Derandomizing Concentration Inequalities with dependencies and their combinatorial applications

Dissertation
in fulfillment of the requirements for the degree of Dr. rer. nat.
of the Faculty of Mathematics and Natural Sciences
at Kiel University

submitted by

Mayank

Kiel, 2021

First Examiner: Prof. Dr. Anand Srivastav
Second Examiner: Prof. Dr. Sören Christensen

Date of Oral Examination: 02 July, 2021

Acknowledgements

I would like to take the opportunity to thank the people who have supported me through my PhD Thesis.

I would like to thank Revered Prof. Dr. Prem Saran Satsangi Sahab, Chairman, Advisory Committee on Education, Dayalbagh, Agra, India. It would have been impossible to complete my PhD without his supreme guidance.

I profusely thank Prof. Dr. Anand Srivastav for taking me through this journey.

I also would like to thank Jan Schiemann, Christian Schielke, Jan Geest, Volkmar Sauerland, Mourad El Ouali (Now in Morocco) and Axel Wedemeyer. I had extremely nice mathematical and non-mathematical discussions with them.

Lastly, I would like to thank my parents, my wife Sugandha and my two lovely daughters Radha and Ragini for all their support.

Contents

Deutsche Zusammenfassung

Sowohl in der Kombinatorik als auch bei dem Entwurf und der Analyse von randomisierten Algorithmen für kombinatorische Optimierungsprobleme wird vielfach die berühmte Bounded-Differences-Inequality von C. McDiarmid (1989), die auf der Martingal-Ungleichung von K. Azuma (1967) beruht, verwendet, um zu zeigen, dass eine echt positive Erfolgswahrscheinlichkeit vorliegt. In dem Fall, wo Summen von unabhängigen Zufallsvariablen vorliegen, kommt man mit den Ungleichungen von Chernoff (1952) und Hoeffding (1964) aus und kann sogar eine effiziente Derandomisierung erreichen, also das Ereignis in deterministischer, polynomieller Zeit konstruieren (Srivastav und Stangier 1996), was natürlich wünschenswert ist. Aus dem probabilistischen Existenzresultat oder dem randomisierten Algorithmus erhält man so eine wirkliche Konstruktion der gesuchten kombinatorischen Struktur oder einen effizienten, deterministischen Algorithmus.
Dieses Derandomisierungsproblem war bislang für die Bounded-Differences-Inequality von C. McDiarmid offen. Das Hauptresultat in Kapitel 3 ist eine effiziente Derandomisierung der Bounded-Differences-Inequality, wobei die Zeit zur Berechnung von bedingten Erwartungswerten der Zielfunktion in die Komplexität eingeht. Damit wird das genannte offene Problem erstmalig gelöst. In den nachfolgenden Kapiteln 4 bis 7 wird die Stärke und Reichweite dieser Derandomisierung demonstriert.
In Kapitel 5 derandomisieren wir die Zufallsstrategie des Makers in dem Maker-Breaker-Subgraphenspiel, die in der für das Gebiet grundlegenden Arbeit von Bednarska und Luczak (2000) gegeben und mit der Konzentrationsungleichung von Janson, Luczak und Rucinski analysiert wurde. Da wir aber die Bounded-Differences-Inequality verwenden, ist es erforderlich, einen neuen Beweis der Existenz von Subgraphen in G(n,M)-Zufallsgraphen zu geben, was in Kapitel 4 geleistet wird.
In Kapitel 6 derandomisieren wir den zweistufigen randomisierten Algorithmus für das Set-Multicover Problem von El Ouali, Munstermann und Srivastav (2014).
In Kapitel 7 zeigen wir, dass sich der Algorithmus von Bansal, Caprara und Sviridenko (2009) für das multidimensionale Bin-Packing Problem mit unserer derandomisierten Form der Bounded-Differences-Inequality elegant derandomisieren lässt, während die genannten Autoren eine Potentialfunktion, die auf das Problem adaptiert ist, verwendet haben, was zu einer recht aufwändigen Analyse führte.
In Kapitel 8 wird der Algorithmus von Ahuja und Srivastav (2002) für das Constrained Hypergraph Coloring Problem, das sowohl das Property B Problem für die nicht-monochromatische 2-Färbung von Hypergraphen, als auch das multidimensionale Bin-Packing Problem verallgemeinert, derandomisiert, aber mit vorheriger Analyse mit der Bounded-Differences-Inequality statt dem Lovasz-Local-Lemma.

In Kapitel 9 wenden wir uns der Konzentrationsungleichung von Janson (1994) zu, wo beim Vorliegen von Summen von Zufallsvariablen, die nicht unabhängig, aber teilweise abhängig sind, oder anders gesagt, in bestimmten Gruppen unabhängig sind, die die bekannte Ungleichung von Hoeffding (1964) verallgemeinert. Wir zeigen eine ähnliche Ungleichung unter partieller Dependenz der zugrundeliegenden Zufallsvariablen, die die bekannte Konzentrationsungleichung von Alon und Spencer (1991) verallgemeinert.
In Kapitel 10 derandomisieren wir die Ungleichung von Alon und Spencer. Die Derandomisierung unserer verallgemeinerten Alon-Spencer-Ungleichung unter partiellen Dependenzen verbleibt ein interessantes, offenes Problem.

Summary

Both in combinatorics and design and analysis of randomized algorithms for combinatorial optimization problems, we often use the famous bounded differences inequality by C. McDiarmid (1989), which is based on the martingale inequality by K. Azuma (1967), to show positive probability of success. In the case of sum of independent random variables, the inequalities of Chernoff (1952) and Hoeffding (1964) can be used and can be efficiently derandomized, i.e. we can construct the required event in deterministic, polynomial time (Srivastav and Stangier 1996). With such an algorithm one can construct the sought combinatorial structure or design an efficient deterministic algorithm from the probabilistic existentce result or the randomized algorithm.

The derandomization of C. McDiarmid's bounded differences inequality was an open problem. The main result in Chapter 3 is an efficient derandomization of the bounded differences inequality, with the time required to compute the conditional expectation of the objective function being part of the complexity. The following chapters 4 through 7 demonstrate the generality and power of the derandomization framework developed in Chapter 3.

In Chapter 5, we derandomize the Maker's random strategy in the Maker-Breaker subgraph game given by Bednarska and Luczak (2000), which is fundamental for the field, and analyzed with the concentration inequality of Janson, Luczak and Rucinski. But since we use the bounded differences inequality, it is necessary to give a new proof of the existence of subgraphs in G(n,M)-random graphs (Chapter 4).

In Chapter 6, we derandomize the two-stage randomized algorithm for the set-multicover problem by El Ouali, Munstermann and Srivastav (2014).

In Chapter 7, we show that the algorithm of Bansal, Caprara and Sviridenko (2009) for the multidimensional bin packing problem can be elegantly derandomized with our derandomization framework of bounded differences inequality, while the authors use a potential function based approach, leading to a rather complex analysis.

In Chapter 8, we analyze the constrained hypergraph coloring problem given in Ahuja and Srivastav (2002), which generalizes both the property B problem for the non-monochromatic 2-coloring of hypergraphs and the multidimensional bin packing problem using the bounded differences inequality instead of the Lovasz local lemma. We also derandomize the algorithm using our framework.

In Chapter 9, we turn to the generalization of the well-known concentration inequality of Hoeffding (1964) by Janson (1994), to sums of random variables, that are not independent, but are partially dependent, or in other words, are independent in certain groups. Assuming the same dependency structure as in Janson (1994), we generalize the well-known concentration inequality of Alon and Spencer (1991).

In Chapter 10, we derandomize the inequality of Alon and Spencer. The derandomization of our generalized Alon-Spencer inequality under partial dependencies remains an interesting, open problem.

Chapter 1

Introduction

Derandomization is the task of turning a probabilistic existence result or a randomized algorithm into a deterministic, polynomial-time algorithm, preferably achieving nearly the same guarantee. With the evident power of randomized algorithms, derandomization became an important task in algorithmic discrete mathematics. This is desirable, since deterministic algorithms give hard guarantees on performance and running times.

The Problem. The derandomization problem considered in this thesis is as follows:

Derandomization Problem Let F be a finite set and let $\Omega = F^n$ be a probability space with probability measure $\mathbb{P}$. Let $E_1, \ldots, E_m$ be events in Ω. We assume that $\mathbb{P}(\bigcap_{i=1}^m E_i) \geqslant \delta$ for some $\delta > 0$. Hence $\bigcap_{i=1}^m E_i$ is not empty and derandomization is the task of constructing a point in $\bigcap_{i=1}^m E_i$ in time polynomial in n, m and $\ln(1/\delta)$.

If the probability of the event E_i^c, where the superscript c denotes the complementary event, can be bounded by Chernoff bounds or the Lovász Local Lemma, the derandomization problem has been solved and many applications to combinatorial packing, covering and coloring problems have been given e.g. Raghavan [Rag88], Spencer [Spe94], Srivastav, Stangier [SS96], Beck [Bec91], Srinivasan [Sri99], Moser, Tardos [MT10]). For obvious reasons let us call such a procedure, the derandomized (algorithmic) version of underlying large deviation bound. Further highlights in derandomization are derandomized counterparts of semidefinite programming based algorithms, e.g. for Max Cut by Mahajan and Ramesh [MR99] and for Bansal's [Ban10] low-discrepancy computing algorithm by Bansal and Spencer [BS11]. In computational geometry derandomization has been advanced by the work of Jiří Matoušek, e.g. [Mat96].

1.1 Derandomizing the Generalized Bounded Differences Inquality

Erdős and Selfridge [ES73], Beck and Fiala [BF81] and Spencer [Spe77] suggested for derandomization a very general technique, *the conditional probability method*. Unfortunately, in many cases conditional probabilities under consideration cannot be computed efficiently. Here a milestone for algorithmic progress has been the introduction of so called pessimistic estimators by Raghavan [Rag88], which are computable upper bounds on the conditional probabilities sharing the properties of them. One key fact for the success of the conditional probability method is that pessimistic estimators can be constructed, whenever linear objective functions are involved, because the proofs of Chernoff and Hoeffding type inequalities deliver the right pessimistic estimators. In fact, for a polynomial-time implementation of the method, approximations of these estimators by Taylor polynomials must be used, too e.g. [SS96].

The efficient derandomization has not been known, when the event probabilities are estimated by the famous and frequently applied bounded differences inequality (BDI) and its generalized form due to C. McDiarmid [McD89] which is based on the martingale inequality of K. Azuma [Azu67]. A. Srivastav and P. Stangier [SS93] could derandomize the Azuma inequality in the very special application to quadratic lattice approximation, but this is quite far from the generalized bounded differences inequality. Also, merging of different derandomized algorithms was not investigated. In consequence, largely for all randomized algorithms using the BDI in its analysis, derandomized counterparts are not known. To the best of our knowledge the only exception is the derandomization of the randomized algorithm of Bansal et al. [BCS09] for the multidimensional bin packing problem, where a potential function approach was used in this specific context. The potential function given there has a problem adapted form, in particular it is *set oblivious*, and thus this approach cannot be lifted to the general BDI situation, where we only know that the function under consideration is Lipschitz bounded.

On the other hand, since the BDI has a wide range of applications, for example to the analysis of hybrid randomized algorithms consisting of several subroutines, it is desirable to derandomize such algorithms with an easily applicable derandomized BDI framework in a short and elegant way without setting up problem adapted potential functions, whose construction is non-trivial and differs from case to case.

In order to formulate previous and our results in a mathematical way, we need some technical notions. A common setting is the following: let $X_1, \dots, X_n$ be real-valued random variables. For $i = 1, \dots, m$, let $\psi_i = \psi_i(X_1, X_2, \cdots, X_n)$ be a function of X_j's for $1 \leqslant j \leqslant n$. Given rational parameters $\lambda_i > 0$, denote for $i = 1, \dots, m$ by E_i either the event "$\psi_i \leqslant \mathbb{E}(\psi_i) + \lambda_i$" or the event "$\psi_i \geqslant \mathbb{E}(\psi_i) - \lambda_i$". If ψ_i is a linear function or if it has bounded range or Lipschitz bounded in every coordinate, thus satisfies the assumptions of the bounded differences inequality (BDI), the various types of large deviation bounds due to Chernoff [Che52], Hoeffding [Hoe63], Azuma [Azu67] and McDiarmid [McD89] can be summarized by the inequality

$$\mathbb{P}(E_i^c) \leqslant e^{-t_i \lambda_i} \mathbb{E}(e^{t_i \psi_i}) \leqslant f(\lambda_i). \tag{1.1}$$

An optimal choice of the parameter $t_i > 0$ gives the sharpest possible upper bound $f(\lambda_i)$. If $\sum_{i=1}^m f(\lambda_i) < 1 - \delta$ for some $0 < \delta < 1$, then $\mathbb{P}(\bigcap_{i=1}^m E_i) \geqslant \delta > 0$, hence $\bigcap_{i=1}^m E_i$ is not empty and derandomization is the task of constructing a point in $\bigcap_{i=1}^m E_i$ efficiently, i.e. in polynomial time in $n, m, \ln(1/\delta)$.

In settings where the ψ_i's have bounded range as in generalized BDI or are Lipschitz bounded as in the BDI, not even appropriate pessimistic estimators are known.

In chapter 3, we resolve the derandomization problem for Azuma's inequality and the generalized BDI (consequently BDI) in the probabilistic setting described above. The only assumptions we must make is the boundedness and the polynomial-time computability of the functions ψ_i, their expectations $\mathbb{E}(\psi_i)$ and their conditional expectations. We feel that these are natural assumptions. In fact, in the setting of randomized algorithms with a linear objective function these assumptions are either automatically satisfied or can easily be proved (see for example our applications).

The hard part of the derandomization is the construction of suitable pessimistic estimators for the generealized BDI. Since the generalized BDI is a kind of corollary to Azuma's martingale bound, we turn the proof of Azuma's inequality into an algorithmic form, and this means that we extract the appropriate pessimistic estimator by passing through the proof. Here an important aspect is to reduce for some event C, the computation of the conditional probabilities of the moment generating function $\mathbb{E}(e^{\psi_i}|C)$, which is not known, to the computation of the conditional probabilities $\mathbb{E}(\psi_i|C)$, which by our assumption can be done in polynomial time. We show this by using properties of a finite martingale. Note that these arguments do not apply for non-finite probability measures. Now we have found candidates for the pessimistic estimators, but they are transcendental real-valued functions. Therefore they cannot be computed in polynomial time, which is a pre-requisite for any polynomial-time derandomization. Of course, one can approximate them, but then the precision and time for the approximation has to be taken into account, and even more the concern is that these approximations might not be pessimistic estimators. The first problem we solve by fast Taylor polynomial approximations, while the second issue is resolved by an additive perturbation of the approximations ensuring that the perturbed functions become pessimistic estimators.

