this claim for a moment, we see that the probability that a fixed witness tree τ occurs in the execution log is at most the probability that the τ-check passes. The latter can be trivially calculated to be $\prod_{v \in V(\tau)} \Pr[[v]]$, which proves part (ii).

To prove the claim, we first look at one step of the τ-check: Fix a witness tree $\tau = \tau_C(t)$, a vertex $v \in V(\tau)$ and a random variable $P \in \mathrm{vbl}([v])$. Moreover, let $S_v(P)$ denote the set of all $w \in V(\tau)$ such that $d(w) > d(v)$ and $P \in \mathrm{vbl}([w])$. Then $S_v(P)$ contains vertices that were considered before v in course of the τ-check, and more precisely it contains all of those where P was resampled: According to the first part of this lemma, no vertex w of equal depth can have $P \in \mathrm{vbl}([w])$. Hence resampling P at vertex v in course of the τ-check gives the value is $P^{(|S_v(P)|)}$ (note that the list starts at $P^{(0)}$), and this is true for all $P \in \mathrm{vbl}([v])$.

Similarly, we want to check the values of all $P \in \mathrm{vbl}([v])$ when the algorithm does step $q(v)$. Note that all vertices w satisfying $P \in \mathrm{vbl}([w]) \cap \mathrm{vbl}([v])$ and corresponding to events that were resampled before $[v]$ by the algorithm are of depth at least $d(v)+1$ because they could at least be attached to v. Equally, no vertex w satisfying $P \in \mathrm{vbl}([w]) \cap \mathrm{vbl}([v])$ and corresponding to an event that was resampled after $[v]$ and before step t by the algorithm can have depth larger than or equal to $d(v)$: If so, when adding v during the construction of τ, we could have attached v to w.

Hence $S_v(P)$ precisely contains all vertices w satisfying $P \in \mathrm{vbl}(w) \cap \mathrm{vbl}(v)$ corresponding to events that were resampled before $[v]$ by the algorithm – so before the algorithm resamples $[v]$, all $P \in \mathrm{vbl}([v])$ have the value $P^{(|S_v(P)|)}$ (note that here $P^{(0)}$ is used for the initial sampling). The necessity of resampling the variables in $\mathrm{vbl}([v])$ comes from $[v]$ being violated by the values $P^{(|S_v(P)|)}$, and this is exactly the condition for the τ-check to pass at vertex v, which proves the claim. □

A way of randomly constructing labelled trees is via a so-called *Galton-Watson process*, a multi-step process defined as follows: Fix an event $A \in \mathcal{A}$ and real numbers $(x_B)_{B \in \mathcal{A}} \in (0, 1)^{|\mathcal{A}|}$. Start by producing a single vertex labelled A in the first step.

In each of the following steps, go through all vertices v added in the previous step and add children to them in the following way: For all events $B \in N_G([v])$, we attach a child labelled B to v with probability x_B, and with probability $1 - x_B$, we do not. This process may stop at some point because there are no new vertices added in a step, or it may continue infinitely.

One can observe that a Galton-Watson process can produce a proper witness tree, but this is not necessarily always the case. The next lemma calculates the probability of getting a fixed proper witness tree.

Lemma 4.7 ([17]). *Let $\mathcal{A}$ be a finite family of events determined by mutually independent random variables in $\mathcal{P}$ and let G be their intersection graph. Moreover, let τ be a witness tree such that its root is labelled by $A \in \mathcal{A}$ and let $(x_B)_{B\in\mathcal{A}} \in (0,1)^{|\mathcal{A}|}$. The probability p_τ that the Galton-Watson process described above constructs the tree τ is given by*

$$p_\tau = \frac{1 - x_A}{x_A} \prod_{v\in V(\tau)} x'_{[v]},$$

where $x'_B \coloneqq x_B \cdot \prod_{C\in N_G(B)}(1 - x_C)$ for all $B \in \mathcal{A}$.

Proof. For every vertex $v \in V(\tau)$, let $D_v \subseteq N_G^+(v)$ be the subset of all events that occur as a label of a child node of v. In the step of the Galton-Watson process where we add children of v, the probability that they appear as in τ is given by $\prod_{[u]\in D_v} x_{[u]} \cdot \prod_{N_G^+([v])\setminus D_v}(1 - x_{[u]})$, and extending this to all probabilistic choices by multiplying the expressions for all $v \in V(\tau)$, we get

$$\begin{aligned} p_\tau &= \prod_{v\in V(\tau)} \left(\prod_{[u]\in D_v} x_{[u]} \cdot \prod_{[u]\in N_G^+([v])\setminus D_v} \left(1 - x_{[u]}\right) \right) = \\ &= \prod_{v\in V(\tau)} \left(\prod_{[u]\in D_v} \frac{x_{[u]}}{1 - x_{[u]}} \cdot \prod_{[u]\in N_G^+([v])} \left(1 - x_{[u]}\right) \right) = \\ &= \frac{1 - x_A}{x_A} \cdot \prod_{v\in V(\tau)} \left(\frac{x_{[v]}}{1 - x_{[v]}} \cdot \prod_{[u]\in N_G^+([v])} \left(1 - x_{[u]}\right) \right), \end{aligned}$$

where the last equation follows from the fact that taking a product over all children of all vertices in τ is the same as taking a product over all vertices but the root, which is not a child of any other vertex.

Replacing inclusive by exclusive neighbourhoods and using the definition of x'_B, this gives

$$p_\tau = \frac{1 - x_A}{x_A} \cdot \prod_{v\in V(\tau)} \left(x_{[v]} \cdot \prod_{[u]\in N_G([v])} \left(1 - x_{[u]}\right) \right) = \frac{1 - x_A}{x_A} \cdot \prod_{v\in V(\tau)} x'_{[v]},$$

which is what the lemma claimed. □

In the proof of Theorem 4.3, we now connect the above two lemmas.

Proof of Theorem 4.3. Let $\mathcal{P}$, $\mathcal{A}$ and G be as in the theorem statement and take real numbers $(x_A)_{A\in\mathcal{A}} \in (0,1)^{|\mathcal{A}|}$ such that the assumption (4.1) holds.

Let N_A be the the random variable (depending on the random resamplings of the variables in $\mathcal{P}$ by the algorithm) counting how may times the event A is resampled during one execution of the resampling algorithm. We already observed before that N_A is exactly the number of times that A appears in the execution log C, and every such appearance corresponds to a witness tree with root labelled A.

Moreover, it is easy to see that if $\tau_C(i)$ and $\tau_C(j)$ with $i > j$ are two witness trees with root labelled A, then they are different: The part of the log until step i has for sure strictly more occurrences of the event A than the part until step j, and hence the same is true for the corresponding witness trees.

Consequently, if we let $\mathcal{T}_A$ denote the set of all proper witness trees with root labelled A, we can calculate the expected value of N_A by summing the probability that a witness tree occurs in the execution log over all trees in $\mathcal{T}_A$:

$$\mathbb{E}[N_A] = \sum_{\tau\in\mathcal{T}_A} \Pr[\tau \text{ occurs in } C] \leqslant \sum_{\tau\in\mathcal{T}_A} \prod_{v\in V(\tau)} \Pr[[v]] \leqslant \sum_{\tau\in\mathcal{T}_A} \prod_{v\in V(\tau)} x'_{[v]},$$

where the first inequality follows from part (ii) in Lemma 4.6 and the second is assumption (4.1). By Lemma 4.7, we can replace the last product to get

$$\mathbb{E}[N_A] \leqslant \sum_{\tau\in\mathcal{T}_A} \frac{x_A}{1-x_A} p_\tau \leqslant \frac{x_A}{1-x_A},$$

where the last inequality follows from the fact that each Galton-Watson process defines exactly one tree, so the sum over all probabilities p_τ is at most one.

This proves that each event A is resampled at most an expected number of $\frac{x_A}{1-x_A}$ times, and linearity of expectation gives a total expected number of resampling steps of $\sum_{A\in\mathcal{A}} \frac{x_A}{1-x_A}$, proving part (ii) of the theorem.

For part (i), note that the above implies that there are random sources on which the algorithm terminates in one of the desired assignments of the variables in $\mathcal{P}$ after at most $\sum_{A\in\mathcal{A}} \frac{x_A}{1-x_A}$ many steps, which is a finite number and hence finishes the algorithmic proof of the Local Lemma. □

4.2 Acyclic Edge-Colouring

While the goal of the previous section was to derive a rather general algorithmic approach to the Local Lemma of, another strategy is to focus on one particular problem and make use of the extra structure coming with the specification. In this section, we do the latter and turn to acyclic edge-colouring, which is defined as follows.

Definition 4.8. *Let $G = (V, E)$ be a graph.*

(i) An acyclic edge-colouring *is a proper colouring of the edges E such that there is no two-coloured cycle.*

(ii) The acyclic chromatic index *of G, denoted $a'(G)$ is the smallest number of colours needed in an acyclic edge-colouring of G.*

In other words, an acyclic edge-colouring is a colouring such that the union of any two colour classes is a forest.

In [8], Fiamčík conjectured that for a graph G with maximum degree Δ, the acyclic chromatic index satisfies $a'(G) \leqslant \Delta + 2$, i. e. a bound only one larger than Vizing's bound for proper edge-colourings.

Using the Lovász Local Lemma, one can prove bounds on the acyclic chromatic index of a graph G depending on the maximum degree Δ, as for example $a'(G) \leqslant 16\Delta$ ([14]). For an algorithmic approach, we design an algorithm that colours the edges more or less randomly, always conducting correction steps if acyclicity is violated. Using the idea of *entropy compression* enables us to analyse the algorithm and prove conditions under which we get a positive probability that the algorithm halts, giving rather strong bounds on $a'(G)$ algorithmically.

4.2.1 The Algorithmic Approach

The algorithm for generating an acyclic edge-colouring is Algorithm 4.2. Its idea is simple: For a graph $G = (V, E)$ with ordered edges $E = \{e_1, \ldots, e_m\}$ and maximum degree Δ and a real number $\gamma > 1$, we colour the graph from $K := \lceil (2+\gamma)(\Delta - 1) \rceil$ many colours by an iterative process. Throughout the algorithm, we maintain the set X_i of uncoloured edges and a colouring $\Phi : E \to [K] \cup \{0\}$, where $\Phi(e_j) = 0$ encodes that e_j is not yet coloured.

More precisely, we choose an uncoloured edge $e_j \in X_i$ for colouring at step i. There may be colours appearing on edges incident to e_j and colours generating two-coloured cycles. We a priori want to avoid colours on incident edges and those generating a 2-coloured 4-cycle. There may still be larger 2-coloured cycles generated, but we perform a separate correction step in this case.

Definition 4.9. *For a graph $G = (V, E)$ and a partial colouring $\Phi \colon E \to [K] \cup \{0\}$, the set $\mathrm{col}_\Phi(e_j, G)$ of* neighbouring colours *of an edge $e_j = \{u, v\} \in E$ contains all colours $\Phi(e) \neq 0$, where $e = \{x, y\} \in E$, such that either $|e \cap e_j| = 1$, or $\{x, u\}, \{y, v\} \in E$ and $\Phi(\{x, u\}) = \Phi(\{y, v\}) \neq 0$.*

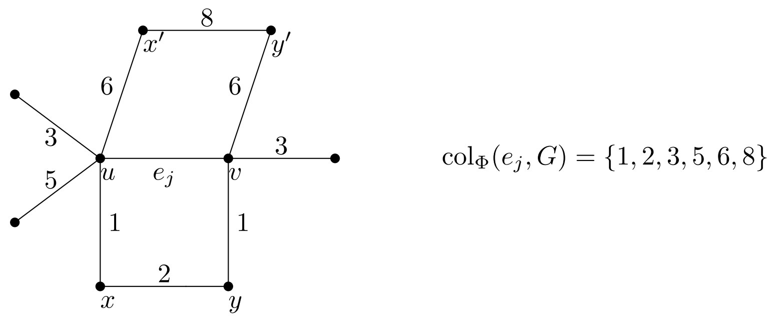

Figure 4.1: Neighbouring colours of an edge e_j.

Definition 4.9 captures exactly what we indicated before: The neighbouring colours are those that we want to avoid a priori. In the example in Figure 4.1, colours $\{1,3,5,6\}$ are forbidden because of edges incident to e_j and colours $\{2,8\}$ are forbidden because they would generate cycles of length 4.

Coming back to the algorithm, we colour the edge e_j using a random colour from $[K] \setminus \mathrm{col}_\Phi(e_j, G)$. As already indicated, this may still generate two-coloured cycles of length at least 6. If so, we choose one of them and uncolour all but the second and third edge after e_j. By the deterministic orientation of the chosen cycle, this is unambiguous.

This particular "correction step" of uncolouring all but two edges on a cycle on the one hand guarantees that after each step, the partial colouring remains acyclic, and on the other, we will see that it has certain advantages when it comes to keeping a useful record of the steps done by the algorithm. The same is true for the particular choice of a non-coloured edge e_j.

Algorithm 4.2: Acyclic edge-colouring ([7])

```
function acyclic_colouring(G, γ)
    K ← ⌈(2 + γ)(Δ − 1)⌉,   X ← E,   Φ(E) ← 0;
    while X ≠ ∅ do
        pick e_j ∈ X with minimum j;
        Φ(e_j) ← random colour from [K] \ col_Φ(e_j, G);
        if (G, Φ) has a 2−coloured cycle C do
            let C = e_j e_{i_2} ... e_{i_{2k}} such that i_2 < i_{2k};
            Φ({e_j, e_{i_4}, ..., e_{i_{2k}}}) ← 0;
            X ← X ∪ {e_{i_4}, ..., e_{i_{2k}}};
        else
            X ← X \ {e_j};
        end if
    end while
    return Φ
```

Note that if there is a 2-coloured cycle C, it can always be represented as $C = e_j e_{i_2} \dots e_{i_{2k}}$ with $i_2 < i_{2k}$ by choosing the right orientation. The particular choice of C is not important – this could be made deterministic by implementing any concrete rule as for example choosing the cycle C with lexicographically smallest $(i_2, \dots, i_{2k})$.

Of course, nothing forces Algorithm 4.2 to terminate – if the input γ is too small, it may even be the case that it never terminates. However, we will prove the following theorem.

Theorem 4.10 ([7]). *Let $\ell \geqslant 1$ be a fixed integer and let $k = \max\{2, \ell\}$. Let τ be the unique root of the polynomial $P(x) = (2k-3)x^{2k+2} - (2k-1)x^{2k} + x^4 - 2x^2 + 1$ in the open interval $(0,1)$. Then there is a positive probability for Algorithm 4.2 to halt if $\gamma = \frac{\tau^{2k} - \tau^2 + 1}{\tau - \tau^3}$ and G has maximum degree Δ and girth at least $2\ell+1$. In particular, every such graph has an acyclic edge-colouring with at most $\lceil (2+\gamma)(\Delta-1) \rceil$ colours.*

4.2.2 Entropy Compression & Analysis of the Algorithm

To start with, we describe the idea of the so-called *entropy compression* method ([18]). We are given an algorithm that modifies an object A step-by-step, using random choices to improve A.

We model these random choices by a random input string F such that in each step, the algorithm cuts off a prefix of F, leaving F' for the subsequent iterations, and deterministically improves A to A' using the bits cut off from F as a random source.

The key idea is to make sure that no information is lost when doing such a step – which can be achieved by keeping a history H' of all previous steps such that A and F can be reconstructed from A', F' and H'.

In general, the goal is that A', F' and H' compress the information of A and F. This can be measured using concepts as for example *Shannon entropy*, which – loosely speaking – measures the amount of information contained in a random string. If in each step, this entropy is reduced by an additive constant, we see that the process cannot go on forever, forcing the algorithm to halt.

In our specific situation, we get to the conclusion by slightly changing the argument. Note that it depends only on the random input string F whether or not the modified algorithm halts. We are able to show that for large enough t, the set $\mathcal{F}_t$ of strings F such that the algorithm did not halt at step t is of smaller cardinality than the number of possible states and histories (A', F', H') at step t, allowing the conclusion that there are strings F such that the algorithm halts after at most t steps.

Adjusting the Algorithm for Entropy Compression

For a given input (G, γ), the only random choices that Algorithm 4.2 does is to colour an edge e_j randomly from the set $[K] \setminus \mathrm{col}_\Phi(e_j, G)$ in each step. We first prove that this set of available colours is large.