Finally, we combined pessimistic estimators for the BDI with pessimistic estimators for other concentration inequalities. For this purpose we generalize the notion of the pessimistic estimator to so called weak pessimistic estimators introducing a kind of convexity of a family of functions. This appears at the first moment only as a technical and mathematical generalization, but it turns out to be an easily applicable framework, and pays off in later constructions.

1.2 The Subgraph Containment Problem

In chapter 4, we upper bound the probability of non-existence of a fixed subgraph in larger random graph using the generalized BDI framework. Let $n \in \mathbb{N}$, $M \in \left[\binom{n}{2}\right]$ and $p \in [0,1]$. $\mathcal{G}_{n,M}$ is the set of graphs chosen uniformly at random from the family of all subgraphs of K_n with exactly M edges and n nodes and $\mathcal{G}_{n,p}$ denotes the set of graphs obtained by adding edges of K_n with probability p, independently for each edge. We identify a graph in the model $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$ with the model itself. So we may say that a fixed graph H is a subgraph of $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$ meaning that H occurs as a subgraph in a random graph from $\mathcal{G}_{n,M}$ resp.

$\mathcal{G}_{n,p}$.

For a fixed graph G, the well known subgraph containment problem can be stated as : Does $\mathcal{G}_{n,M}$ contain a copy of G ?

Let e_G resp. v_G be the number of edges resp. nodes of G. Define

$$d(G) := \frac{e_G - 1}{v_G - 2} \text{ and } m(G) = \max\left\{d(H) : H \subseteq G, v_H \geqslant 3\right\}$$

$m(G)$ is a measure of graph density frequently appearing in the theory of random graphs [JŁR11]. A graph G is called *strictly K_2-balanced* if $d(H) < d(G)$, where $H \subset G, v_H \geqslant 3$.

The following theorem mentioned in Janson, Łuczak and Ruckinski [JŁR11] is the $\mathcal{G}_{n,M}$ counterpart of Theorem 4.2 proved by Janson, Rucinski and Łuczack [JŁR90], the best known upper bound on the probability of non-existence of fixed G in $\mathcal{G}_{n,p}$.

Theorem 1.1. For every fixed graph G containing a cycle, there exists a constant $c'_1 > 0$ and $n'_1 \in \mathbb{N}$ such that for every $n \geqslant n'_1, n \in \mathbb{N}$, and $M = n^{2-\frac{1}{m(G)}}$, we have

$$\mathbb{P}(\mathcal{G}_{n,M} \not\supset G) \leqslant e^{-c'_1 M} \tag{1.2}$$

We prove an analogue of Theorem 4.3 with the generalized bounded differences inequality. Given the derandomized framework for generalized BDI in chapter 3, we derandomize Maker's random strategy in chapter 5.

1.3 Derandomizing Maker's Strategy for the Maker-Breaker Subgraph Game

The Maker-Breaker subgraph game $\mathbf{G}(G, n, b)$, where $b, n \in \mathbb{N}$ is played on the complete graph K_n on n vertices by two players, Maker and Breaker. In each round of the game, Maker chooses an edge of K_n, which has not been claimed previously and breaker responds by selecting at most b edges from K_n. The variable b denotes the *bias* of the game. The game ends if each of $\binom{n}{2}$ edges of K_n is claimed by either of the players. If the subgraph constructed during the game by the Maker contains a copy of G, then he wins, otherwise he looses and Breaker wins the game. The game $\mathbf{G}(G, n, b)$ is a special case of positional games on graphs with G as a fixed graph.

In their outstanding work, Bednarska and Łuczack [BŁ00] used results of the appearance of a fixed graph in a random graph to show that for the game $\mathbf{G}(G, n, b)$, a random strategy for Maker is asymptotically optimal in n, but to find the right matching constants is a major open problem in this area. They modified the above game so that Maker cannot see Breaker moves, but Breaker has all the information about the moves of the Maker. They proved that even with this weak strategy of Maker, where Maker chooses his edges uniformly at random among all edges, which have not been claimed by him so far, Maker wins with a probability of at least $\frac{1}{3}$. The main result in [BŁ00] is the following theorem:

Theorem 1.2. For every graph G which contains at least 3 non-isolated vertices, there exist positive constants c_0, n_0 such that for every $n \geqslant n_0$ and $b \leqslant c_0 n^{1/m(G)}$, $n, b \in \mathbb{N}$, Maker has a winning strategy for the game $\mathsf{G}(G, n, b)$.

In this chaper 5, we present a deterministic Maker strategy derandomizing the existential result of Bednarska and Łuczack using the subgraph containment result proved in chapter 4 and derandomization framework of generalized BDI in the chapter 3.

1.4 Derandomized Hybrid Algorithm for the Set-Multicover Problem

In chapter 6, we study a generalized version of the set cover problem, SET b-MULTICOVER. The problem can be naturally expressed in the context of hypergraphs. A hypergraph $\mathcal{H} = (V, \mathcal{E})$ consists of a finite set V and a set $\mathcal{E}$ of subsets of V. We call elements of V as vertices and the elements of $\mathcal{E}$ (hyper-)edges. We fix $n := |V|$, $m := |\mathcal{E}|$. For $b \in \mathbb{N}$, a set b-multicover in $\mathcal{H}$ is a set of edges $C \subseteq \mathcal{E}$ such that every vertex in V belongs to at least b edges in C. SET b-MULTICOVER is the problem of finding a set b-multicover of minimum cardinality.

The *edge size* is the cardinality of the edge. Let l be the maximum edge size and let Δ be the maximum vertex degree, where the degree of a vertex is the number of edges containing that vertex. Define $\delta := \Delta - b + 1$.

We derandomize the presently best set b-multicover approximation algorithm, $b \in \mathbb{N}$, published in [EMS14]. It is a challenging candidate, because here randomized rounding is combined with a repairing routine, and thus it is an inherently hybrid algorithm. Furthermore, the various event probabilities in the analysis of this randomized algorithm are estimated not only by the BDI, but also with some kind of Chernoff bounds, the Angluin-Valiant inequality. So we must use here the full power of our derandomized BDI joined with pessimistic estimators for the Angluin-Valiant inequality. Since all constructions are time optimized, the total derandomization has complexity of only $\mathcal{O}(m^2 n)$.

1.5 Derandomized Algorithm for the Multidimensional-Bin Packing Problem

In theoritical computer science, the bin packing problem had profound impact on the field of approximation algorithms. In the *classical bin packing* problem, assume we have n items, $\{i_1, i_2, \cdots, i_n\}$, where the size of each item $i_k \in (0, 1]$ for all $k \in [n]$ and bins have capacity 1. The problem is to partition items into minimal number of subsets such that items in each subset can be packed into a bin. The problem was proved to be $\mathcal{NP}$-Hard by Garey and Johnson in 1979 [GJ79]. Bin packing problem can be naturally extended to higher dimensions, namely *vector bin packing* problem and *geometric bin packing* problem.

In the d-dimensional vector bin packing problem, each bin and item has d dimensions and we need to partition the items such that we can pack them in minimum number of bins. For example, we can think of each job as an item with CPU, RAM, disk, network requirements etc. as its d dimensions. The goal is to assign all the the jobs to minimum number of computing devices, which can be considered as d-dimensional bins with bounded amount of d resources required by the jobs.

Bansal et al. [BCS09] gave the *Round and Approx* framework (R&A) and used it to construct algorithms for two-dimensional geometric bin packing problem and vector bin packing problems. In their paper, Bansal et al. derandomized the R&A framework using the potential function approach.

In chapter 7, we demonstrate that the derandomization of Bansal's algorithm follows as a straightforward corollary from our derandomized BDI framework which we developed in the chapter 3.

1.6 Derandomized Approximation of Constrained Hypergraph Coloring (CHC)

Ahuja and Srivastav [AS02] introduced the *constrained hypergraph coloring problem*(CHC) as the generalization of the property B hypergraph coloring problem. It also models special cases of multidimensional bin packing (MDBP) problem and the resource constrained scheduling (RCS) problem.

Consider the hypergraph $\mathcal{H} = (V, \mathcal{E})$ with $V = \{1, 2, \cdots, n\}, \mathcal{E} = \{E_1, \ldots, E_m\}$, and $l = \max_{1 \leqslant i \leqslant m} |E_i|$. Let $b = (b_1, b_2, \ldots, b_m)^t$ be a vector. The problem is to partition the vertex set into a minimum number of sets such that there are at most b_i vertices in E_i of any partition set for all $i \in [m]$. We may color the vertices of each partition set with one color and call the set *color class.*

CHC reduces to extensively studied well known problems based on the underlying combinatorial structure and value of b_i. For a simple graph and $b_i = 1$ for all $i \in [m]$, CHC reduces to the graph coloring problem. The hypergraph $\mathcal{H}$ is said to be c-colorable iff there is a function $V \to \{1, 2, \cdots, c\}$ such that no edge is monochromatic. Hypergraph 2-colorability is the famous property B [Erd63]. For a hypergraph $\mathcal{H}$ with property B and $b_i = |E_i| - 1$ for all $i \in [m]$, CHC is equivalent to the problem of finding a non-monochromatic 2-coloring of $\mathcal{H}$.

In chapter 8, we have analyzed the CHC problem using independent BDI and designed a randomized polynomial time algorithm which outputs a solution for the CHC problem with $\lceil (1+\epsilon)\, C \rceil$ colors with probability at least $1 - \frac{1}{c_1}$, $c_1 > 0$ a constant, provided that $b_i \geqslant \frac{1+\epsilon}{\epsilon}\sqrt{\frac{l \ln(c_1 mn)}{2}}$ for all $i \in [m]$, where C is the optimal solution for linear program for constrained hypergraph coloring problem. We have also derandomized the algorithm using our derandomized BDI framework which we developed in the chapter 3.

1.7 Concentration Bounds with Partially Dependent Random Variables

Svante Janson [Jan04] extended the well known Hoeffding's bound (Theorem 2.1) for sums of independent random variables to obtain concentration bounds for sums of dependent random variables with a defined dependency structure. The method is based on breaking the sum of non-independent random variables into sums of independent random variables. Janson applied the framework to U-Statistics, random strings and random graphs. In chapter 9, we extend the Alon-Spencer [AS04] concentration bound for sums of independent random variables, which generalizes Hoeffding's bound, to obtain concentration bounds for the sum of dependent random variables with similar dependency structure as defined in [Jan04].

1.8 Derandomizing Alon-Spencer Concentration Inequality

In Chapter 9, we have stated the Alon-Spencer bound (Theorem 9.2). In this chapter, we derandomize the Alon-Spencer inequality assuming all random variables are mutually independent.

Chapter 2

Concentration Inequalities and Proofs

The Probabilistic Method has developed intensively into one of the most powerful tools widely used in Combinatorics. The method was developed majorly because of increased used of randomness in Theoretical Computer Science, a field which has resulted in many combinatorial problems. The goal of any application of the Probabilistic Method is to show that “good events” occur with positive probability or equivalently "bad events" occur with probability less than 1. Frequently we tend to bound the tail distribution, the probability that a random variable takes on values far from the expectation. In the context of analysis of algorithms, these bounds are tools for estimating the failure probability of the algorithms. Concentration bounds for tail distribution hence are the most important tools for showing that probability of the event is very small, not merely less than 1.

In this chapter, we cover Hoeffding bounds in case we have independent random variables, the Azuma-Hoeffding inequality and its closely related generalized independent bounded differences bounds. The Chernoff bound is used for tail distribution of a sum of independent $0-1$ random variables. Hoeffding's bound extends the Chernoff bound technique to bounded independent random variables.

2.1 The Hoeffding Bound

Theorem 2.1. (Hoeffding's Bound) Let $Z_1, Z_2, \cdots, Z_n$ be independent random variables with $a_i \leqslant Z_i \leqslant b_i$ for each i, for suitable constants a_i, b_i. Let $S_n = \sum_{i=1}^n Z_i$ and let $\mathbb{E}[S_n] = \mu$. Then for any $t \geqslant 0$,

$$\mathbb{P}(|S_n - \mu| \geqslant t) \leqslant 2e^{-2t^2/\sum_{i=1}^n (b_i - a_i)^2} \tag{2.1}$$

We need the below lemma to prove the theorem.

Lemma 2.2. Let the random variable Z satisfy $\mathbb{E}[Z] = 0$ and $a \leqslant Z \leqslant b$, where a and b are constants. Then for any $h > 0$,

$$\mathbb{E}\left[e^{hZ}\right] \leqslant e^{\frac{1}{8}h^2(b-a)^2}$$

Proof. Since e^{hz} is a convex function of z, for $a \leqslant z \leqslant b$,

$$e^{hz} \leqslant \frac{z-a}{b-a}e^{hb} + \frac{b-z}{b-a}e^{ha}$$

Now we consider $e^{\lambda Z}$ and take expectations on both sides.

$$\begin{aligned}
\mathbb{E}\left[e^{hZ}\right] &\leqslant \frac{b}{b-a}e^{ha} - \frac{a}{b-a}e^{hb} \text{ (as } \mathbb{E}\left[Z\right] = 0) \\
&= (1-p)e^{-py} + pe^{(1-p)y} \\
&= e^{-py}(1-p+pe^{y}) = e^{f(y)}
\end{aligned}$$

where $p = -a/(b-a)$, $y = (b-a)h$ and $f(z) = -pz + \ln(1-p+pe^{z})$.

$$f'(z) = -p + \frac{pe^{z}}{(1-p)+pe^{z}} = -p + \frac{p}{p+(1-p)e^{-z}}$$

and

$$f''(z) = \frac{p(1-p)e^{-z}}{(p+(1-p)e^{-z})^2} \leqslant \frac{1}{4} \tag{2.2}$$

(since the geomteric mean is at most the arithmetic mean). Also $f(0) = f'(0) = 0$, and hence by Taylor's theorem, for any $z > 0$ there is a $z' \in \{0, z\}$ such that

$$f(z) = f(0) + zf'(0) + \frac{1}{2}z^2 f''(z') \tag{2.3}$$

Hence

$$f(y) \leqslant \frac{1}{8}y^2 = \frac{1}{8}(b-a)^2h^2 \text{ (by 2.2) },$$

which gives the desired inequality. □

Now we give the proof of Theorem 2.1.