Lemma 4.11 ([7]). *For every graph $G = (V, E)$ with maximum degree Δ and a partial edge-colouring $\Phi : E \to [K] \cup \{0\}$, we have $\mathrm{col}_\Phi(e, G) \leqslant 2(\Delta - 1)$ for all $e \in E$.*

Proof. Recall that the neighbouring colours $\mathrm{col}_\Phi(e, G)$ are those appearing either on edges incident to e or those that – assigned to e – would generate a 2-coloured 4-cycle.

Let $e = \{u, v\}$. The maximum degree is Δ, so both u and v are incident to at most $\Delta - 1$ edges other than e, which gives an upper bound of $2(\Delta - 1)$ different colours of the first type. Note that this bound is achieved only if all colours appearing on edges incident to u are pairwise different from colours appearing on edges incident to v.

For the second type, we need to count colours on edges $e' = \{x, y\}$ such that both $\{u, x\}, \{v, y\} \in E$ and $\Phi(\{u, x\}) = \Phi(\{v, y\}) \neq 0$. In particular, for each such edge e', we have edges $\{u, x\}$ and $\{v, y\}$ incident to u and v of the same colour – so this colour was counted twice among colours of the first type. Altogether, this proves $\mathrm{col}_\Phi(e, G) \leqslant 2(\Delta - 1)$. ☐

Having Lemma 4.11 at hand, we see that in each colouring step, the algorithm chooses a random colour from a set of $|[K] \setminus \mathrm{col}_\Phi(e, G)| \geqslant \lceil (2+\gamma)(\Delta - 1)\rceil - 2(\Delta - 1) = \lceil \gamma(\Delta - 1)\rceil$ many colours. Hence for colouring an edge e_j in step i, it is sufficient to have a random number $F_i \in \{1, \dots, \lceil \gamma(\Delta - 1)\rceil\}$ and to assign the F_i-th smallest element of $[K] \setminus \mathrm{col}_\Phi(e_j, G)$ to the edge e_j.

Consequently, the random choices in t steps of the algorithm can be modelled by inputting a random vector $F \in \{1, \dots, \lceil \gamma(\Delta-1)\rceil\}^t$ and choosing colours as described above.

Keeping a Record of the Steps

For given inputs, let X_i and Φ_i denote the uncoloured edges and the partial colouring after step i, respectively. Knowing the uncoloured edges determines the edge e_j coloured at step i because we choose it such that j is minimal. So the only thing we need to document is whether there the colouring in step i generates a 2-coloured cycle or not.

To do so, we define a vector R with entries R_i corresponding to steps i as follows: If the edge in e_j chosen in the step i is coloured by the algorithm without generating a 2-coloured cycle, we set $R_i = \emptyset$. Otherwise, there is a 2-coloured cycle $C = e_j e_{i_2} \dots e_{i_{2k}}$ of length $2k$ for some $k \geqslant 3$ on which the algorithm does the correction step. Note that the algorithm orients the cycle such that $i_2 < i_{2k}$.

There are at most $(\Delta - 1)^{2k-2}$ oriented cycles of the form $e_j e_{i_2} \dots e_{i_{2k}}$ with $i_2 < i_{2k}$, so ordering them lexicographically, we can identify C as the ℓ-th cycle for some $\ell \leqslant (\Delta - 1)^{2k-2}$. In this case, we set $R_i = (k, \ell)$.

The following two lemmas prove that if we run the modified algorithm on an input vector $F \in \{1, \dots, \lceil \gamma(\Delta - 1)\rceil\}^t$ for t steps, then F is uniquely determined by the record R and the colouring Φ_t.

Lemma 4.12 ([7]). *After each step $i \in [t]$, the set X_i is uniquely determined by the record $(R_j)_{j \leqslant i}$.*

Proof. We proceed by induction on i. For $i = 1$, i. e. after the first step, edge e_1 is coloured, so $X_1 = E \setminus \{e_1\}$.

Assume we know X_i for some $i < t$ and let $j = \min\{j' \mid e_{j'} \in X_i\}$. If $R_{i+1} = \emptyset$, then the algorithm coloured the edge e_j, so $X_{i+1} = X_i \setminus \{e_j\}$.

If $R_{i+1} = (k, \ell)$, then we can find the ℓ-th oriented cycle of length $2k$ in G containing e_j, say $C = e_j e_{i_2} \dots e_{i_{2k}}$, with $i_2 < i_{2k}$. We know that in this case, $X_{i+1} = X_i \setminus \{e_{i_4}, \dots, e_{i_{2k}}\}$. □

Lemma 4.13 ([7]). *For all $i \in [t]$, the map $(F_j)_{j \leqslant i} \mapsto ((R_j)_{j \leqslant i}, \Phi_i)$, namely the map assigning the input F to the record and the partial colouring after i steps of the algorithm, is injective.*

Proof. We again proceed by induction on i. After one step, corresponding to $i = 1$, edge e_1 is coloured using colour $\Phi_1(e_1)$. But this is exactly F_1.

Assume that for some $i < t$, $(F_j)_{j\leqslant i}$ is uniquely determined by $((R_j)_{j\leqslant i}, \Phi_i)$. The record $(R_j)_{j\leqslant i+1}$ uniquely determines X_i and X_{i+1} by Lemma 4.12. In particular, we know the edge e_j chosen for colouring in step $i+1$: Its index j is given by $j = \min\{j' \mid e_{j'} \in X_i\}$. To conclude by induction, we need to reconstruct F_{i+1} and Φ_i.

To start with, assume that $R_{i+1} = \emptyset$. Then the algorithm colours e_j using the F_{i+1}-th colour from $[K] \setminus \mathrm{col}_{\Phi_i}(e_j, G)$ in step $i+1$. Turning this around, we can see that

$$F_{i+1} = \Phi_{i+1}(e_j) - |\{c \in \mathrm{col}_{\Phi_{i+1}}(e_j, G) \mid c < \Phi_{i+1}(e_j)\}|, \tag{4.2}$$

where we used that $\mathrm{col}_{\Phi_i}(e_j, G) = \mathrm{col}_{\Phi_{i+1}}(e_j, G)$. Moreover, Φ_i is obviously reconstructed from Φ_{i+1} by setting the colour of e_j to zero.

If $R_{i+1} = (k, \ell)$, then a cycle $C = e_j e_{i_2} \dots e_{i_{2k}}$ that we can reconstruct from k and ℓ as in the proof of Lemma 4.12 was created by colouring e_j in step $i+1$. Since in this case, the algorithm uncolours all the edges in C but e_{i_2} and e_{i_3}, we know that $\Phi_i = \Phi_{i+1}$ on $E \setminus C$, $\Phi_i(e_{i_2}) = \Phi_i(e_{i_4}) = \dots = \Phi_i(e_{i_{2k}}) = \Phi_{i+1}(e_{i_2})$ and $\Phi_i(e_{i_3}) = \Phi_i(e_{i_5}) = \dots = \Phi_i(e_{i_{2k-1}}) = \Phi_{i+1}(e_{i_3})$, which reconstructs Φ_i from Φ_{i+1}. Moreover, we know that the colour assigned to e_j in step $i+1$ before the correction step must have been $\Phi_{i+1}(e_{i_3})$, so using $\Phi_{i+1}(e_{i_3})$ instead of $\Phi_{i+1}(e_j)$ in (4.2) reconstructs F_{i+1} and hence finishes the proof. □

Turning back to how we formulated the method of entropy compression in the introduction, Lemma 4.13 shows that keeping $H = ((R_j)_{j\leqslant t}, \Phi_t)$ as a history of the first t steps of the algorithm is a good choice when it comes to reconstructing the input. Moreover, the subsequent section shows that this "bookkeeping" provides enough compression to conclude.

Counting non-terminating Inputs: Dyck Words and Rooted Trees

Let $\mathcal{F}_t$ be the set of input vectors $F \in \{1, \dots, \lceil\gamma(\Delta-1)\rceil\}^t$ such that the algorithm does not halt after t steps. As pointed out in the introduction, the goal is to bound $|\mathcal{F}_t|$. If we denote by $\mathcal{R}_t$ the set of records generated by inputs from $\mathcal{F}_t$, we can prove the following lemma.