Proof. For $h > 0$, we have

$$\begin{aligned}
\mathbb{E}\left[e^{h(S_n-\mu)}\right] &= \mathbb{E}\left[\prod_{i=1}^{n} e^{h(Z_i-\mathbb{E}[Z_i])}\right] \\
&= \prod_{i=1}^{n} \mathbb{E}\left[e^{h(Z_i-\mathbb{E}[Z_i])}\right] \text{ (by Independence of } Z_i,\ i \in [n]) \\
&\leqslant e^{\frac{1}{8}h^2\sum_{i=1}^{n}(b_i-a_i)^2} \text{ (by Lemma 2.2)}
\end{aligned}$$

By Markov's inequality,

$$\begin{aligned}\mathbb{P}(S_n - \mu \geqslant t) &\leqslant e^{-ht}\mathbb{E}\left[e^{h(S_n-\mu)}\right] \\ &\leqslant e^{-ht+\frac{1}{8}h^2\sum_{i=1}^n (b_i-a_i)^2}\end{aligned}$$

To minimize the RHS, set $h = 4t/\sum_{i=1}^n (b_i - a_i)^2$ to obtain,

$$\mathbb{P}(S_n - \mu \geqslant t) \leqslant e^{-2t^2/\sum_{i=1}^n (b_i-a_i)^2}$$

Replacing Z_i by $-Z_i$ to obtain

$$\mathbb{P}(S_n - \mu \leqslant -t) \leqslant e^{-2t^2/\sum_{i=1}^n (b_i-a_i)^2},$$

and we have completed the proof.

In one of our applications, for the sum of independent $\{0,1\}$-random variables, we also use the large deviation inequality due to Angluin and Valiant:

Theorem 2.3. (Angluin-Valiant Bound)[McD98] Let $Z_1, \ldots, Z_n$ be independent $\{0,1\}$-random variables. Let $Z = \sum_{i=1}^n Z_i$. For every $\beta > 0$ it holds that

(i) $\mathbb{P}(Z \geqslant (1+\beta) \cdot \mathbb{E}[Z]) \leqslant \exp\left(-\frac{\beta^2\mathbb{E}[Z]}{3}\right)$

(ii) $\mathbb{P}(Z \leqslant (1-\beta) \cdot \mathbb{E}[Z]) \leqslant \exp\left(-\frac{\beta^2\mathbb{E}[Z]}{2}\right)$

□

2.2 Preliminary Facts about Martingales

We will frequently apply Azuma's inequality throughout this thesis. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and $\mathbb{Q}$ a second probability measure on Ω absolutely continuous with respect to $\mathbb{P}$, i.e. $\mathbb{P}(A) = 0$ implies $\mathbb{Q}(A) = 0$ for all $A \in \Sigma$. Then $\mathbb{Q}$ possesses a density function with respect to $\mathbb{P}$. This is just the celebrated Radon-Nikodym theorem [Nik30]:

Theorem 2.4. (Radon-Nikodym)
There exists an integrable function $X : \Omega \to \mathbb{R}$, $X \geqslant 0$ such that

$$\mathbb{Q}(A) = \int_A X \, d\mathbb{P}$$

for all $A \in \Sigma$.

Let $\Sigma_0 \subseteq \Sigma$ be a sub σ-algebra of Σ and let $X : \Omega \to \mathbb{R}$ be any $\mathbb{P}$-integrable function.

Definition 2.5. Let $Y : \Omega \to \mathbb{R}$ be a Σ_0-measurable function satisfying

$$\int_A Y \, d\mathbb{P} = \int_A X \, d\mathbb{P}$$

for all $A \in \Sigma_0$. Then Y is called the conditional expectation of X subject to Σ_0, and is denoted by $\mathbb{E}[X|\Sigma_0]$.

Note that for every non-negative $\mathbb{P}$-integrable function X and every $\Sigma_0 \subseteq \Sigma$ the conditional expectation $\mathbb{E}[X|\Sigma_0]$ exists, is unique, integrable and non-negative almost surely. We shall need some useful properties of conditional expectations. The following theorem can be found in any textbook of probability theory, for example in [BB96].

Theorem 2.6. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, $\Sigma_0 \subseteq \Sigma$ a sub-σ-algebra and X, Y non-negative, integrable random variables. Then the following properties of $\mathbb{E}[X|\Sigma_0]$ are true

- (i) $\mathbb{E}[X|\Sigma_0] = \mathbb{E}[X]$ (X and Σ_0 are independent)
- (ii) If X is Σ_0-measurable, then $\mathbb{E}[X|\Sigma_0] = X$ a.s.
- (iii) If $X = Y$ a.s., then $\mathbb{E}[X|\Sigma_0] = \mathbb{E}[Y|\Sigma_0]$ a.s.
- (iv) If $X = \alpha =$ const., then $\mathbb{E}[X|\Sigma_0] = \alpha$ a.s.
- (v) $\mathbb{E}[\alpha X + \beta Y|\Sigma_0] = \alpha\mathbb{E}[X|\Sigma_0] + \beta\mathbb{E}[Y|\Sigma_0]$ a.s. for $\alpha, \beta \in \mathbb{R}$.
- (vi) If $X \leqslant Y$ a.s. then $\mathbb{E}[X|\Sigma_0] \leqslant \mathbb{E}[Y|\Sigma_0]$ a.s.
- (vii) If $Z : \Omega \to \mathbb{R}$ is Σ_0-measurable and bounded, then $\mathbb{E}[ZX|\Sigma_0] = Z\mathbb{E}[X|\Sigma_0]$.
- $(viii)$ If $\mathcal{B} = \{B,\ B \in \Sigma_0\}$ is a partition of Ω and $\mathcal{B}$ generates Σ_0, then

$$\mathbb{E}[X|\Sigma_0] = \sum_{B \in \mathcal{B}} \mathbb{E}[X|B] \, 1_B,$$

where 1_B is the indicator function for the set B and $\mathbb{E}[X|B]$ is the average $\mathbb{E}[X|B] = \frac{1}{\mathbb{P}(B)} \int_B X \, d\mathbb{P}$.

Let Σ be a σ-algebra over Ω. A sequence $\left(\Sigma_k\right)_{k=0}^{\infty}$ of sub-σ-algebras of Σ is called a filtration, if $\Sigma_k \subseteq \Sigma_{k+1}$ for all k.

Definition 2.7. (Martingale)
Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, $\left(\Sigma_k\right)_{k=0}^{\infty}$ a filtration and $\left(X_k\right)_{k=0}^{\infty}$ a sequence of integrable random variables.

(i) Then $\left(X_k\right)_{k=0}^{\infty}$ is called a martingale, if

$$\mathbb{E}\left[X_{k+1}|\Sigma_k\right] = X_k$$

for all $k = 0, 1, \ldots,$.
(ii) The sequence $\left(Y_k\right)_{k=1}^{\infty}$ with $Y_k = X_k - X_{k-1}$ is the martingale difference sequence of $\left(X_k\right)_{k=0}^{\infty}$.
Most useful is a special martingale, the Doob-martingale.

Proposition 2.8. Let $f : \Omega \to \mathbb{R}$ be an integrable function, Σ, $\left(X_k\right)_{k=0}^{\infty}$ as above.

1. The sequence $\left(X_k\right)_k$ with $X_k = \mathbb{E}\left[f|\Sigma_k\right]$ is a martingale, called Doob's martingale.
2. If the filtration is finite such that $\{\phi, \Omega\} = \Sigma_0 \subseteq \cdots \subseteq \Sigma_n = \Sigma$, then $X_0 = \mathbb{E}\left[f\right]$ and $X_n = f$.

A finite filtration can be realized by a sequence of partitions of Ω. Suppose that $|\Omega| = n$. For each integer k, $0 \leqslant k \leqslant n-1$ let $\mathcal{P}_k$ be a partition of Ω where

(a) $\mathcal{P}_0 = \{\Omega\}$
(b) $\mathcal{P}_n = \{\{\omega\}; \omega \in \Omega\}$.
(c) $\mathcal{P}_{k+1}$ is finer than $\mathcal{P}_k$.

Then the σ-algebras Σ_k generated by the $\mathcal{P}_k$'s form a finite filtration with

$$\{\phi, \Omega\} = \Sigma_0 \subseteq \cdots \subseteq \Sigma_n = \mathcal{P}(\Omega)$$

where $\mathcal{P}(\Omega)$ is the power set of Ω. For an $\mathbb{P}$-integrable function $f : \Omega \to \mathbb{R}$, let $X_k = \mathbb{E}\left[f|\Sigma_k\right]$ be the Doob martingale. Joel Spencer gave a nice interpretation of this martingale ([Spe94], p. 56): the martingale process $\left(X_k\right)$ can be viewed as a process exhibiting more and more information about the function f. In the initial state of the process all information is hidden in the average $X_0 = \mathbb{E}\left[f\right]$, but in the final state all information is available, because $X_n = f$. So, for combinatorial functions f such a martingale is a "scanning machine" which provides us with information about f step by step.

2.3 Azuma Inequality and Bounded Differences Inequality (BDI)

We are interested in bounding large deviations. Let $\left(Y_k\right)_{k=1}^{n}$ be the martingale differences of the martingale $\left(X_k\right)_{k=0}^{n}$, $X_k = \mathbb{E}\left[f|\Sigma_k\right]$. Then

$$f - \mathbb{E}\left[f\right] = \sum_{k=1}^{n} Y_k$$

and the deviations of $|f - \mathbb{E}[f]|$ are exactly the deviations of the sum of the martingale differences. Deviations of the sum of martingale differences can be bounded by an exponentially decreasing function. This is achieved by Azuma's inequality [Azu67], a generalization of the Chernoff inequality. We now state the famous martingale inequality of K. Azuma.

Theorem 2.9. (Azuma-Hoeffding Inequality) Let $X_0, X_1, X_2, \ldots, X_n$ be a martingale sequence with bounded differences, i.e. $|X_k - X_{k-1}| \leqslant c_k$ for each $k = \{1, 2, \cdots, n\}$. Then for any $t > 0$, it holds

$$\mathbb{P}(|X_n - X_0| \geqslant t) \leqslant 2\exp\left(\frac{-t^2}{2\sum_{k=1}^{n} c_k^2}\right) \tag{2.4}$$

Proof. We will first derive an upper bound for $\mathbb{E}\left[e^{\alpha(X_n - X_0)}\right]$. We already defined $Y_k = X_k - X_{k-1}$ for $k = 1, 2, \cdots, n$. Since $X_0, X_1, \cdots, X_n$ is a martingale,

$$\begin{aligned}\mathbb{E}[Y_k \mid X_0, X_1, \cdots, X_{k-1}] &= \mathbb{E}[X_k - X_{k-1} \mid X_0, X_1, \cdots, X_{k-1}] \\ &= \mathbb{E}[X_k \mid X_0, X_1, \cdots, X_{k-1}] - X_{k-1} \\ &= 0\end{aligned}$$

Now consider

$$\mathbb{E}\left[e^{\alpha Y_k} \mid X_0, X_1, \cdots, X_{k-1}\right] \tag{2.5}$$

Writing

$$Y_k = -c_k \frac{1 - Y_k/c_k}{2} + c_k \frac{1 + Y_k/c_k}{2} \tag{2.6}$$

and using convexity of $e^{\alpha Y_k}$, we have that

$$\begin{aligned}e^{\alpha Y_k} &\leqslant \frac{1 - Y_k/c_k}{2} e^{-\alpha c_k} + \frac{1 + Y_k/c_k}{2} e^{\alpha c_k} \\ &= \frac{e^{\alpha c_k} + e^{-\alpha c_k}}{2} + \frac{Y_k}{2c_k}(e^{\alpha c_k} - e^{-\alpha c_k})\end{aligned}$$

Since $\mathbb{E}[Y_k \mid X_0, X_1, \cdots, X_{k-1}] = 0$, we have

$$\begin{aligned}\mathbb{E}\left[e^{\alpha Y_k} \mid X_0, X_1, \cdots, X_{k-1}\right] &\leqslant \mathbb{E}\left[\frac{e^{\alpha c_k} + e^{-\alpha c_k}}{2} + \frac{Y_k}{2c_k}(e^{\alpha c_k} - e^{-\alpha c_k}) \mid X_0, X_1, \cdots, X_{k-1}\right] \\ &= \frac{e^{\alpha c_k} + e^{-\alpha c_k}}{2}\end{aligned}$$

$$\leqslant e^{(\alpha c_k)^2/2}$$

It follows that

$$\begin{aligned}
\mathbb{E}\left[e^{\alpha(X_n - X_0)}\right] &= \mathbb{E}\left[\prod_{k=1}^{n} e^{\alpha Y_k}\right] \\
&= \mathbb{E}\left[\prod_{k=1}^{n-1} e^{\alpha Y_k}\right] \mathbb{E}\left[e^{\alpha Y_n} \mid X_0, X_1, \cdots, X_{n-1}\right] \\
&\leqslant \mathbb{E}\left[\prod_{k=1}^{n-1} e^{\alpha Y_k}\right] e^{(\alpha c_n)^2/2} \\
&\leqslant e^{\alpha^2 \sum_{k=1}^{n} c_k^2/2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{P}(X_n - X_0 \geqslant t) &= \mathbb{P}(e^{\alpha(X_n - X_0)} \geqslant e^{\alpha t}) \\
&\leqslant \frac{\mathbb{E}\left[e^{\alpha(X_n - X_0)}\right]}{e^{\alpha t}} \\
&\leqslant e^{\alpha^2 \sum_{k=1}^{n} c_k^2/2 - \alpha t} \\
&\leqslant e^{\frac{-t^2}{\left(2\sum_{k=1}^{n} c_k^2\right)}}
\end{aligned}$$

where the last inequality holds by choosing $\alpha = t/\sum_{k=1}^{n} c_k^2$. A similar argument gives the bound for $\mathbb{P}(X_n - X_0 \leqslant -t)$, by replacing X_n by $-X_n$, hence proving the theorem.

$\square$

Colin McDiarmid in his seminal paper [McD89] proved the generalized bounded differences inequality. The generalized bounded difference inequality (BDI) is a martingale-free formulation of Azuma's inequality. Let $Z = (Z_1, Z_2, \cdots, Z_n)$ be family of random variables with Z_k taking values in a set A_k. Let f be a real-valued function defined on $\prod A_k$.

Let V be an event that $Z_j = z_j$ where $z_j \in A_j$ for each $j = 1, 2, 3, \cdots, k-1$. Let the random variable W be distributed like Z_k conditioned on the event V. For $z \in A_k$, let

$$g(z) = \mathbb{E}\left[f(Z) \mid V, Z_k = z\right] - \mathbb{E}\left[f(Z) \mid V\right],$$

The function $g(z)$ measures how much the expected value of $f(Z)$ changes if it is revealed that Z_k takes the value z. We can easily see that $\mathbb{E}\left[g(W)\right] = 0$. Let *range* of $g(W)$, $ran(z_1, z_2, \cdots, z_{k-1})$ be defined as the function $\sup\{|g(z) - g(u)| : z, u \in A_k\}$. Let us define *maximum range* of $g(W)$, $\hat{r}_k$, as $\sup_{z_1,z_2,\cdots,z_{k-1}} ran(z_1, z_2, \cdots, z_{k-1})$ for all k. Let $\bar{\mathrm{r}}^2$, the *sum of squared maximum ranges*, be the sum of the values $\hat{r}_k^2$. So $\bar{\mathrm{r}}^2 = \sum_{k=1}^n \hat{r}_k^2$. We now state the generalized bounded differences inequality.

Theorem 2.10. (Generalized BDI) [McD98] Let $Z = (Z_1, Z_2, \cdots, Z_n)$ be the family of random variables that are not necessarily independent, with Z_k taking on values in a set A_k, and let f be a real-valued function defined on $\prod A_k$. Let μ denote the mean of $f(Z)$, and let $\bar{\mathrm{r}}^2$ denote the sum of squared maximum ranges. Then for any $t \geqslant 0$,

$$\mathbb{P}(f(Z) - \mu \geqslant t) \leqslant e^{-2t^2/\bar{\mathrm{r}}^2} \tag{2.7}$$

If the random variables $Z_1, Z_2, \cdots, Z_n$ are independent then the above theorem reduces to the independent bounded differences inequality (BDI) due to C. Mcdiarmid [McD89].

Theorem 2.11. (Independent BDI) Let $Z = (Z_1, Z_2, \ldots, Z_n)$ be a family of of independent random variables with Z_k taking on values from set A_k for each k. Suppose that the real valued function f defined on $\prod_{k=1}^n A_k$, satisfies $|f(Z) - f(Z')| \leqslant c_k$ for all $k \in [n]$ if the vectors Z and Z' differ only in the k-th coordinate. Then for any $t > 0$, it holds

(i) $\mathbb{P}(f(Z) \geqslant \mathbb{E}[f(Z)] + t) \leqslant \exp\left(\frac{-2t^2}{\sum_{k=1}^n c_k^2}\right)$

(ii) $\mathbb{P}(f(Z) \leqslant \mathbb{E}[f(Z)] - t) \leqslant \exp\left(\frac{-2t^2}{\sum_{k=1}^n c_k^2}\right)$

Chapter 3

Derandomizing the Generalized Bounded Differences Inequality

3.1 Derandomization Concept

Derandomization is the task of turning a probabilistic existence result or a randomized algorithm into a deterministic, polynomial-time algorithm, preferably achieving nearly the same guarantee. With the evident power of randomized algorithms, derandomization became an important task in algorithmic discrete mathematics. This is important, since deterministic algorithms give hard guarantees on performance and running times.

In derandomization the algorithmic problem often is to find points in the probability space satisfying events whose existence is guaranteed by large deviation bounds. Erdős and Selfridge [ES73], Beck and Fiala [BF81] and Spencer [Spe77] suggested for derandomization a very general technique, *the conditional probability method*. Unfortunately, in many cases conditional probabilities under consideration cannot be computed efficiently. A milestone for algorithmic progress has been the introduction of so called pessimistic estimators by Raghavan [Rag88], which are computable upper bounds on the conditional probabilities sharing the properties of them. One key fact for the success of the conditional probability method is that pessimistic estimators can be constructed, whenever linear objective functions are involved, because the proofs of Chernoff and Hoeffding type large deviation bound inequalities for sums of independent random variables deliver the right pessimistic estimators. In fact, for a polynomial-time implementation of the method, approximations of these estimators by Taylor polynomials must be used, too e.g. [SS96].

When the probability of the events under consideration can be estimated by Chernoff bounds or the Lovász Local Lemma (LLL), the problem has been solved e.g. Raghavan [Rag88], Beck [Bec91], Srivastav & Stangier [SS96], Matoušek [Mat96], Srinivasan et al. [Sri99], Moser & Tardos 2010 [MT10]. For obvious reasons let us call such a procedure, the derandomized (algorithmic) version of underlying large deviation bound. In this chapter, we resolve the derandomization problem for the Azuma inequality and the generalized bounded differences inequality (BDI) of C. McDiarmid [McD89] resp. using the conditional probability and Pessimistic estimator approach in a finite, but quite general probabilistic setting.

Our derandomized BDI forms a general and easily applicable framework for the derandomziation of randomized algorithms, where the BDI is used in the analysis. We further show, introducing the notion of a weak pessimistic estimator, that pessimistic estimators of the BDI can be joined with pessimistic estimators of other derandomized inequalities, like the Chernoff-Hoeffding bounds or the Lovász Local Lemma, and thus further extend the range of applicability.

3.2 Statement of the Derandomized Generalized BDI

We define first the probability space under consideration. Let $A_1, A_2, \cdots, A_n$ be finite sets with $|A_i| = N_i \in \mathbb{N}$ for all i, and set $N := \max_{1 \leqslant i \leqslant n} N_i$. Let $\Omega = \prod_{i=1}^{n} A_i$ be the sample space.

Let $\mathbb{P}$ be an arbitrary probability measure on Ω and $(\Omega, \mathbb{P})$ is a probability space with the powerset $\mathcal{P}(\Omega)$ of Ω as the σ-algebra. Let $Z_1, Z_2, \cdots, Z_n$ be n random variables, where Z_i takes value in the set A_i, $i \in [n]$. In this setting we do not assume that the $Z_1, Z_2, \cdots, Z_n$ are independent.

We say $\omega, \omega' \in \Omega$ are *k-equivalent*, i.e. $\omega \cong_k \omega'$ if and only if $\omega_j = \omega'_j$ for all $1 \leqslant j \leqslant k$. *$k$-equivalency* defines an equivalency relation on Ω and induces for each k a partition $\mathcal{P}_k$ of Ω with

$$\{\Omega\} = \mathcal{P}_0 \subset \ldots \subset \mathcal{P}_n = \{\{\omega\}; \omega \in \Omega\}$$

Let Σ_k be the σ-algebra generated by $\mathcal{P}_k$. Then the sequence $(\Sigma_k)_{k=0}^{n}$ is a filtration, where $\Sigma_0 = \{\varnothing, \Omega\}$ and $\Sigma_n = \mathcal{P}(\Omega)$.

Definition 3.1. Let $\psi_1, \ldots, \psi_m$ be rational-valued functions on Ω with $\mathbb{E}(\psi_i) = \mu_i$. For $i \in [m]$, let $\lambda_i > 0$ be rational numbers and define the event $E_i^{(+)}$ by $\psi_i \leqslant \mu_i + \lambda_i$ and let $E_i^{(-)}$ denote the event $\psi_i \geqslant \mu_i - \lambda_i$. Furthermore set $E = \bigcap_{i=1}^{m} E_i$ where E_i is either $E_i^{(+)}$ or $E_i^{(-)}$. W.l.o.g we can assume $E_i = E_i^{(+)}$.

Let $Z := (Z_1, Z_2, \cdots, Z_n)$. Let $X_{i0}, X_{i1}, \cdots, X_{in}$ be the martingale obtained by setting $X_{ik} = \mathbb{E}[\psi_i(Z) \,|\, \Sigma_k]$ for all $i \in [m], k \in [n]$. Let $Y_{i1}, Y_{i2}, \cdots, Y_{in}$ be the corresponding martingale difference sequence obtained by $Y_{ik} = X_{ik} - X_{i,k-1}$. Let us assume that

$$||Y_{ik}||_\infty \leqslant d_{ik} \text{ for all } k \in [n]. \tag{3.1}$$

Set for $l \in [n]$, $D_{il} := d_{il}^2 + d_{i,l+1}^2 + \cdots + d_{in}^2$. Then for any $\lambda_i > 0$, by Azuma's inequality (2.9), $\mathbb{P}(E_i^c) \leqslant f(\lambda_i)$ where $f(\lambda_i) = \exp(-\frac{2\lambda_i^2}{D_{i1}})$ for all $i \in [m]$. We assume that for some $\delta \in (0, 1)$

$$\sum_{i=1}^{m} f(\lambda_i) \leqslant 1 - \delta. \tag{3.2}$$

Then, the union bound gives

$$\mathbb{P}(\bigcap_{i=1}^{m} E_i) \geqslant 1 - \mathbb{P}(\bigcup_{i=1}^{m} E_i^c) \geqslant 1 - \bigcup_{i=1}^{m} \mathbb{P}(E_i^c) \geqslant 1 - \sum_{i=1}^{m} f(\lambda_i) \geqslant \delta > 0, \text{ thus } \bigcap_{i=1}^{m} E_i \neq \varnothing. \quad (3.3)$$

In this setting we can show the following general theorem:

Theorem 3.2. (Derandomized Azuma Inequality) Let $\psi_1, \ldots, \psi_m$ be rational valued functions on Ω and let $E_1, \ldots, E_m$ be associated events as in Definition 3.1 satisfying (3.2). Let P, Q denote functions in n, m, N and suppose that the following conditions are satisfied:

(1) $\max_{\omega \in \Omega} |\psi_i(\omega)| \leqslant P$ for all $i \in [m]$.

(2) For every $i \in [m]$, $\omega \in \Omega$ and $k \in [n]$, the conditional expectation $\mathbb{E}\left[\psi_i | \Sigma_k\right](\omega)$ can be computed in $\mathcal{O}(Q)$ time.

Then a vector $x \in \bigcap_{i=1}^{m} E_i$ can be constructed in $\mathcal{O}(mnN[P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q])$ time.

Remark 3.3. While the computation time in its general form here may depend on the arbitrary functions P and Q, in most of our applications P and Q will be polynomials. However, in the application to the Maker Breaker game, we will see that especially Q is not a polynomial, but superexponentially dependent on n. So, it is necessary and useful to specify P and Q only in a quite general form. Further, note that the theorem makes sense and is an efficient algorithm only, if P and Q are small compared to $|\Omega|$, because otherwise by brute force one could try all points of Ω and find the right one, where by 3.3 exists.

C. McDiarmid in his seminal paper [McD89] proved the generalized bounded differences inequality. The generalized bounded difference inequality (BDI) is a martingale-free formulation of Azuma's inequality. We use the above context to state the generalized bounded differences inequality.

Let V_{k-1} be an event that $Z_j = z_j$ where $z_j \in A_j$ for each $j = 1, 2, 3, \cdots, k-1$, $k \in [n]$. Let the random variable W be distributed like Z_k conditioned on the event V_{k-1}. For $z \in A_k$, let

$$g_{ik}(z) = \mathbb{E}\left[\psi_i(Z) \,|\, V_{k-1}, Z_k = z\right] - \mathbb{E}\left[\psi_i(Z) \,|\, V_{k-1}\right],$$

for all $i \in [m]$ and $k \in [n]$.

The function $g_{ik}(z)$ measures how much the expected value of $\psi_i(Z)$ changes if it is revealed that Z_k takes the value z. We can easily see that $\mathbb{E}\left[g_{ik}(W)\right] = 0$. Let *range* of $g_{ik}(W)$ is defined by

$$ran_k^{(i)}(z_1, z_2, \cdots, z_{k-1}) := \max\{|g_{ik}(z) - g_{ik}(u)| : z, u \in A_k\}$$

for all $i \in [m]$ and $k \in [n]$.

Let us define the *maximum range* of $g_{ik}(W)$ by,

$$\hat{r}_{ik} := \max_{z_1, z_2, \cdots, z_{k-1}} ran_k^{(i)}(z_1, z_2, \cdots, z_{k-1})$$

for all $i \in [m]$ and $k \in [n]$.

Let $\bar{\mathrm{r}}_i^2$, the *sum of squared maximum ranges*, be the sum of the values $\hat{r}_{ik}^2$. So

$$\bar{\mathrm{r}}_i^2 = \sum_{k=1}^{n} \hat{r}_{ik}^2$$

.

We now state the generalized bounded differences inequality.

Theorem 3.4. (Generalized BDI) [McD98] Let $Z = (Z_1, Z_2, \cdots, Z_n)$ be the family of random variables that are not necessarily independent, with Z_k taking on values in a set A_k, and let ψ_i be real-valued functions defined on $\prod A_k$ for all $i \in [m]$. Let μ_i denote the mean of $\psi_i(Z)$, and let $\bar{\mathrm{r}}_i^2$ denote the sum of squared maximum ranges for all $i \in [m]$. Then for any $\lambda_i \geqslant 0$,

$$\mathbb{P}(\psi_i(Z) - \mu_i \geqslant \lambda_i) \leqslant e^{-2\lambda_i^2/\bar{\mathrm{r}}_i^2} \tag{3.4}$$

for all $i \in [m]$.

If the random variables $Z_1, Z_2, \cdots, Z_n$ are independent and let $d_{ik} \geqslant 0$ denote the Lipschitz bounds, so for every k-th coordinate of the function ψ_i, we have

$$|\psi_i(Z_1, \ldots, Z_{k-1}, Z_k, Z_{k+1}, \ldots, Z_n) - \psi_i(Z_1, \ldots, Z_{k-1}, Z_k^{'}, Z_{k+1}, \ldots, Z_n)| \leqslant d_{ik} \tag{3.5}$$

We observe that

$$\hat{r}_{ik} = d_{ik}$$

for all $i \in [m], k \in [n]$.

We now state the bounded differences inequality (BDI) due to C. Mcdiarmid [McD89] for independent random variables. This corollary of generalized BDI is used in many combinatorial applications.

Theorem 3.5. (Independent BDI) Let $Z = (Z_1, Z_2, \ldots, Z_n)$ be a family of of independent random variables with Z_k taking on values from set A_k for each k. Suppose that the real valued functions ψ_i defined on $\prod_{k=1}^{n} A_k$ for all $i \in [m]$, satisfies $|\psi_i(Z) - \psi_i(Z')| \leqslant d_{ik}$ if the vectors Z and Z' differ only in the k-th coordinate. Then for any $t > 0$, it holds

(i) $\mathbb{P}(\psi_i(Z) \geqslant \mathbb{E}[\psi_i(Z)] + t) \leqslant \exp\left(\frac{-2t^2}{\sum_{k=1}^{n} d_{ik}^2}\right)$

(ii) $\mathbb{P}(\psi_i(Z) \leqslant \mathbb{E}[\psi_i(Z)] - t) \leqslant \exp\left(\frac{-2t^2}{\sum_{k=1}^{n} d_{ik}^2}\right)$

for all $i \in [m]$.