Lemma 4.14 ([7]). *Let G be a graph on m edges and let $\mathcal{F}_t$ be defined as above. Then $|\mathcal{F}_t| \leqslant (K+1)^m \cdot |\mathcal{R}_t|$.*

Proof. By Lemma 4.13, there is an injection

$$\mathcal{F}_t \to \mathcal{R}_t \times \{\Phi_t \mid \Phi_t \text{ is a partial colouring of } G\},$$

and as there are $(K+1)^m$ partial colourings of m edges, the postulated inequality $|\mathcal{F}_t| \leqslant (K+1)^m \cdot |\mathcal{R}_t|$ follows. □

For an easier counting, we convert records $R \in \mathcal{R}_t$ to binary strings following the approach of Esperet and Parreau ([7]). To do so, we need an auxiliary function θ_k on words $w = w_1 w_2 \dots w_{2k-2}$ of length $2k-2$ with $w_i \in [\Delta - 1]$, namely

$$\theta_k(w) = 1 + \sum_{i=1}^{2k-2} (w_i - 1)(\Delta - 1)^{i-1}.$$

This function θ_k maps the words of the prescribed form to $\{1, \dots, (\Delta-1)^{2k-2}\}$, and it is bijective. Using this function, we construct three maps whose concatenation transforms a record $R \in \mathcal{R}_t$ to a binary string, namely

$$R \mapsto R^* \mapsto R^\bullet \overset{\kappa}{\mapsto} R^\circ,$$

defined by the following:

- $R^* = (R_i^*)_{i \leqslant t}$ is obtained from $R = (R_i)_{i \leqslant t}$ by setting $R_i^* = 0$ if $R_i = \emptyset$ and $R_i^* = 0 \circ \theta_k^{-1}(\ell)$ if $R_i = (k, \ell)$, where $\circ$ denotes the concatenation of words. Remember that $\ell \leqslant (\Delta-1)^{2k-2}$, so this is well-defined.
- $R^\bullet$ is obtained from R^* by concatenation of all elements of R^*, which can be seen as a word over the alphabet $\{0, \dots, \Delta-1\}$.
- R° is obtained from the word $R^\bullet$ by keeping zeros and replacing non-zero letters by 1, i. e. by element-wise application of κ defined by $\kappa(x) = 0$ if $x = 0$ and $\kappa(x) = 1$ otherwise.

The resulting binary strings R° are not completely arbitrary, they are so-called *partial Dyck words*, as we prove in Lemma 4.16.

Definition 4.15. *(i) A* partial Dyck word *is a word w on the alphabet $\{0,1\}$ such that any prefix of w contains at least as many zeros as ones.*

(ii) A Dyck word *is a partial Dyck word that contains the same number of zeros and ones.*

(iii) A descent *in a partial Dyck word is a maximal sequence of consecutive ones.*

Dyck words have an interpretation as paths, where a zero indicates "up" and a one indicates "down", which is why we call a sequence of ones *descent*.

Lemma 4.16 ([7]). *If Algorithm 4.2 coloured r edges and generated a record $R \in \mathcal{R}_t$ after t steps, then the word R° is a partial Dyck word with t zeros and $t-r$ ones. Moreover, all descents in R° are even, and if G has girth at least $2\ell+1$ for some $\ell \geqslant 1$, then all descents in R° have length at least* $\max\{4, 2\ell\}$.

Proof. Note that by construction of R°, its zeros correspond exactly to the zeros in R^*, and ones correspond to non-zero letters in R^*. Recall that every R_i^* consists of exactly one zero and, if there was a 2-coloured cycle of length $2k$ generated at step i, an additional word of length $2k-2$ on the alphabet $\{1, \dots, \Delta-1\}$.

Consequently, R° is a concatenation of words of the form 01^{2k-2}, which can be interpreted as follows: The zero corresponds to colouring an edge, while each one corresponds to uncolouring an edge, following the algorithm chronologically. Taking this into account, it is clear that R° is a partial Dyck word: Edges can only be uncoloured if they were coloured previously, so every prefix has to contain more zeros than ones.

For the total number of zeros, note that there is one zero in each R_i^*, so there are t in total. The number of ones equals the total number of uncolouring steps, which is equal to $t-r$: t edges were coloured in total, and r remain coloured in the end.

The above also implies that all descents are of the form 1^{2k-2}, where $2k$ is the length of a generated two-coloured cycle in G, i. e. $2k \geqslant 6$ and of course $2k \geqslant 2\ell + 2$ by the condition on the girth of G. Altogether, descents are all even and of length at least $\max\{4, 2\ell\}$. □

For a non-empty set D of positive integers and positive integers t and $r \leqslant t$, let $C_{t,r,D}$ denote the number of partial Dyck words with t zeros, $t - r$ ones and all descents having lengths in D. Similarly, let $C_{t,D}$ be the number of Dyck words with all descents having lengths in D. Using this notation, the above lemma immediately leads to the following extension of Lemma 4.14:

Lemma 4.17. *Let G be a graph on m edges with girth at least $2\ell + 1$, let $k = \max\{2, \ell\}$ and let $\mathcal{F}_t$ be defined as before. Then with $D = 2\mathbb{N} + 2k$,*

$$|\mathcal{F}_t| \leqslant (K+1)^m \cdot \sum_{r=0}^{m-1} (\Delta - 1)^{t-r} \cdot C_{t,r,D}.$$

Proof. By Lemma 4.14, we have $|\mathcal{F}_t| \leqslant (K+1)^m \cdot |\mathcal{R}_t|$, so it suffices to prove the inequality $|\mathcal{R}_t| \leqslant \sum_{r=0}^{m-1} (\Delta - 1)^{t-r} \cdot C_{t,r,D}$.

Note that the maps associated to $R \mapsto R^*$ and $R \mapsto R^\bullet$ are both bijective, while $\kappa : R^\bullet \mapsto R^\circ$ is many-to-one. More precisely, we have $\kappa^{-1}(1) = \{1, \ldots, \Delta - 1\}$ and $\kappa^{-1}(0) = 0$. By Lemma 4.16, every R° is a partial Dyck word containing $t - r$ ones for some $r \in \{0, \ldots, m-1\}$, hence there are at most $(\Delta - 1)^{t-r}$ different pre-images $R^\bullet$. Moreover, Lemma 4.16 also proves that descents in R° are even and of length at least $2k$, i. e. have lengths in D.

Altogether, summing up over all different R° with $R \in \mathcal{R}_t$ and grouping them according to the value of r gives the desired $|\mathcal{R}_t| \leqslant \sum_{r=0}^{m-1} (\Delta - 1)^{t-r} \cdot C_{t,r,D}$. □

By the above lemma, we now need to bound $C_{t,r,D}$, the number of partial Dyck words with some properties. The subsequent lemma reduces this to counting slightly longer Dyck words and further transforms this to counting rooted trees on a fixed number of vertices.

Lemma 4.18. *Let t and $r \leqslant t$ be positive integers and let $D \neq \{1\}$ be a non-empty set of positive integers. Then:*

(i) If $s = \min D \setminus \{1\}$, we have $C_{t,r,D} \leqslant C_{t+r(s-1),D}$.

(ii) The number $C_{t,D}$ is equal to the number of rooted trees on $t + 1$ vertices such that the number of children of each vertex is in $D \cup \{0\}$.

Proof. For part (i), note that from any partial Dyck word w with t zeros, $t - r$ ones and all descents in D, we can construct a Dyck word of length $2t + 2r(s - 1)$ by concatenating w with $(0^{s-1}1^s)^r$. This adds $r(s - 1)$ zeros and rs ones, so the total length is $2t - r + r(s - 1) + rs = 2t + 2r(s - 1)$. This identification is injective because each Dyck word of lenght $2t + 2r(s - 1)$ ending in $(0^{s-1}1^s)^r$ corresponds uniquely to one of the prescribed partial Dyck words, hence $C_{t,r,D} \leqslant C_{t+r(s-1),D}$, proving part (i).

For part (ii), start with a rooted tree on $t+1$ vertices such that the number of children is in D for all vertices. Passing the vertices in a reversed breadth-first order, encode a vertex with i children by 1^i0, and when coming to the root, leave out the last zero. Concatenating all those strings gives a Dyck word of length $2t$ because every vertex but the root, which is not counted at all, is counted as a parent by adding a zero to the string before it is counted as a child of some other vertex by adding a one. Moreover, a maximal sequence of ones corresponds to children of some vertex, hence its length is in D.