If random variables $Z_1, Z_2, \cdots, Z_n$ are not independent, then define

$$d_{ik} := \hat{r}_{ik}$$

for all $i \in [m], k \in [n]$.

Set for $l \in [n]$, $D_{il} := d_{il}^2 + \cdots + d_{in}^2$. Then for any $\lambda_i > 0$, by the generalized BDI (3.4), $\mathbb{P}(E_i^c) \leqslant f(\lambda_i)$ where $f(\lambda_i) := \exp(-\frac{2\lambda_i^2}{D_{i1}})$ for all $i \in [m]$. We assume that for some $\delta \in (0,1)$

$$\sum_{i=1}^{m} f(\lambda_i) \leqslant 1 - \delta, \tag{3.6}$$

so using the union bound

$$\mathbb{P}(\bigcap_{i=1}^{m} E_i) \geqslant 1 - \mathbb{P}(\bigcup_{i=1}^{m} E_i^c) \geqslant \delta > 0, \text{ thus } \bigcap_{i=1}^{m} E_i \neq \varnothing. \tag{3.7}$$

The derandomized generalized BDI is:

Theorem 3.6. (Derandomized Generalized BDI) Let $\psi_1, \ldots, \psi_m$ be rational valued functions on $\Omega = \prod A_k$ and let $E_1, \ldots, E_m$ be associated events as in Definition 3.1 satisfying (3.6). Let P, Q denote functions in n, m, N and suppose that the following conditions are satisfied.

(1) $\max_{\omega \in \Omega} |\psi_i(\omega)| \leqslant P$ for all $i \in [m]$.

(2) For every $i \in [m]$, $\omega \in \Omega$ and $k \in [n]$, the conditional expectation $\mathbb{E}\left[\psi_i | \Sigma_k\right](\omega)$ can be computed in $\mathcal{O}(Q)$ time.

Then a vector $x \in \bigcap_{i=1}^{m} E_i$ can be constructed in $\mathcal{O}(mnN[P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q])$ time.

With Theorem 3.2 we can immediately prove the derandomized generalized BDI.

Proof. Let the σ-field Σ_k in the filtration be the σ-field generated by $Z_1, Z_2, \cdots, Z_k$ where $\Sigma_0 = \{\varnothing, \Omega\}$. Let $X_{i0}, X_{i1}, \cdots, X_{in}$ be the martingale obtained by setting $X_{ik} = \mathbb{E}\left[\psi_i(Z) \,|\, \Sigma_k\right]$ for all $i \in [m], k \in [n]$. Let $Y_{i1}, Y_{i2}, \cdots, Y_{in}$ be the corresponding martingale difference sequence obtained by $Y_{ik} = X_{ik} - X_{i,k-1}$ for all $i \in [m], k \in [n]$.

We need to upper bound Y_{ik} in terms of the maximum range function defined earlier. We observe Y_{ik} is uppper bounded by

$$\max_{a \in A_k} \mathbb{E}\left[\psi_i(Z) \,|\, \Sigma_{k-1}, Z_k = a\right] - \mathbb{E}\left[\psi_i(Z) \,|\, \Sigma_{k-1}\right]$$

and bounded from the below by

$$\min_{b \in A_k} \mathbb{E}\left[\psi_i(Z) \mid \Sigma_{k-1}, Z_k = b\right] - \mathbb{E}\left[\psi_i(Z) \mid \Sigma_{k-1}\right].$$

Thus if we can upper bound the below quantity then we are done.

$$\max_{a \in A_k} \mathbb{E}\left[\psi_i(Z) \mid \Sigma_{k-1}, Z_k = a\right] - \min_{b \in A_k} \mathbb{E}\left[\psi_i(Z) \mid \Sigma_{k-1}, Z_k = b\right]$$

The above quantity is equivalent to:

$$\max_{a,b \in A_k} \mathbb{E}\left[\psi_i(Z) \mid \Sigma_{k-1}, Z_k = a\right] - \mathbb{E}\left[\psi_i(Z) \mid \Sigma_{k-1}, Z_k = b\right]$$

Hence by definition of range and maximum range, we establish

$$||Y_{ik}||_\infty \leqslant \hat{r}_{ik}$$

So, the associated martingale satisfies (3.1), thus the derandomized Azuma inequality implies the derandomized generalized BDI. □

We use the derandomized independent bounded differences inequality for various applications, hence we will now state the same. The proof is same as the proof for generalized BDI, hence not stated here. Observe that $|A_i| = N$ for all $i \in [n]$.

Theorem 3.7. (Derandomized Independent BDI) Let $\psi_1, \ldots, \psi_m$ be rational valued functions on $\Omega = \prod A_k$ and let $E_1, \ldots, E_m$ be associated events as in Definition 3.1 satisfying (3.6). Let P, Q denote functions in n, m, N and suppose that the following conditions are satisfied.

(1) $\max_{\omega \in \Omega} |\psi_i(\omega)| \leqslant P$ for all $i \in [m]$.

(2) For every $i \in [m]$, $\omega \in \Omega$ and $k \in [n]$, the conditional expectation $\mathbb{E}\left[\psi_i | \Sigma_k\right](\omega)$ can be computed in $\mathcal{O}(Q)$ time.

Then a vector $x \in \bigcap_{i=1}^m E_i$ can be constructed in $\mathcal{O}(mnN[P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q])$ time.

We define the pessimistic estimators in the next section and use them to prove derandomized Azuma inequality.

3.3 General Form of Pessimistic Estimators

The basic notion of pessimistic estimators was introduced by Raghavan [Rag88]. We generalize this definition in order to cope with different concentration bounds.

Definition 3.8. Let $\mathcal{U}$ be a family of functions of the form $U_l : [N]^l \mapsto \mathbb{Q}$, $l \in [n]$ plus a function U_0 . Let $(\Omega, \mathbb{P})$ be a probability space, $E \subset \Omega$ an event and $0 < \delta < 1$.

(i) $\mathcal{U}$ is called a *weak* pessimistic estimator for the event E, if for each $l \in [n]$ the following conditions are satisfied:

(a) $\mathbb{P}(E^c|\omega_1, \ldots, \omega_l) \leqslant U_l(\omega_1, \ldots, \omega_l)$ for all $\omega_1, \ldots, \omega_l \in [N]$.

(b) Given $\omega_1, \ldots, \omega_{l-1}$, there exists an ω_l such that $U_l(\omega_1, \ldots, \omega_l) \leqslant U_{l-1}(\omega_1, \ldots, \omega_{l-1})$.

(c) $U_0 \leqslant 1 - \delta$.

(ii) $\mathcal{U}$ is called a pessimistic estimator, if $\mathcal{U}$ is a weak pessimistic estimator and each value $U_l(\omega_1, \ldots, \omega_l)$ and U_0 can be computed in time polynomially bounded in n, N and $\log \frac{1}{\delta}$.

In the next definition, we join the weak pessimistic estimators of finitely many events to a weak pessimistic estimators of their intersection. The crucial property is a kind of convexity.

Definition 3.9. 1. A family $\mathcal{U}$ as in Definition 3.8 is called *convex*, if there are real numbers $\mu_j > 0$, $j \in [N]$, with $\mu_1 + \ldots + \mu_N = 1$ such that $\sum_{j=1}^{N} \mu_j U_l(\omega_1, \ldots, \omega_{l-1}, j) \leqslant U_{l-1}(\omega_1, \ldots, \omega_{l-1})$ for all $\omega_1, \ldots, \omega_l \in [N]$ and all $l \in [n]$.

2. The sum $\bigoplus_{i=1}^{m} \mathcal{U}_i$ of m families $\mathcal{U}_i$ is the set of functions U_l with $U_l := \sum_{i=1}^{m} U_l^{(i)}$, $U_l^{(i)} \in \mathcal{U}_i$, $l \in [n]$.

Proposition 3.10. If $\mathcal{U}_i$ is a weak pessimistic estimator for an event E_i, $i \in [m]$, $\bigoplus_{i=1}^{m} \mathcal{U}_i$ is convex and $\sum_{i=1}^{m} U_0^{(i)} \leqslant 1 - \delta$ for some $\delta > 0$, then $\bigoplus_{i=1}^{m} \mathcal{U}_i$ is a weak pessimistic estimator for the event $\bigcap_{i=1}^{m} E_i$.

Proof. We need to establish the conditions specified in Definition 3.8: condition (c) is true by the assumption, thus we need to prove conditions (a) and (b). Let $l \in [n]$ and $\omega_1, \ldots, \omega_l \in [N]$ be arbitrary, but fixed. Then

$$\begin{aligned}
\mathbb{P}((\bigcap_{i=1}^{m} E_i)^c \,|\, \omega_1, \ldots, \omega_l) = \mathbb{P}(\bigcup_{i=1}^{m} E_i^c \,|\, \omega_1, \ldots, \omega_l) &\leqslant \sum_{i=1}^{m} \mathbb{P}(E_i^c|\omega_1, \ldots, \omega_l) \\
&\leqslant \sum_{i=1}^{m} U_l^{(i)}(\omega_1, \ldots, \omega_l) \\
&= U_l(\omega_1, \ldots, \omega_l)
\end{aligned}$$

Condition (b) is a direct consequence of the convexity of $\bigoplus_{i=1}^{m} \mathcal{U}_i$. Let $\bigoplus_{i=1}^{m} \mathcal{U}_i$ be the functions $(U_l)_l$ as in Definition 3.9 (2). Define ω_l as the minimizer of $j \mapsto U_l(\omega_1, \ldots, \omega_{l-1}, j)$, so

$$U_l(\omega_1, \ldots, \omega_{l-1}, \omega_l) \leqslant \sum_{j=1}^{N} \mu_j U_l(\omega_1, \ldots, \omega_{l-1}, j) \leqslant U_l(\omega_1, \ldots, \omega_{l-1})$$

□

With pessimistic estimators, we have a polynomial time implementation of the conditional probability method.

Algorithm DERAND

INPUT: An event $E \subset \prod A_k$ and a pessimistic estimator $\mathcal{U}$ for E.

OUTPUT: A vector $x \in E$.

ALGORITHM: For $l = 0, \ldots, n-1$ do:
If $x_1 \ldots, x_{l-1}$ have been selected, choose $x_l \in A_l$ as the minimizer of the function $\omega \to U_{l+1}(E^c \mid x_1 \ldots, x_{l-1}, \omega), \quad \omega \in A_l$.

Suppose we can compute each evaluation of the function $U_l, l \in [n] \cup \{0\}$, in $\mathcal{O}(t(U))$ time. The striking observation is

Proposition 3.11. The algorithm DERAND computes $x \in E$ in $\mathcal{O}(nNt(U))$ time.

Proof. Since $\mathcal{U}$ is a pessimistic estimator for the event E, each $U_l(x_1, \ldots, x_{l-1}, \omega)$, $\omega \in A_l$, can be computed in $\mathcal{O}(t(U))$ time, thus the minimizer x_l can be computed in $\mathcal{O}(Nt(U))$ time. The vector $x = (x_1, \ldots, x_n)$ satisfies

$$\begin{aligned} \mathbb{P}(E^c \,|\, x_1, \ldots, x_n] \leqslant U_n(x_1, \ldots, x_n) &\leqslant U_{n-1}(x_1, \ldots, x_{n-1}) \\ &\vdots \\ &\leqslant U_0 \leqslant 1 - \delta < 1, \end{aligned}$$

so $\mathbb{P}(E^c \,|\, x_1, \ldots, x_n) = 0$, and $x \in E$. Since we iterate the minimizer computation over all n variables, we consume $\mathcal{O}(nNt(U))$ time.

□

The next statement will be used to ensure that a suitable approximation of a weak pessimistic estimator is still a weak pessimistic estimator. Now, if the approximation is computable in polynomial time, we get the desired pessimistic estimator.

Proposition 3.12. Let $\mathcal{V} = (V_l)$, $l \in [n] \cup \{0\}$ be a weak pessimistic estimator for an event $E \subset \prod A_k$ and let $\delta > 0$. Suppose that $V_0 \leqslant 1 - \delta$. Let $\gamma < \frac{\delta}{4n+1}$ and let $\mathcal{W} = (W_l)$, $l \in [n] \cup \{0\}$ be a family of functions with

$$|V_l - W_l| \leqslant \gamma \text{ for all } \ [n] \cup \{0\}.$$

Then the family $\mathcal{U} = (U_l)$, $l \in [n] \cup \{0\}$, with $U_l := W_l + 2(2n - l)\gamma$ is a weak pessimistic estimator for E.

Proof. We must show that (U_l), $l \in [n] \cup \{0\}$, satisfies all conditions of Definition 3.8.

Condition (a): For arbitrary $l \in [n]$, let $\tilde{\omega} := (\omega_1, \ldots, \omega_l)$ where $\omega_1, \ldots, \omega_l \in [N]$ are arbitrary, but fixed. We have

$$\begin{aligned} \mathbb{P}(E^c \mid \omega_1, \omega_2, \ldots, \omega_l) &\leqslant V_l(\tilde{\omega}) \\ &= V_l(\tilde{\omega}) - W_l(\tilde{\omega}) - 2(2n-l)\gamma + W_l(\tilde{\omega}) + 2(2n-l)\gamma \\ &\leqslant W_l(\tilde{\omega}) + 2(2n-l)\gamma \\ &= U_l(\tilde{\omega}). \end{aligned}$$

Condition (b): Let $l \in [n]$ and $\tilde{\omega} := (\omega_1, \ldots, \omega_l)$, where $\omega_1, \ldots, \omega_l \in [N]$ are arbitrary, but fixed. If $x_1, x_2, \ldots, x_l \in [N]$ are already chosen and $U_l(x_1, x_2, \ldots, x_l)$ is known, let x_{l+1} be the value that minimizes the function $\omega \mapsto U_{l+1}(x_1, x_2, \ldots, x_l, \omega)$, $\omega \in [N]$. We have

$$\begin{aligned} U_l(\tilde{\omega}) - U_{l+1}(\tilde{\omega}, x_{l+1}) &= W_l(\tilde{\omega}) + 2(2n-l)\gamma - W_{l+1}(\tilde{\omega}, x_{l+1}) - 2(2n-(l+1))\gamma \\ &= W_l(\tilde{\omega}) - W_{l+1}(\tilde{\omega}, x_{l+1}) + 2\gamma \\ &= W_l(\tilde{\omega}) - V_l(\tilde{\omega}) + V_{l+1}(\tilde{\omega}, x_{l+1}) - W_{l+1}(\tilde{\omega}, x_{l+1}) \\ &+ V_l(\tilde{\omega}) - V_{l+1}(\tilde{\omega}, x_{l+1}) + 2\gamma \\ &\geqslant V_l(\tilde{\omega}) - V_{l+1}(\tilde{\omega}, x_{l+1}) \geqslant 0. \end{aligned}$$

Hence condition (b) of Definition 3.8 is established with the choice of x_{l+1}.

Condition (c):

For $l = 0$, we show that $U_0 \leqslant 1 - \beta$ where $\beta := \delta - (4n+1)\gamma$:

$$\begin{aligned} U_0 &= W_0 + 4n\gamma \;=\; W_0 - V_0 + V_0 + 4n\gamma \\ &\leqslant V_0 + (4n+1)\gamma \;\leqslant\; 1 - \delta + (4n+1)\gamma \\ &= 1 - \beta. \end{aligned}$$

□

3.4 Proofs for the Derandomized Inequalities

Srivastav and Stangier [SS96] constructed pessimistic estimators for events which are governed by concentration inequalities like Angluin-Valiant bound (Theorem 2.3) for the sum of independent $\{0,1\}$-random variables and is a variation of Chernoff or Hoeffding's bound. Using Proposition (3.10), we can construct pessimistic estimators for mixed situation where events are governed by generalized BDI and Angluin-Valiant bound. We will set the context for derandomized Angluin-Valiant bound. Let $X_1, \ldots, X_n$ be independent $\{0,1\}$-random variables. Let $\phi_i = \sum_{j=1}^n a_{ij} X_j$ with $a_{ij} \in [0,1]$, $i \in [l], j \in [n]$.

Definition 3.13. For $\beta_i > 0$, let us define the event $F_i^{(+)}$ by $\phi_i \leqslant \mathbb{E}[\phi_i](1+\beta_i)$ and let $F_i^{(-)}$ denote the event $\phi_i \geqslant \mathbb{E}[\phi_i](1-\beta_i)$. Let F_i be one of these events. W.l.o.g we can assume $F_i = F_i^{(+)}$.

Then by the Angluin-Valiant inequality (2.3), $\mathbb{P}(F_i^c) \leqslant g(\beta_i)$ with $g(\beta_i) = \exp\left(-\frac{\beta_i^2 \mathbb{E}[\phi_i]}{p}\right)$ for all i, and some $p \in \{2,3\}$. We assume that for some $\epsilon \in (0,1)$,

$$\sum_{i=1}^{l} \mathbb{P}(F_i^c) \leqslant \sum_{i=1}^{l} g(\beta_i) \leqslant 1-\epsilon,$$

$$\text{so using the union bound, } \mathbb{P}(\bigcap_{i=1}^{l} F_i) \geqslant 1 - \mathbb{P}(\bigcup_{i=1}^{l} F_i^c) \geqslant \epsilon > 0, \text{ thus } \bigcap_{i=1}^{l} F_i \neq \varnothing \tag{3.8}$$

For the mixed situation, where $\mathcal{E} = \{E_1, \ldots, E_m\}$ is a set of m events in the generalized BDI setting and $\mathcal{F} = \{F_1, \ldots, F_l\}$ is a set of l events in the Angluin-Valiant setting as above, satisfying (3.7) resp. (3.8). We furthermore assume

$$\sum_{i=1}^{m} \mathbb{P}(E_i^c) + \sum_{j=1}^{l} \mathbb{P}(F_j^c) \leqslant \sum_{i=1}^{m} f(\lambda_i) + \sum_{j=1}^{l} g(\beta_j) \leqslant 1-\delta+1-\epsilon < 1 \tag{3.9}$$

Thus

$$\mathbb{P}(\bigcap_{i=1}^{m} E_i \cap \bigcap_{j=1}^{l} F_j) > \epsilon + \delta - 1 > 0, \text{ so } \bigcap_{i=1}^{m} E_i \cap \bigcap_{j=1}^{l} F_j \neq \varnothing. \tag{3.10}$$

We have

Theorem 3.14. Consider $\mathcal{E}$ and $\mathcal{F}$ as above satisfying (3.10). Let P, Q denote polynomials in n, m, N satisfying the conditions of Theorem 3.6. Then a vector $x \in (\bigcap_{i=1}^m E_i \cap \bigcap_{j=1}^l F_j)$ can be constructed in $\mathcal{O}(mnN[P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q] + Nn^2 l \log \frac{Nln}{\epsilon})$ time.

Note that for $\mathcal{E} = \varnothing$ we have the derandomized Angluin-Valiant bound which is proved in [SS96]. Theorem 3.14 is thus the mixed situation.

Proof. By Proposition 3.10, we can combine the pessimistic estimators of the Angluin-Valiant bound (constructed in [SS96]) and generalized BDI resp. to get the pessimistic estimator for the event $\bigcap_{i=1}^m E_i \cap \bigcap_{j=1}^l F_j$. The time complexity to derandomize the event $\bigcap_{j=1}^l F_j$ is $\mathcal{O}(Nn^2 l \log \frac{Nln}{\epsilon})$ [SS96] and for the event $\bigcap_{i=1}^m E_i$ is $\mathcal{O}(mnN[P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q])$ (by Theorem 3.6) and by adding these time complexities we get our result. □

We now define the pessimistic estimators for the derandomization of Azuma's inequality and proceed to the proof of Theorem 3.2. With the next lemma, we reduce the computation of the conditional expectation of the moment generating function to the computation of the conditional expectation of the objective function.

Lemma 3.15. Let $t > 0$ and $i \in [m]$. For all $k \in [n]$, let $C \in \mathcal{P}_k$. We have

$$\mathbb{E}\left[e^{tX_{ik}}|C\right] = e^{t\mathbb{E}[\psi_i|C]}.$$

Proof. Let $\mathcal{P}_k = \{C_1, \ldots, C_l\}$ and for $C \in \mathcal{P}_k$, let $\mathbf{1}_C$ be the characteristic function of C. Because $(X_k)_k$ is a discrete, finite martingale, we have according to the definition of the Doob martingale and Theorem 2.6 (viii), $X_{ik} = \sum_{C \in \mathcal{P}_k} \mathbf{1}_C \mathbb{E}\left[\psi_i|C\right]$.

Hence for an arbitary, but fixed $C \in \mathcal{P}_k$ we have for $\omega \in C$, $X_{ik}(\omega) = \mathbb{E}\left[\psi_i|C\right]$, so

$$\begin{aligned}\mathbb{E}\left[e^{tX_{ik}}|C\right] = \frac{1}{\mathbb{P}(C)}\sum_{\omega\in C} e^{tX_{ik}(\omega)}\mathbb{P}(\omega) &= \frac{1}{\mathbb{P}(C)}\sum_{\omega\in C} e^{t\mathbb{E}[\psi_i|C]}\mathbb{P}(\omega) \\ &= e^{t\mathbb{E}[\psi_i|C]}\frac{1}{\mathbb{P}(C)}\sum_{\omega\in C}\mathbb{P}(\omega) = e^{t\mathbb{E}[\psi_i|C]}.\end{aligned}$$

□

The following lemma will be used frequently in upcoming proofs.

Lemma 3.16. ([McD98], Lemma 2.6)
Let X be a random variable with $\mathbb{E}[X] = 0$ and $a \leqslant X \leqslant b$ where a, b are constants. Then for any $t > 0$

$$\mathbb{E}\left[e^{tX}\right] \leqslant \exp\left(\frac{1}{8}t^2(b-a)^2\right).$$

Next we define the functions which will form the weak pessimistic estimators for the events E_i. Let s_i be the signum of the event E_i: we set $s_i = 1$ iff $E_i = E_i^+$ and $s_i = -1$ iff $E_i = E_i^-$.

Definition 3.17. For each $i \in [m]$, let $\mathcal{V}_i$ be a family of functions $V_{il} : \prod_{i=1}^{l} A_i \mapsto \mathbb{Q}$, $l \in [n] \cup \{0\}$ defined as follows. For $\omega_1, \ldots, \omega_l \in [N]$, set

(i) $V_{il}(\omega_1, \ldots, \omega_l) = e^{-t_i(\lambda_i + s_i X_{i0})} e^{\frac{t_i^2 D_{i,l+1}}{8}} e^{t_i s_i \mathbb{E}(\psi_i|\omega_1,\ldots,\omega_l)}$

(ii) $V_{i0} = e^{-t_i\lambda_i} e^{\frac{t_i^2 D_{i1}}{8}}$.

Theorem 3.18. Under the assumption (3.3) and with $t_i := \frac{4\lambda_i}{D_{i1}}$, the family $\bigoplus_{i=1}^{m} \mathcal{V}_i$ is a weak pessimistic estimator for the event $\bigcap_{i=1}^{m} E_i$.

Proof. We first show that each $\mathcal{V}_i$ is a weak pessimistic estimator for the event E_i. Then we prove the convexity of $\bigoplus_{i=1}^m \mathcal{V}_i$ and by Proposition 3.10 we will be done.

Fix an arbitrary $i \in [m]$. W.l.o.g we assume that $E_i = E_i^+$, so $s_i = 1$. Let $l \in [n]$ and $\omega_1, \ldots, \omega_l \in \prod_{i=1}^l A_i$. Define

$$C := \{\omega' \in \Omega;\ \omega'_k = \omega_k \ \text{ for } \ k = 1, \ldots, l\},$$

and for $j \in A_{l+1}$

$$C_j := \{\omega' \in \Omega;\ \omega'_k = \omega_k \ \text{ for } \ k = 1, \ldots, l \ \text{ and } \ \omega'_{l+1} = j\}.$$

We first show that $\mathcal{V}_i$ is a weak pessimistic estimator for E_i. Let us check conditions (a) – (c) of Definition 3.8.

Condition (a): We have

$$\begin{aligned}
\mathbb{P}(E_i^c|\omega_1, \ldots, \omega_l) &= \mathbb{P}(E_i^c|C) \\
&= \mathbb{P}(X_{in} - X_{i0} > \lambda_i|C) \\
&\leqslant e^{-t_i\lambda_i}\mathbb{E}\left[e^{t_i(X_{in}-X_{i0})}|C\right] \\
&= e^{-t_i\lambda_i}\mathbb{E}\left[\mathbf{1}_C e^{t_i(X_{in}-X_{i0})}\right] \cdot \mathbb{P}(C)^{-1} \\
&= e^{-t_i\lambda_i}\mathbb{E}\left[\mathbb{E}(\mathbf{1}_C e^{t_i(X_{in}-X_{i0})}|\Sigma_{n-1})\right] \cdot \mathbb{P}(C)^{-1} \\
&= e^{-t_i\lambda_i}\mathbb{E}\left[\mathbb{E}(\mathbf{1}_C e^{t_i(X_{i,n-1}-X_{i0})}e^{t_iY_{in}}|\Sigma_{n-1})\right] \cdot \mathbb{P}(C)^{-1} \\
&= e^{-t_i\lambda_i}\mathbb{E}\left[\mathbf{1}_C e^{t_i(X_{i,n-1}-X_{i0})}\mathbb{E}(e^{t_iY_{in}}|\Sigma_{n-1})\right] \cdot \mathbb{P}(C)^{-1} \text{ (Theorem 2.6 (vii))} \\
&\leqslant e^{-t_i\lambda_i}e^{\frac{t_i^2 d_{in}^2}{8}} \cdot \mathbb{E}\left[\mathbf{1}_C\, e^{t_i(X_{i,n-1}-X_{i0})}\right] \cdot \mathbb{P}(C)^{-1} \text{ (Lemma 3.16)} \\
&= e^{-t_i\lambda_i}e^{\frac{t_i^2 d_{in}^2}{8}} \cdot \mathbb{E}\left[e^{t_i(X_{i,n-1}-X_{i0})}|C\right] \\
&\leqslant e^{-t_i\lambda_i}e^{\frac{t_i^2 D_{i,l+1}}{8}} \cdot \mathbb{E}\left(\mathbf{1}_C e^{t_i(X_{il}-X_{i0})}\right) \cdot \mathbb{P}(C)^{-1} \\
&= e^{-t_i\lambda_i}e^{\frac{t_i^2 D_{i,l+1}}{8}} e^{t_i\mathbb{E}[\psi_i|C]}e^{-t_iX_{i0}} \text{ (Lemma 3.15)} \\
&= V_{il}(\omega_1, \ldots, \omega_l)\,.
\end{aligned}$$

Condition (b): We first show the convexity of $\mathcal{V}_i$ for all $i \in [m]$. For $i \in [m]$ and $j \in [N]$ put $\mu_{ij} := \frac{\mathbb{P}(C_j)}{\mathbb{P}(C)}$. So μ_{ij} does not depend on i. Then for any choice of $\omega_1, \ldots, \omega_l \in [N]$

$$\sum_{j=1}^N V_{i,l+1}(\omega_1, \ldots, \omega_l, j)\mu_{ij} = \underbrace{e^{-t_i\lambda_i}e^{\frac{t_i^2 D_{i,l+2}}{8}} \sum_{j=1}^N \mu_{ij}e^{t_i\mathbb{E}[\psi_i|C_j]-t_iX_{i0}}}_{:= \mathrm{T}} \tag{3.11}$$

With Lemma 3.15
$$\exp(t_i\mathbb{E}\left[\psi_i|C_j\right]) = \mathbb{E}\left[e^{\,t_iX_{i,l+1}}|C_j\right],$$
so we have
$$\sum_{j=1}^{N}\mu_{ij}e^{\,t_i\mathbb{E}[\psi_i|C_j]} = \sum_{j=1}^{N}\mu_{ij}\mathbb{E}\left[e^{\,t_iX_{i,l+1}}|C_j\right] = \frac{1}{\mathbb{P}(C)}\sum_{j=1}^{N}\mathbb{P}(C_j)\mathbb{E}\left[e^{\,t_iX_{i,l+1}}|C_j\right] = \mathbb{E}\left[e^{\,t_iX_{i,l+1}}|C\right]. \tag{3.12}$$

We continue
$$\begin{aligned}
T &= e^{-t_i\lambda_i}e^{\frac{t_i^2D_{i,l+2}}{8}}\left(\sum_{j=1}^{N}\mu_{ij}e^{\,t_i\mathbb{E}[\psi_i|C_j]}\right)e^{-t_iX_{i0}}\\
&= e^{-t_i\lambda_i}e^{\frac{t_i^2D_{i,l+2}}{8}}\mathbb{E}\left[e^{\,t_iX_{i,l+1}}|C\right]\cdot e^{-t_i\mathbb{E}[\psi_i]} \text{ (By 3.12)}\\
&\leqslant e^{-t_i\lambda_i}e^{\frac{t_i^2D_{i,l+1}}{8}}\mathbb{E}\left[e^{\,t_iX_{il}}|C\right]e^{-t_iX_{i0}} \text{ (Similar to the proof of (a))}\\
&= e^{-t_i\lambda_i}e^{\frac{t_i^2D_{i,l+1}}{8}}e^{\,t_i[\mathbb{E}[\psi_i|C]-X_{i0}]}\\
&= V_{il}(\omega_1,\ldots,\omega_l).
\end{aligned}$$

Condition (c): For $i\in[m]$ put $t_i = \frac{4\lambda_i}{D_{i1}}$. Then the proof of Theorem 2.9 gives
$$V_{i0} = e^{-t_i\lambda_i}e^{\frac{t_i^2D_{i1}}{8}} \leqslant f(\lambda_i).$$

We can now conclude the proof. We have $\sum_{i=1}^m V_{i0} \leqslant \sum_{i=1}^m f(\lambda_i) \leqslant 1-\delta$ according to (3.3). Since the $\mu_{ij} = \frac{\mathbb{P}(C_j)}{\mathbb{P}(C)}$, $i\in[m]$, $i\in[m]$ do not depend on i, the convexity of $\mathcal{V}_i$'s implies the convexity of $\bigoplus_i\mathcal{V}_i$. Hence by Proposition 3.10, $\bigoplus_i\mathcal{V}_i$ is a weak pessimistic estimator for the event $\bigcap_{i=1}^m E_i$.