This construction can also be reversed: Cut a given Dyck word into blocks of the form 1^i0. For each such block, choose the i "oldest" vertices that do not yet have a parent and join them at a common parent. If $i = 0$, then simply add a new leaf vertex. This concludes part (ii). □

In order to count rooted trees with the properties described above, we apply a corollary of a result by Drmota. For a proof of the theorem, we point to [5], but we deduce the corollary that we need here.

Theorem 4.19 ([5, Thm. 5]). *Let $\varphi(x) = \sum_{n\geqslant 0} \varphi_n x^n$ be a power series with non-negative coefficients such that $\varphi_0 > 0$ and $\varphi_j > 0$ for some $j \geqslant 2$ and let its radius of convergence be $R > 0$. Set $d = \gcd\{j > 0 \mid \varphi_j > 0\}$ and suppose that there exists $\tau \in (0, R)$ such that $\tau\varphi'(\tau) = \varphi(\tau)$. Moreover, let $y(x) = \sum_{n\geqslant 0} y_n x^n$ satisfy $y(x) = x\varphi(y(x))$. Then $y_n = 0$ for $n \not\equiv 1 \pmod d$, and else*

$$y_n = d\sqrt{\frac{\varphi(\tau)}{2\pi\varphi''(\tau)}}\frac{\varphi'(\tau)^n}{n^{3/2}}\left(1 + \mathcal{O}\left(n^{-1}\right)\right). \tag{4.3}$$

Corollary 4.20. *Let $D \neq \{1\}$ be a non-empty set of positive integers and define $\varphi_D(x) = 1 + \sum_{i\in D} x^i$. Assume that there exists $\tau \in (0, 1)$ such that $\tau\varphi'(\tau) = \varphi(\tau)$. Then τ is the unique solution of this equation in $(0, 1)$ and there is a constant c_D such that $C_{t,D} \leqslant c_D\gamma^t t^{-3/2}$, where $\gamma = \varphi'_D(\tau)$.*

Proof. We first address uniqueness of τ. Note that

$$x\varphi'_D(x) - \varphi_D(x) = \sum_{i\in D}(i-1)x^i - 1$$

is strictly increasing because $D \neq \{1\}$, so if there is a solution of $x\varphi'_D(x) - \varphi_D(x) = 0$ in $(0, 1)$, it is unique. Obviously, $\varphi_D(x)$ satisfies the conditions of Theorem 4.19. If D is finite, φ_D is a polynomial, which has of course infinite radius of convergence; if D is infinite, the radius of convergence equals 1.

Let $y(x) = \sum_{t\geqslant 0} C_{t,D}x^{t+1}$. By Lemma 4.18, this is the generating function for the number of rooted trees such that the number of children of each vertex is in $D \cup \{0\}$. But such a tree either is a single root, or it is a single root with i trees attached to it, where $i \in D$ and in each of the attached trees, the number of children of every vertex is in D. Translating this to a functional equation gives

$$y(x) = x \cdot \left(1 + \sum_{i\in D} y(x)^i\right) = x\varphi_D(x),$$

completing the check of the assumptions of Theorem 4.19. Hence (4.3) in Theorem 4.19 yields $C_{t,D} \leqslant c_D \gamma^t t^{-3/2}$, where all constants independent from t were absorbed into c_D, and $\gamma = \varphi'_D(\tau)$ as claimed. □

With all the above lemmas at hand, we can finally prove Theorem 4.10, the main theorem of this section.

Proof of Theorem 4.10. Remember that we still want to show that $|\mathcal{F}_t| < \lceil\gamma(\Delta - 1)\rceil^t$ because this gives the existence of an input F for the modified algorithm such that it halts after at most t steps.

Let $k = \max\{2, \ell\}$ and $D = 2\mathbb{N} + 2k$, then Lemma 4.17 together with Lemma 4.18 gives

$$\begin{aligned} |\mathcal{F}_t| &\leqslant (K+1)^m \cdot \sum_{r=0}^{m-1} (\Delta - 1)^{t-r} \cdot C_{t,r,D} \leqslant \\ &\leqslant (K+1)^m (\Delta - 1)^t \cdot \sum_{r=0}^{m-1} C_{t+r(2k-1),D}. \end{aligned} \tag{4.4}$$

In order to bound $C_{t+r(2k-1),D}$, we apply Corollary 4.20 with $D = 2\mathbb{N} + 2k$, i. e. $\varphi_D(x) = 1 + \frac{x^{2k}}{1-x^2}$. In this case, the equation $\varphi_D(x) = x\varphi'_D(x) = 0$ is equivalent to

$$\begin{aligned} & 1 + \frac{x^{2k}}{1-x^2} = x \cdot \frac{2kx^{2k-1}(1-x^2) + 2x^{2k+1}}{(1-x^2)^2} \\ \iff \quad & (2k-3)x^{2k+2} - (2k-1)x^{2k} + x^4 - 2x^2 + 1 = 0, \end{aligned}$$

which is the equation $P(x) = 0$ with P as in Theorem 4.10. By Corollary 4.20, there is a unique solution $\tau \in (0, 1)$ for the above equation, and there exists a constant c_D such that

$$C_{t+r(2k-1),D} \leqslant c_D \gamma^{t+r(2k-1)} (t + r(2k-1))^{-3/2} \leqslant c_D \gamma^{t+r(2k-1)} t^{-3/2},$$

where $\gamma = \varphi'_D(\tau) = \varphi_D(\tau)/\tau = \frac{\tau^{2k} - \tau^2 + 1}{\tau - \tau^3}$. Plugging the above into (4.4) gives

$$\begin{aligned} |\mathcal{F}_t| &\leqslant (K+1)^m (\Delta - 1)^t \cdot c_D t^{-3/2} \cdot \sum_{r=0}^{m-1} \gamma^{t+r(2k-1)} \leqslant \\ &\leqslant (K+1)^m \lceil\gamma(\Delta - 1)\rceil^t \cdot c_D t^{-3/2} \cdot \frac{\gamma^{m(2k-1)} - 1}{\gamma^{2k-1} - 1}, \end{aligned}$$

so that we get

$$\frac{|\mathcal{F}_t|}{\lceil\gamma(\Delta - 1)\rceil^t} \leqslant c_D (K+1)^m \cdot \frac{\gamma^{m(2k-1)} - 1}{\gamma^{2k-1} - 1} \cdot t^{-3/2} \to 0$$

for $t \to \infty$. In particular, the above fraction is smaller than 1 for t large enough, giving $|\mathcal{F}_t| < \lceil\gamma(\Delta - 1)\rceil^t$ and hence the existence of an input vector F such that the algorithm halts on F after at most t steps and gives an acyclic edge-colouring of G using at most $K = \lceil(\gamma + 2)(\Delta - 1)\rceil$ colours. □

We can calculate the constants occurring in the above proof for various lower bounds g on the girth of the graph G. This results in Table 4.1 with values of the constant c (rounded up to four decimals) such that we have $a'(G) \leqslant \lceil c \cdot (\Delta - 1) \rceil$ for all graphs G with maximum degree Δ. Empirically, we can observe that for this construction, c tends to 3 from above as g grows.

g	D	$P(x)$	τ	γ	c
3	$2\mathbb{N} + 4$	$x^6 - 2x^4 - 2x^2 + 1$	0.6180	2	4
10	$2\mathbb{N} + 8$	$7x^{12} - 9x^{10} + x^4 - 2x^2 + 1$	0.7279	1.4958	3.4959
50	$2\mathbb{N} + 48$	$47x^{52} - 49x^{50} + x^4 - 2x^2 + 1$	0.8933	1.1391	3.1392
100	$2\mathbb{N} + 98$	$97x^{102} - 99x^{100} + x^4 - 2x^2 + 1$	0.9344	1.0797	3.0798
200	$2\mathbb{N} + 198$	$197x^{202} - 199x^{200} + x^4 - 2x^2 + 1$	0.9609	1.0454	3.0454

Table 4.1: Acyclic chromatic indices for graphs with different girths.

The results are collected in the subsequent corollary. Note that they are way stronger than for example the bound of 16Δ that can be proved by a simple application of the Lovász Local Lemma mentioned in the beginning.

Corollary 4.21. *Let G be a graph of maximum degree Δ and girth g. Then we have the following:*

(i) If G is simple, then $a'(G) \leqslant 4(\Delta - 1)$.

(ii) If $g \geqslant 10$, then $a'(G) \leqslant \lceil 3.4959(\Delta - 1) \rceil$.