□

We state here the technical Lemma 2.7 from [SS96] which will be used in the proof of Theorem 3.2. It is an extension of Brent's approximation [Bre76] of the $\exp(x)$ function on a compact interval to arbitrary rational numbers involving their encoding length.

Lemma 3.19. Let y be a rational number with encoding length L and let $\gamma_1\in(0,1)$ be a positive real number. Let q be a positive integer with $q\geqslant 8\lceil|y|\rceil+\lceil\log\frac{1}{\gamma_1}\rceil$. Then the q-th degree Taylor polynomial, $T_q(y)=\sum_{k=0}^{q}\frac{y^k}{k!}$ of $\exp(y)$ has encoding length $\mathcal{O}(Lq+q\log q)$, can be computed in $\mathcal{O}(q)$ time and the inequality $|\exp(y)-T_q(y)|\leqslant\gamma_1$ holds.

Proof of Theorem 3.2: By Theorem 3.18, we have constructed a weak pessimistic estimator. We approximate them by efficiently computable polynomials using Lemma 3.19. Let us define $\mathcal{V} := \bigoplus_i\mathcal{V}_i$ where the $\mathcal{V}_i$ are as in Theorem 3.18. Recall that s_i, $i\in[m]$, is the sign of the event E_i, i.e. $s_i=+1$ or $s_i=-1$,$i\in[m]$. For $(\omega_1,\ldots,\omega_l)\in\prod_{i=1}^{l}A_i$ each function $V_{il}\in\mathcal{V}_i$ has the form

$$V_{il}(\omega_1,\ldots,\omega_l) = \exp\left(-t_i(\lambda_i + s_i\mathbb{E}(\psi_i)) + \frac{t_i^2 D_{i,l+1}}{8} + t_i s_i \mathbb{E}\left[\psi_i | \omega_1,\ldots,\omega_l\right]\right). \tag{3.13}$$

Denote the exponent in the r.h.s of (3.13) by g_{il}. We first bound $|g_{il}|$. By the assumption (i) of Theorem 3.2, $||\psi_i||_\infty \leqslant P$ for all $i \in [m]$, where P is a polynomial in n, m, N. We can assume that $\lambda_i \leqslant P$ for $i \in [m]$ (otherwise, if for some i, $\lambda_i > P$, the inequality $\psi_i \leqslant \mathbb{E}[\psi_i] + \lambda_i$ is trivial and can be neglected). Then using $t_i = \frac{4\lambda_i}{D_{i1}}$ and the assumption (i) in Theorem 3.2 we get

$$\max_{i,l} |g_{il}| = \mathcal{O}(P \max_i \frac{\lambda_i}{D_{i1}}).$$

Set $\gamma := \frac{\delta}{m(4n+1)}$. Let T be the q-th degree Taylor polynomial of the exponential function with $q = \alpha(P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{1}{\gamma})$ and let $\alpha > 0$ be a constant. Invoking Lemma 3.19 with a sufficiently large α we have for all $i \in [m]$ and $\omega_1,\ldots,\omega_l \in [N]$

$$|V_{il}(\omega_1,\ldots,\omega_l) - T(g_{il})| \leqslant \gamma. \tag{3.14}$$

Note that T depends upon $\omega_1,\ldots,\omega_l$ as well. Define the family $\mathcal{W}_i = (W_l^{(i)})_l$, $i \in [m]$, by $W_l^{(i)}(\omega_1,\ldots,\omega_l) := T(g_{il})$. Further define the family $\mathcal{U}_i = (U_l^{(i)})_l$ by

$$U_l^{(i)}(\omega_1,\ldots,\omega_l) := W_l^{(i)}(\omega_1,\ldots,\omega_l) + 2(2n-l)\gamma, \tag{3.15}$$

Set $U := \bigoplus_i \mathcal{U}_i$. By Theorem 3.18, $\mathcal{V}$ is a weak pessimistic estimator for $\bigcap_{i=1}^m E_i$. Since $U_l = W_l + 2(2n-l)\frac{\delta}{4n+1}$ for all $l \in [n] \cup \{0\}$, by Proposition 3.12, $\mathcal{U}$ is a a weak pessimistic estimator for $\bigcap_{i=1}^m E_i$ as well.

We show now the claimed time complexity. According to Lemma 3.19 one evaluation of T for exponents as given in (3.13) takes

$$\mathcal{O}(q) = \mathcal{O}(P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta}) \tag{3.16}$$

time. Let $i \in [m]$ and $l \in [n] \cup \{0\}$. In order to compute $W_l^i(\omega_1,\ldots,\omega_l)$ we have to compute the exponent of (3.13). By the assumption (ii) in Theorem 3.2, the computation of $\mathbb{E}[\psi_i \mid \omega_1,\ldots,\omega_l]$ takes $\mathcal{O}(Q)$ time. Thus the named exponent can be computed in $\mathcal{O}(Q)$ time. For fixed $i \in [m]$ and $l \in [n] \cup \{0\}$ the time to compute $T(g_{il})$ is therefore $\mathcal{O}(P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q)$, and this is also the time to compute each evaluation of $U_l^{(i)}$. Thus for each $l \in [n] \cup \{0\}$, the computation time for U_l is $\mathcal{O}(m(P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q))$. We can run the algorithm DERAND with $\mathcal{U}$ and get by Proposition 3.11 a total running time of $\mathcal{O}(mnN(P \max_i \frac{\lambda_i}{D_{i1}} + \log \frac{mn}{\delta} + Q))$, and this completes the proof of Theorem 3.2.

□

Chapter 4

The Subgraph Containment Problem

Let us recall the standard models of random graphs. Let $n \in \mathbb{N}$, $M \in \left[\binom{n}{2}\right]$ and $p \in [0,1]$. $\mathcal{G}_{n,M}$ is the set of graphs chosen uniformly at random from the family of all subgraphs of K_n with exactly M edges and n nodes. It is the famous random graphs model invented by Erdős and Rényi [ER59]. $\mathcal{G}_{n,p}$ denotes the set of graphs obtained by adding edges of K_n with probability p, independently for each edge. This is the random graph model introduced by Gilbert [Gil59]. We may identify a graph in the model $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$ with the model itself. So we may say that a fixed graph H is a subgraph of $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$ meaning that H occurs as a subgraph in a random graph from $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$. The famous subgraph containment problem can be stated as follows.

Problem 4.1. (Subgraph Containment Problem in $\mathcal{G}_{n,M}$) Let G be a fixed graph. Does $\mathcal{G}_{n,M}$ contain a copy of G ?

In this chapter, we give a new proof of the containment of a fixed subgraph in $\mathcal{G}_{n,M}$ with the generalized bounded differences inequality (Theorem 2.10).

4.1 Basic Definitions and Known Results

Let e_G resp. v_G be the number of edges resp. nodes of G. Define

$$d(G) := \frac{e_G - 1}{v_G - 2} \text{ and } m(G) = \max\left\{d(H) : H \subseteq G, v_H \geqslant 3\right\}$$

$m(G)$ is a measure of graph density frequently appearing in the theory of random graphs [JŁR11]. A graph G is called *strictly K_2-balanced* if $d(H) < d(G)$, where $H \subset G, v_H \geqslant 3$. Complete graphs, complete bipartite graphs, cycles etc are examples of strictly K_2-balanced graphs. Let us use the notation "$G \subseteq \mathcal{G}_{n,M}$" resp. "$G \nsubseteq \mathcal{G}_{n,M}$", if G is a subgraph of $\mathcal{G}_{n,M}$ resp. is not a subgraph of $\mathcal{G}_{n,M}$. The same notation should be valid for the $\mathcal{G}_{n,p}$ model.

The theory of random graphs delivers an asymptotically precise answer to the subgraph containment problem. Janson, Rucinski and Łuczack [JŁR90] proved a remarkable and up to date best upper bound on the probability of non-existence of fixed G in $\mathcal{G}_{n,p}$:

Theorem 4.2. For every fixed graph G containing a cycle, there exists a constant $c_1 > 0$ and $n_1 \in \mathbb{N}$ such that for every $n \geqslant n_1, n \in \mathbb{N}$, and $n^{-\frac{1}{m(G)}} \leqslant p \leqslant 3n^{-\frac{1}{m(G)}}$

$$\mathbb{P}(\mathcal{G}_{n,p} \not\supset G) \leqslant e^{-c_1 n^2 p}. \tag{4.1}$$

The following theorem mentioned in Janson, Łuczak and Ruckinski [JŁR11] is the $\mathcal{G}_{n,M}$ counterpart of Theorem 4.2.

Theorem 4.3. For every fixed graph G containing a cycle, there exists a constant $c_1' > 0$ and $n_1' \in \mathbb{N}$ such that for every $n \geqslant n_1', n \in \mathbb{N}$, and $M = n^{2-\frac{1}{m(G)}}$, we have

$$\mathbb{P}(\mathcal{G}_{n,M} \not\supset G) \leqslant e^{-c_1' M} \tag{4.2}$$

4.2 Subgraph Containment via the BDI

We will prove an analogue of Theorem 4.3 with the generalized bounded differences inequality, which will become the basis for derandomizing Maker's random strategy in chapter 5. We generate a graph in the $\mathcal{G}_{n,M}$ model by choosing M edges randomly in an iterative way. Let $Z_1, Z_2, \cdots, Z_M$ be the random variables corresponding to M iterations. In the k-th iteration, Z_k is chosen from the set $E_k := \{E(K_n) \setminus (e_1, e_2, \cdots, e_{k-1})\}$ uniformly at random, where $(e_1, e_2, \cdots, e_{k-1})$ are the edges already chosen in the first $k-1$ iterations. So, the $E_1, E_2, \cdots, E_M$ are the edge sets depending on the $Z_1, Z_2, \cdots, Z_M$. The Z_i's are dependent random variables, of course. Note that every graph generated in this way has exactly n nodes and M edges, and appears with the *same* probability. Therefore, it is a random graph in the $\mathcal{G}_{n,M}$ model. The main result of this chapter is the following theorem.

Theorem 4.4. For every fixed, strictly K_2-balanced graph G containing a cycle, there exists constants $\tilde{c} > 0$ and $\tilde{n} \in \mathbb{N}$ such that for every $n \geqslant \tilde{n}$, $n \in \mathbb{N}$, and $M = \Theta(n^{2-1/m(G)})$, we have

$$\mathbb{P}(\mathcal{G}_{n,M} \not\supset G) \leqslant 2\exp\Big(-\tilde{c}M\Big(\frac{M}{n^{(2-1/m(G))}}\Big)^{2(e_G-1)}\Big). \tag{4.3}$$

We first prove a lemma which is essential in the proof of Theorem 4.4. It is a well known fact for $\mathcal{G}_{n,p}$ [FK15], but has to be established for $\mathcal{G}_{n,M}$ as well. We prove the lemma by the asymptotic equivalence of the model $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,M}$ for $p = \frac{M}{\binom{n}{2}}$.