(iii) If $g \geqslant 50$, then $a'(G) \leqslant \lceil 3.1392(\Delta - 1) \rceil$.

(iv) If $g \geqslant 100$, then $a'(G) \leqslant \lceil 3.0798(\Delta - 1) \rceil$.

(v) If $g \geqslant 200$, then $a'(G) \leqslant \lceil 3.0454(\Delta - 1) \rceil$.

Appendix A

Bounding Probabilities

In Chapter 2, we repeatedly need to bound the probability that a certain random variable exceeds a given lower bound. This appendix shall derive the two methods used for the specific situations in Chapter 2 without aiming for the most general statements. More details are given in [2], which is at the same time the main source for this appendix.

A.1 Chernoff Bounds

The probabilities in (2.1) and (2.6) are both of the same spirit: We are given a binomially distributed random variable X with expected value $\mathbb{E}[X]$ and a positive number $a > 0$. The goal is to bound the probability $\Pr[X > \mathbb{E}[X] + a]$ from above.

A first attempt could be to directly apply Markov's inequality, which gives the bound

$$\Pr[X > \mathbb{E}[X] + a] \leqslant \frac{\mathbb{E}[X]}{\mathbb{E}[X] + a},$$

but this bound is not sufficiently sharp for our purposes. The trick of the so-called Chernoff bounds is to use monotonicity of the map $x \mapsto \exp(x)$, which implies that for any $\lambda > 0$, we have

$$\Pr[X > \mathbb{E}[X] + a] = \Pr\Big[e^{\lambda \cdot X} > e^{\lambda \cdot (\mathbb{E}[X]+a)}\Big]. \tag{A.1}$$

It turns out that applying Markov's inequality after this transformation gives way better results in the sense that the upper bounds are exponentially decreasing in a.

Theorem A.1. *Let X be a binomially distributed random variable with parameters n and p and let $a > 0$. The following two bounds hold:*

(i) $\Pr[X > np + a] \leqslant e^{a - a\ln\left(1+\frac{a}{np}\right) - np\ln\left(1+\frac{a}{np}\right)}$. *[2, Thm. A.1.10]*

(ii) $\Pr[X > np + a] \leqslant e^{-\frac{2a^2}{n}}$. *[2, Thm. A.1.4]*

Proof. As indicated in (A.1), for every $\lambda > 0$ we have

$$\Pr[X > np + a] = \Pr\Big[e^{\lambda X} > e^{\lambda(np+a)}\Big]$$

and applying Markov's inequality to the right-hand side yields

$$\Pr[X > np + a] \leqslant \mathbb{E}\Big[e^{\lambda X}\Big] \cdot e^{-\lambda(np+a)}. \tag{A.2}$$

By definition of the binomial distribution, we have

$$\mathbb{E}\Big[e^{\lambda X}\Big] = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{\lambda k} = \sum_{k=0}^{n} \binom{n}{k} (pe^{\lambda})^k (1-p)^{n-k} = \Big(pe^{\lambda} + 1 - p\Big)^n,$$

and plugging this into (A.2) gives

$$\Pr[X > np + a] \leqslant \Big(pe^{\lambda} + 1 - p\Big)^n \cdot e^{-\lambda(np+a)}.$$

To derive (i) and (ii), we need to bound the above expression appropriately.

For the first bound, we start with substitution $\lambda = \ln(1 + a/np)$ and then apply the well-known inequality $(1 + a/n)^n \leqslant e^a$ to get

$$\begin{aligned}\Pr[X > np + a] &\leqslant \Big(pe^{\lambda} + 1 - p\Big)^n \cdot e^{-\lambda(np+a)} = \\ &= \Big(1 + \frac{a}{n}\Big)^n \cdot e^{-\lambda(np+a)} \leqslant e^{a - a\ln\left(1+\frac{a}{np}\right) - np\ln\left(1+\frac{a}{np}\right)}.\end{aligned}$$

For the second one, we claim that

$$pe^{\lambda} + (1-p) \leqslant e^{\lambda^2/8 + \lambda p}, \tag{A.3}$$

implying

$$\Pr[X > np + a] \leqslant \Big(pe^{\lambda} + 1 - p\Big)^n \cdot e^{-\lambda(np+a)} \leqslant e^{\frac{\lambda^2 n}{8} + \lambda np} \cdot e^{-\lambda(np+a)} = e^{\frac{\lambda^2 n}{8} - \lambda a}.$$

Setting $\lambda = 4a/n$, we have $\frac{\lambda^2 n}{8} - \lambda a = -\frac{2a^2}{n}$, so

$$\Pr[X > np + a] \leqslant e^{-\frac{2a^2}{n}}$$

as claimed in (ii). It remains to prove (A.3). Taking logarithms, we see that the inequality is equivalent to

$$\ln\Big(1 - p + pe^{\lambda}\Big) \leqslant \frac{\lambda^2}{8} + \lambda p.$$

This inequality holds for $\lambda = 0$, so it is sufficient to show that the derivatives with respect to λ of the left- and right-hand side satisfy the same inequality, namely

$$\frac{pe^{\lambda}}{1 - p + pe^{\lambda}} \leqslant \frac{\lambda}{4} + p.$$

Repeating the same reasoning, we see that this inequality holds for $\lambda = 0$, so it suffices to prove

$$\frac{(1-p)pe^{\lambda}}{(1 - p + pe^{\lambda})^2} \leqslant \frac{1}{4},$$

and that last inequality is true by the inequality of arithmetic and geometric means applied in the form

$$(1-p) \cdot pe^{\lambda} \leqslant \left(\frac{(1-p) + pe^{\lambda}}{2}\right)^2. \qquad \square$$

A.2 A Deviation Inequality for Martingales

For the bound in (2.7), the Chernoff estimate methods from the previous chapter are no longer applicable because the distribution of the random variable considered is more difficult.

However, a similar idea combined with martingales leads to another class of upper bounds for deviations.

Definition A.2. *(i) A sequence $X_0, \dots, X_m$ of random variables in an arbitrary probability space is a* martingale *if for $0 \leqslant i < m$,*

$$\mathbb{E}[X_{i+1} \mid X_i, \dots, X_0] = X_i.$$

(ii) For $n \in \mathbb{Z}^+$ and $p \in (0,1)$, $\mathcal{G}(n,p)$ denotes the probability space of random graphs on the vertex set $[n]$ where each edge independently appears with probability p.

Moreover, let $\mathcal{G}_k$ be the partition of $\mathcal{G}(n,p)$ induced by identifying elements inducing the same graph on the first k vertices.

(iii) For a graph theoretic function f, the sequence $X_0, \dots, X_n$ of random variables on $\mathcal{G}(n,p)$ defined by

$$X_j \coloneqq \mathbb{E}[f \mid \mathcal{G}_j]$$

is the vertex exposure martingale *for f.*

Intuitively, the vertex exposure martingale for a function f is the sequence of conditional expectations of f given the information of the graph induced by the first i vertices. In particular, $X_1 = \mathbb{E}[f]$ is constant because a single vertex does not expose any edges and $X_m(H) = f(H)$ for every graph H because the whole graph has been exposed.

Proposition A.3. *The vertex exposure martingale is a martingale.*

Proof. Fix a graph theoretic function f and let $X_0, \dots, X_n$ be the vertex exposure martingale for f. We need to show that for $0 \leqslant i < n$,

$$\mathbb{E}[X_{i+1} \mid X_i, \dots, X_0] = X_i.$$

Plugging in the definition of the vertex exposure martingale and using the fact that the random variables $X_i, \dots, X_0$ induce the partition $\mathcal{G}_i$, we may rewrite this as

$$\mathbb{E}\Big[\mathbb{E}[f \mid \mathcal{G}_{i+1}] \;\Big|\; \mathcal{G}_i\Big] = \mathbb{E}[f \mid \mathcal{G}_i].$$

By definition of conditional probability, for all $\omega \in \mathcal{G}(n,p)$,

$$\mathbb{E}\Big[\mathbb{E}[f \mid \mathcal{G}_{i+1}] \;\Big|\; \mathcal{G}_i\Big](\omega) = \sum_{\substack{G \in \mathcal{G}_i,\\ \Pr[G] > 0}} \frac{\mathbb{E}[\mathbb{E}[f \mid \mathcal{G}_{i+1}] \cdot \mathbb{1}_G]}{\Pr[G]} \cdot \mathbb{1}_G(\omega).$$

Employing linearity of expectation, we get

$$\mathbb{E}[\mathbb{E}[f \mid \mathcal{G}_{i+1}] \cdot \mathbb{1}_G] = \sum_{\substack{G' \in \mathcal{G}_{i+1},\\ G' \subseteq G}} \mathbb{E}[\mathbb{E}[f \mid \mathcal{G}_{i+1}] \cdot \mathbb{1}_{G'}] = \sum_{\substack{G' \in \mathcal{G}_{i+1},\\ G' \subseteq G}} \mathbb{E}[f \cdot \mathbb{1}_{G'}] = \mathbb{E}[f \cdot \mathbb{1}_G],$$

and plugging this into the above, we conclude

$$\mathbb{E}\Big[\mathbb{E}[f \mid \mathcal{G}_{i+1}] \,\Big|\, \mathcal{G}_i\Big](\omega) = \sum_{\substack{G \in \mathcal{G}_i,\\ \Pr[G]>0}} \frac{\mathbb{E}[f \cdot \mathbb{1}_G]}{\Pr[G]} \cdot \mathbb{1}_G(\omega) = \mathbb{E}[f \mid \mathcal{G}_i]. \qquad \square$$

The following theorem provides the bounds we need in (2.7).

Theorem A.4 (Azuma's Inequality, [2, Thm. 7.2.1]). *Let $X_0, \ldots, X_m$ be a martingale with $X_0 = 0$ and $|X_{i+1} - X_i| \leqslant 1$ for all $i \in [m]$ and let $\lambda > 0$. Then*

$$\Pr\big[X_m > \lambda\sqrt{m}\big] \leqslant e^{-\frac{\lambda^2}{2}}.$$

Proof. The idea of the proof is again to fix $\alpha > 0$ (a concrete value will be chosen later) and use Markov's inequality in the form

$$\Pr\big[X_m > \lambda\sqrt{m}\big] = \Pr\Big[e^{\alpha X_m} > e^{\alpha\lambda\sqrt{m}}\Big] \leqslant \mathbb{E}\Big[e^{\alpha X_m}\Big] \cdot e^{-\alpha\lambda\sqrt{m}}. \tag{A.4}$$

To bound the expected value on the right-hand side, define $Y_i \coloneqq X_i - X_{i-1}$ for $i \in [m]$ and note that we can write

$$\mathbb{E}\Big[e^{\alpha X_m}\Big] = \mathbb{E}\left[\prod_{i=1}^{m} e^{\alpha Y_i}\right] = \mathbb{E}\left[\left(\prod_{i=1}^{m-1} e^{\alpha Y_i}\right) \cdot \mathbb{E}\Big[e^{\alpha Y_m} \,\Big|\, X_{m-1}, \ldots, X_0\Big]\right]. \tag{A.5}$$

because Y_i depends only on the values of $X_{m-1}, \ldots, X_0$ for $i < m$.

By assumption, we have $|Y_i| = |X_i - X_{i-1}| \leqslant 1$. Moreover,

$$\mathbb{E}[Y_i \mid X_{i-1}, \ldots, X_0] = \mathbb{E}[X_i \mid X_{i-1}, \ldots, X_0] - X_{i-1} = 0$$

by definition of a martingale. Define $g(x) \coloneqq \cosh(\alpha) + x\sinh(\alpha)$, then by convexity of $x \mapsto e^{\lambda x}$, the following inequality is true for any $x \in [-1, 1]$:

$$e^{\alpha x} \leqslant \frac{1-x}{2} \cdot e^{-\alpha} + \frac{1+x}{2} \cdot e^{\alpha} = g(x),$$

so for each $i \in [m]$, we have

$$\begin{aligned}
\mathbb{E}\Big[e^{\alpha Y_i} \,\Big|\, X_{i-1}, \ldots, X_0\Big] &\leqslant \mathbb{E}[g(Y_i) \mid X_{i-1}, \ldots, X_0] = \\
&= \cosh(\alpha) + \mathbb{E}[Y_i \mid X_{i-1}, \ldots, x_0] \cdot \sinh(\alpha) = \\
&= \cosh(\alpha) \leqslant \exp\frac{\alpha^2}{2},
\end{aligned} \tag{A.6}$$

where the last inequality follows from comparing the series expansions of exp and cosh around 0.

Plugging (A.6) into (A.5), we get

$$\mathbb{E}\Big[e^{\alpha X_m}\Big] \leqslant \mathbb{E}\left[\prod_{i=1}^{m-1} e^{\alpha Y_i}\right] \cdot e^{\frac{\alpha^2}{2}}.$$

Of course, the splitting in (A.5) can be iterated and the bound in (A.6) holds for all $i \in [m]$, so we get

$$\mathbb{E}\Big[e^{\alpha X_m}\Big] \leqslant \mathbb{E}\left[\prod_{i=1}^{m-1} e^{\alpha Y_i}\right] \cdot e^{\frac{\alpha^2}{2}} \leqslant \mathbb{E}\left[\prod_{i=1}^{m-2} e^{\alpha Y_i}\right] \cdot e^{\frac{2\alpha^2}{2}} \leqslant \ldots \leqslant e^{\frac{\alpha^2 m}{2}}.$$

Plugging the above into (A.4) and finally setting $\alpha = \lambda/\sqrt{m}$, we get the desired bound

$$\Pr\big[X_m > \lambda\sqrt{m}\big] \leqslant \mathbb{E}\Big[e^{\alpha X_m}\Big] \cdot e^{-\alpha\lambda\sqrt{m}} \leqslant e^{\frac{\alpha^2 m}{2} - \alpha\lambda\sqrt{m}} = e^{-\frac{\lambda^2}{2}}. \qquad \square$$

Appendix B

The Lambert W-Function

In this appendix, we proof Proposition 3.13 and Lemma 3.18. For easier reference, we restate both of them here as Proposition B.1 and Lemma B.2, respectively.

Proposition B.1. *For each* $\gamma \in (0, 1/4)$*, the equation*

$$1 = ye^{-\gamma y} \tag{B.1}$$

has a unique solution $y(\gamma) \in [1, 2]$ *and defines a function* $y : (0, 1/4) \to [1, 2]$.

Proof. The equation (B.1) is equivalent to $y = e^{\gamma y}$. For fixed γ, both the left- and right-hand side are continuous and strictly increasing. Note that

$$\begin{aligned} & y\big|_{y=1} = 1 \leqslant e^{\gamma} = e^{\gamma y}\big|_{y=1} \\ \text{and} \quad & y\big|_{y=2} = 2 \geqslant e^{\frac{1}{2}} \geqslant e^{2\gamma} = e^{\gamma y}\big|_{y=2}, \end{aligned}$$

so there is a unique solution $y(\gamma)$ in the interval $[1, 2]$ and y defines a function $y : (0, 1/4) \to [1, 2]$. □

Lemma B.2. *Let* $y : (0, 1/4) \to [1, 2]$ *be as in Proposition B.1. Then, for* $\gamma \in (0, 1/4)$*, we have*

$$y(\gamma) \leqslant 1 + 6\gamma.$$

To proof Lemma B.2, we introduce a related well-known function: The Lambert W-function.

Definition B.3. *The* Lambert W-function *is the inverse of the map*

$$f\colon\ \mathbb{R} \to \mathbb{R},\ W \mapsto We^{W}.$$

It can be observed that the function f has a minimum at $W = -1$ with value $\frac{1}{e}$, it is strictly decreasing on $(-\infty, -1)$ with $\lim_{x\to-\infty} f(x) = 0$ and strictly increasing on $(-1, \infty)$ with $\lim_{x\to\infty} f(x) = \infty$. Consequently, W has two branches:

$$W_0 : (-e^{-1}, \infty) \to (-1, \infty) \quad \text{and} \quad W_1 : (-e^{-1}, 0) \to (-\infty, -1).$$

This behaviour can also be observed in the plot of the Lambert W-function, see Figure B.1. The close relation of the map $y : (0, 1/4) \to [1, 2]$ to the Lambert W-function enables us to prove Lemma B.2.

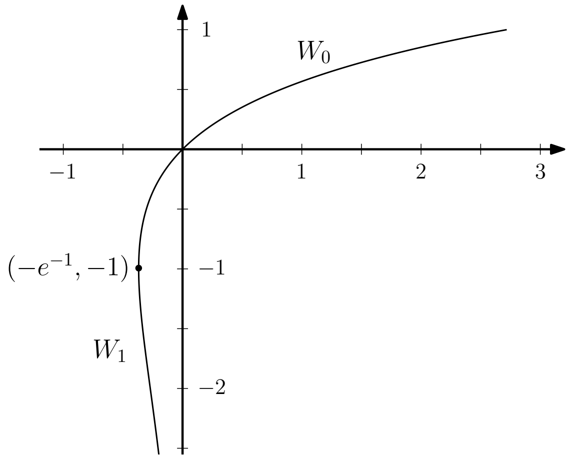

Figure B.1: The Lambert W-function.

Proof of Lemma B.2. For $\gamma \in (0, 1/4)$, let $z(\gamma) := \gamma y(\gamma)$ and note that (B.1) can be transformed as follows:

$$1 = y(\gamma)e^{-\gamma y(\gamma)} \iff -\gamma = -z(\gamma)e^{-z(\gamma)} \iff z(\gamma) = -W_0(-\gamma)$$
$$\iff y(\gamma) = -\frac{W_0(-\gamma)}{\gamma}.$$

Choosing the branch W_0 guarantees that $y(\gamma) \in [1, 2]$. Hence we need to prove

$$-\frac{W_0(-\gamma)}{\gamma} \leqslant 1 + 6\gamma, \quad \text{or equivalently} \quad W_0(x) \geqslant x - 6x^2$$

for $x \in (-1/4, 0)$, where we substituted $x = -\gamma$. Taking into account that $x = W_0(x)e^{W_0(x)}$, the latter can be rewritten as

$$W_0(x) \geqslant W_0(x)e^{W_0(x)} - 6W_0(x)^2 e^{2W_0(x)} \iff 1 \leqslant e^{W_0(x)} - 6W_0(x)e^{2W_0(x)}$$

because $W_0(x) < 0$ for $x < 0$. More precisely, we know that $W_0(x) \in [-1, 0]$, so it suffices to show that for all $t \in [-1, 0]$, we have

$$1 \leqslant e^t - 6te^{2t}.$$

The function $\varphi : t \mapsto e^t - 6te^{2t}$ on the right-hand side has value 1 at $t = 0$ and its derivative is

$$\varphi'(t) = e^t - 6e^{2t} - 12te^{2t} = e^t\Big(1 - 6(1 + 2t)e^t\Big).$$

It can be calculated that $\varphi(-1) = \varphi'(-1) \approx 1.18$ and $\varphi(0) = 1$, so it suffices to show that φ' changes its sign at most once on the interval $[-1, 0]$. Equivalently, it is sufficient to show that $e^t\Big(1 - 6(1 + 2t)e^t\Big) = 0$ has at most one solution in $[-1, 0]$. Note that

$$e^t\Big(1 - 6(1 + 2t)e^t\Big) = 0 \iff 1 = 6(1 + 2t)e^t,$$

and $t \mapsto 6(1 + 2t)e^t$ is strictly increasing on $[-3/2, \infty)$, hence there can be at most one solution. This proves the lemma. □

Appendix C

Notation

Graph theory notation

$G = (V, E)$ ... a graph G with vertices V and edges $E \subseteq \binom{V}{2}$

$e = \{u, v\}$ ... an edge e connecting vertices u and v.

$N_G(v)$ ... the neighbourhood of v in the graph G, also denoted $N(v)$ if unambiguous

$N_G^+(v)$ ... the inclusive neighbourhood of v in the graph G, i. e. $N_G^+(v) := N_G(v) \cup \{v\}$, also denoted $N^+(v)$ if unambiguous

$\deg_G(v)$ ... the degree $|N_G(v)|$ of v in the graph G, also denoted $\deg(v)$ if unambiguous

$G[U]$ ... the subgraph of $G = (V, E)$ induced by the vertices in $U \subseteq V$

$\chi(G)$ ... the chromatic number of the graph G

$\chi'(G)$ ... the edge chromatic number of the graph G

$\mathcal{G}(n, p)$ ... the space of random graphs on n vertices where each edge appears with probability p

Probability theory notation

$\overline{A}$ ... the complement of an event A

$\Pr[A]$ ... probability of an event A

$\Pr[A \mid \mathcal{P}]$ ... conditional probability of an event A given a partition $\mathcal{P}$ of the underlying probability space

$\Pr[X \mid X_1, \ldots, X_n]$ ... conditional probability of a random variable X given the partition of the underlying probability space generated by random variables $X_1, \ldots, X_n$

$\mathbb{E}[X]$ ... expected value of a random variable X

$\mathbb{E}[X \mid \mathcal{P}]$ ... conditional expectation of a random variable X given a partition $\mathcal{P}$ of the underlying probability space

$\mathbb{E}[X \mid X_1, \ldots, X_n]$ ... conditional expectation of a random variable X given the partition of the underlying probability space generated by random variables $X_1, \ldots, X_n$

Other notation

$[n]$ ... the set $\{1, 2, \ldots, n\}$ of the first n positive integers

$\mathbb{1}_X$ ... the indicator function of a set X, i. e. the function with value 1 on X and 0 else

log ... the logarithm with respect to the base e if not indicated otherwise

$(n)_d$ ... the falling factorial, $(n)_d = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$

$n!!$ ... the semifactorial, $n!! = n \cdot (n-2) \cdot \ldots \cdot 3 \cdot 1$ for odd n and $n!! = n \cdot (n-2) \cdot \ldots \cdot 4 \cdot 2$ for even n

$\binom{X}{k}$ ... the set of all subsets of X of cardinality k

Index

Bibliography

[1] Noga Alon, Michael Krivelevich, and Benny Sudakov. Coloring graphs with sparse neighborhoods. *Journal of Combinatorial Theory, Series B*, 77(1):73–82, 1999.

[2] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2011.

[3] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311 – 316, 1980.

[4] Reinhard Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, 2 edition, 2000.

[5] M. Drmota. Combinatorics and asymptotics on trees. *Cubo Journal of Mathematics*, 6(2), 2004.

[6] Paul Erdős and Joel Spencer. Lopsided Lovász local lemma and latin transversals. *Discrete Applied Mathematics: The Journal of Combinatorial Algorithms, Informatics and Computational Sciences*, 30(2):151–154, 1991.

[7] Louis Esperet and Aline Parreau. Acyclic edge-coloring using entropy compression. *European Journal of Combinatorics*, 34(6):1019–1027, 2013.

[8] Iosif Fiamčík. The acyclic chromatic class of a graph (in russian). *Mathematica Slovaca*, 28(2):139–145, 1978.

[9] A. Johansson. Asymptotic choice number for triangle free graphs. *DIMACS Technical Report*, pages 91 – 95, 1996.

[10] Richard M. Karp. Reducibility among combinatorial problems. In Raymond E. Miller and James W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103, 1972.

[11] Linyuan Lu and László Székely. Using lovász local lemma in the space of random injections. *Electronic Journal of Combinatorics*, 14(1):13, 2007.

[12] Linyuan Lu and László Székely. A new asymptotic enumeration technique: the lovász local lemma. *arXiv preprint, arXiv:0905.3983*, 2009.

[13] Austin Mohr. *Applications of the Lopsided Lovász Local Lemma Regarding Hypergraphs*. PhD thesis, University of South Carolina, 2013.

[14] Michael Molloy and Bruce Reed. Further algorithmic aspects of the local lemma. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, pages 524–529, New York, NY, USA, 1998. ACM.

[15] Michael Molloy and Bruce Reed. *Graph colouring and the probabilistic method*, volume 23. Springer Science & Business Media, 2002.

[16] Robin A. Moser. A constructive proof of the Lovász local lemma. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, STOC '09, pages 343–350, New York, NY, USA, 2009. ACM.

[17] Robin A. Moser and Gábor Tardos. A constructive proof of the general Lovász local lemma. *Journal of the ACM*, 57(2):11:1–11:15, 2010.

[18] Terence Tao. Moser's entropy compression argument. Terence Tao's wordpress blog, URL: `http://terrytao.wordpress.com/2009/08/05/mosers-entropy-compression-argument/`, August 2009. Accessed: June 15, 2015.

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

The Lovász Local Lemma: Selected Variants and Applications

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Nägele	Martin

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Zürich, 15. 07. 2015	Martin Nägele

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.