Lemma 4.5. For every fixed graph G containing a cycle, let e_G resp. v_G be the number of edges resp. nodes of G. Let X_G be the number of copies of G in the random graph $\mathcal{G}_{n,M}$. Let $p = \frac{M}{\binom{n}{2}}$. Then, for sufficiently large M, we have

$$\mathbb{E}[X_G] = \Theta(n^{v_G} p^{e_G}) \tag{4.4}$$

Proof. Let N_G^n be the number of copies of G in the complete graph K_n. It is known that $N_G^n = \frac{\binom{n}{v_G} v_G!}{aut(G)}$, where $aut(G)$ is the cardinality of the automorphism group of G [FK15]. Since G is a fixed graph, we may view v_G and $aut(G)$ as constants not depending on n. Thus

$$N_G^n = \Theta(n^{v_G}). \tag{4.5}$$

Let S_G be the set of all copies of G in K_n. For each copy, $H \in S_G$, let

$$I_H = \begin{cases} 1 & \text{: if } H \text{ is a subgraph of } \mathcal{G}_{n,M} \\ 0 & \text{: if } H \text{ is not a subgraph of } \mathcal{G}_{n,M} \end{cases}$$

We first introduce some terminology. A *graph property* is a property that holds for a graph regardless of how its vertices are labeled. We say that a graph property is *monotonically increasing*, when the following assertion holds: if the property is true for the graph $G' = (V, E')$, then it holds for any graph $G'' = (V, E'')$ where $E' \subseteq E''$. Let $P(n, M)$ resp. $P(n, p)$ be the probability that a certain graph property $\mathcal{P}$ holds for a graph in $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$. Let $p^+ = (1 + \epsilon)M/\binom{n}{2}$ and $p^- = (1 - \epsilon)M/\binom{n}{2}$ for a constant $0 < \epsilon < 1$. Then Lemma 5.14 [MU17] says

$$P(n, p^-) - e^{-\mathcal{O}(M)} \leqslant P(n, M) \leqslant P(n, p^+) + e^{-\mathcal{O}(M)} \tag{4.6}$$

In our context, let $P(n, M)$ resp. $P(n, p)$ be the probability of the occurence of the subgraph H in $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$. The property of occurence of H in $\mathcal{G}_{n,\tilde{p}}$ resp. $\mathcal{G}_{n,\tilde{M}}$ is montononically increasing for any $\tilde{p} \in [0, 1]$ and any $\tilde{M} \in [\binom{n}{2}]$. For any $\tilde{p} \in [0, 1]$, we have $P(n, \tilde{p}) = \tilde{p}^{e(H)}$. By (4.6), we get

$$\left(\frac{(1-\epsilon)M}{\binom{n}{2}}\right)^{e_H} - e^{-\mathcal{O}(M)} \leqslant P(n, M) \leqslant \left(\frac{(1+\epsilon)M}{\binom{n}{2}}\right)^{e_H} + e^{-\mathcal{O}(M)} \tag{4.7}$$

For large value of M, the term $e^{-\mathcal{O}(M)}$ tends exponentially fast to 0, and is negligible. Hence we can choose constants $c_1, c_2 > 0$ such that for large M,

$$c_1 p^{e_H} \leqslant P(n, M) \leqslant c_2 p^{e_H} \tag{4.8}$$

We calculate $\mathbb{E}[X_G]$:

$$\mathbb{E}[I_H] = \mathbb{P}(I_H = 1) = P(n, M) \tag{4.9}$$

Since $X_G = \sum_{H \in S_G} I_H$, by linearity of expectation and (4.8), we get for sufficiently large M

$$\mathbb{E}[X_G] = N_G^n P(n, M) = \Theta(n^{v_G} p^{e_G}). \tag{4.10}$$

□

Let $s := c\mathbb{E}[X_G]$ for a suitably chosen constant $c > 0$ and let $\pi = (P_1, P_2, \cdots, P_s)$ be an arbitrary partition of the set of edges of K_n into sets P_i of size $|P_i| < n^2/2s$ for all $i = 1, 2, \cdots, s$. Two copies of G are called *π-disjoint* if for each index i at most one of them has an edge in P_i. Let $D_{\pi,G}(H)$ be the maximal number of pairwise π-disjoint copies of G in a graph H from $\mathcal{G}_{n,M}$.

Since $\mathcal{G}_{n,M}$ is generated by the random variables $Z_1, Z_2, \cdots, Z_M$, we may write

$$f(Z_1, Z_2, \cdots, Z_M) = D_{\pi,G} \tag{4.11}$$

for a suitable function f.

Let $z_1, z_2, \cdots, z_{k-1}$ be some arbitrary values the random variables $Z_1, Z_2, \cdots, Z_{k-1}$ can take. Let $z := (z_1, z_2, \cdots, z_{k-1})$. Set

$$ran_k(z) := \sup_{e,e' \in E_k} \left| \mathbb{E}\left[D_{\pi,G} \,|\, z, Z_k = e\right] - \mathbb{E}\left[D_{\pi,G} \,|\, z, z_{k-1}, Z_k = e'\right] \right| \tag{4.12}$$

ran_k is the range function w.r.t $D_{\pi,G}$ as defined in the context of the generalized BDI (chapter 1, 2.10). Further, let for $k = 1, 2, \cdots, M$

$$\hat{r}_k := \max_z ran_k(z), \text{ and } \bar{\mathrm{r}}^2 := \sum_{k=1}^{M} \hat{r}_k^2. \tag{4.13}$$

$\bar{\mathrm{r}}^2$ is the essential parameter for the generalized BDI. As we wish to prove Theorem 4.4 using the generalized BDI, we have to fix the value of $\bar{\mathrm{r}}^2$ in the following lemma.

Lemma 4.6. For every fixed graph G containing a cycle, let $\bar{\mathrm{r}}^2$ be the sum of squared maximum ranges of the function $D_{\pi,G}$ as in (4.13), where $D_{\pi,G}(H)$ be the maximal number of pairwise π-disjoint copies of G in a random graph H from $\mathcal{G}_{n,M}$. Then $\bar{\mathrm{r}}^2 \leqslant M$.

Proof. **Claim 1:** $\bar{\mathrm{r}}^2 \leqslant M$
This claim follows from the inequality $\hat{r}_k \leqslant 1$ for all $k = 1, 2, \cdots, M$, which on the other hand follows from $ran_k(z) \leqslant 1$ for all z, and this follows from

Claim 2: $\left| \mathbb{E}\left[D_{\pi,G} \,|\, z, Z_k = e\right] - \mathbb{E}\left[D_{\pi,G} \,|\, z, z_{k-1}, Z_k = e'\right] \right| \leqslant 1$ for all z.

For the proof of Claim 2, we estimate the change of $D_{\pi,G}$, when only one edge is changed. So let H be the graph associated to some arbitrary, but now fixed evaluations of the random variables $Z_1, Z_2, \cdots, Z_{k-1}, Z_k, Z_{k+1}, \cdots, Z_M$, and let $\overline{\mathrm{H}}$ be the graph associated to $Z_1, Z_2, \cdots, Z_{k-1}, \overline{\mathrm{Z}}_k, Z_{k+1}, \cdots, Z_M$, where $\overline{\mathrm{Z}}_k$ is an edge e' different from the edge $Z_k = e$.

Claim 3: $|D_{\pi,G}(H) - D_{\pi,G}(\overline{\mathrm{H}})| \leqslant 1$

For the proof, let l resp. l' be the cardinality of maximal sets of π-disjoint copies of G in H resp. $\overline{\mathrm{H}}$. So $D_{\pi,G}(H) = l$ and $D_{\pi,G}(\overline{\mathrm{H}}) = l'$ and Claim 3 says

$$|l - l'| \leqslant 1 \tag{4.14}$$

Assume for a moment that $l - l' \geqslant 2$. W.l.o.g let $l > l'$. Let $S_{max}(H)$ resp. $S_{max}(\overline{\mathrm{H}})$ be sets of maximal π-disjoint copies of G in H resp. $\overline{\mathrm{H}}$. Thus $|S_{max}(H)| = l$ and $|S_{max}(\overline{\mathrm{H}})| = l'$.

Case 1: There is a graph $G^* \in S_{max}(H)$ with $e \in G^*$.

By π-disjointness, $e \notin G'$ for all $G' \in S_{max}(H) \setminus \{G^*\}$, and all of these G' are copies of G in $\overline{\mathrm{H}}$. But their number is $l - 1 > l'$, in contradiction to the maximality of l'.

Case 2: $e \notin G^*$ for all $G^* \in S_{max}(H)$. Then obviously, all these G^* are copies of G in $\overline{\mathrm{H}}$, but their number is $l > l'$, a contradiction to the maximality of l'.

We have proved Claim 3.

Now we proceed to prove Claim 2.

To shorten the notation, let us set $g(\omega) := D_{\pi,G}(\omega)$, where $\omega \in \Omega$, and Ω is the sample space. Claim 3 says that $|g(\omega) - g(\omega')| \leqslant 1$, if ω and ω' differs only at one coordinate. The proof of Claim 2 is in fact a very general argrument relying on Claim 3. Let Σ_k be the σ-field generated by the random variables $Z_1, Z_2, \cdots, Z_k$, $k = 1, 2, \cdots, M$. The equivalence relation "$=_k$", where $\omega =_k \omega'$ iff ω and ω' are equal for the first k coordinates, generates a partition $\mathcal{P}_k$ of Ω, and

$$\mathbb{E}\left[g \,|\, \Sigma_k\right] = \sum_{C \in \mathcal{P}_k} \mathbb{E}\left[g \,|\, C\right] \mathbb{1}_C, \tag{4.15}$$

Let $C_0 \in \mathcal{P}_k$ resp. $C_1 \in \mathcal{P}_k$ be the partition sets with

$$C_0 = \{\omega \in \Omega \,:\, z, \omega_k = e\}, C_1 = \{\omega \in \Omega \,:\, z, \omega_k = e'\},$$

Now

$$\mathbb{E}\left[g \,|\, C_0\right] = \mathbb{E}\left[g \,|\, z, Z_k = e\right] \text{ and } \mathbb{E}\left[g \,|\, C_1\right] = \mathbb{E}\left[g \,|\, z, Z_k = e'\right], \tag{4.16}$$

Let $\varphi : C_0 \to C_1$ be the map defined by

$$\varphi(\omega_1, \cdots, \omega_{k-1}, e, \omega_{k+1}, \cdots, \omega_M) := (\omega_1, \cdots, \omega_{k-1}, e', \omega_{k+1}, \cdots, \omega_M).$$

φ is a bijective function and

$$\mathbb{P}(\varphi(\omega)) = \frac{\mathbb{P}(C_1)}{\mathbb{P}(C_0)} \mathbb{P}(\{\omega\}) \tag{4.17}$$

Note: In the current scenario, we choose $\omega_{k+1}, \omega_{k+2}, \cdots, \omega_M$ uniformly at random from the corresponding sets $E_{k+1}, E_{k+2}, \cdots, E_M$. Hence

$$\mathbb{P}(\{\omega_j\}) = \frac{1}{|E_j|} \tag{4.18}$$

for all $j \in \{k+1, k+2, \cdots, M\}$ devoid of $\omega \in C_0$ or $\omega \in C_1$. So

$$\mathbb{P}(C_0) = \sum_{\omega \in C_0} \mathbb{P}(\{\omega\}) = \sum_{\omega \in C_0} \prod_{j=k+1}^{M} \frac{1}{|E_j|} = \sum_{\omega \in C_1} \prod_{j=k+1}^{M} \frac{1}{|E_j|} = \mathbb{P}(C_1), \tag{4.19}$$

and

$$\begin{aligned}
|\mathbb{E}[g \mid C_0] - \mathbb{E}[g \mid C_1]| &= \left| \frac{1}{\mathbb{P}(C_0)} \sum_{\omega \in C_0} g(\omega)\mathbb{P}(\{\omega\}) - \frac{1}{\mathbb{P}(C_1)} \sum_{\omega \in C_1} g(\omega)\mathbb{P}(\{\omega\}) \right| \\
&= \left| \frac{1}{\mathbb{P}(C_0)} \sum_{\omega \in C_0} g(\omega)\mathbb{P}(\{\omega\}) - \frac{1}{\mathbb{P}(C_1)} \sum_{\omega \in C_0} g(\varphi(\omega))\mathbb{P}(\{\varphi(\omega)\}) \right| \\
&= \left| \frac{1}{\mathbb{P}(C_0)} \sum_{\omega \in C_0} g(\omega)\mathbb{P}(\{\omega\}) - \frac{1}{\mathbb{P}(C_0)} \sum_{\omega \in C_0} g(\varphi(\omega))\mathbb{P}(\{\omega\}) \right| \text{ (by 4.17)} \\
&\leqslant \sum_{\omega \in C_0} \underbrace{|g(\omega) - g(\varphi(\omega))|}_{\leqslant 1, \text{ by Claim 3}} \frac{\mathbb{P}(\{\omega\})}{\mathbb{P}(C_0)} \\
&\leqslant 1
\end{aligned}$$

and Claim 2 is proved. □

In the next lemma, we establish the relationship between $\mathbb{E}[D_{\pi,G}]$ and $\mathbb{E}[X_G]$. As we have an explicit formula for $\mathbb{E}[X_G]$ (see (4.10)), we shall use it to give a bound on $\mathbb{E}[D_{\pi,G}]$.

Lemma 4.7. For every fixed strictly K_2-balanced graph G containing a cycle, let $D_{\pi,G}(H)$ be the maximal number of pairwise π-disjoint copies of G in a random graph H from $\mathcal{G}_{n,M}$. Then for large M, $\mathbb{E}[D_{\pi,G}] = \Theta(\mathbb{E}[X_G])$.

Proof. Obviously $D_{\pi,G} \leqslant X_G$, so $\mathbb{E}[D_{\pi,G}] \leqslant \mathbb{E}[X_G]$. The lower bound for $\mathbb{E}[D_{\pi,G}]$ needs some work:

Let $Y_{\pi,G}$ be the number of non π-disjoint pairs of copies of G in $\mathcal{G}_{n,M}$. Clearly, $D_{\pi,G} \geqslant X_G - Y_{\pi,G}$. By definition, a pair (H', H'') is a non π-disjoint pair of copies of G if either $H' \cap H'' \neq \varnothing$ or $H' \cap H'' = \varnothing$ and H', H'' have an edge in at least one common P_i for some $i \in [s]$.

Case 1: $(H' \cap H'') \neq \varnothing$

By (4.5), for each proper subgraph K of G, $N_K^n = \Theta(n^{v_K})$ and $N_{G-K}^n = \Theta(n^{v_G - v_K})$. Let

$N_{1,K}$ be the number of pairs $(H^{'}, H^{''})$ of copies of graph G in the complete graph K_n with $(H^{'} \cap H^{''})$ isomorphic to K. Then

$$N_{1,K} = N_K^n N_{G-K}^n N_{G-K}^n = \Theta(n^{v_K} n^{2(v_G - v_K)}). \tag{4.20}$$

Case 2: $(H^{'} \cap H^{''}) = \varnothing$
As mentioned above, in this case for $(H^{'}, H^{''})$ to be a non π-disjoint pair of copies of G, $H^{'}, H^{''}$ both must have an edge in at least one common P_i for some $i \in [s]$. Let N_2 be the number of pairs $(H^{'}, H^{''})$ of copies of graph G in the complete graph K_n satisfying this condition. There are s partitions and at most $\frac{n^2}{2s}$ edges in each partition P_i, $i \in [s]$. We can choose edges $e, e^{'}$ in the same parition in $s\binom{n^2/2s}{2}$ ways. By (4.5), we have $N^n_{(H^{'}-e)} = N^n_{(H^{''}-e^{'})} = \Theta(n^{v_G - 2})$ copies of $H^{'} - e$, resp. $H^{''} - e^{'}$ in K_n, and hence $N_2 = s\binom{n^2/2s}{2} N^n_{(H^{'}-e)} N^n_{(H^{''}-e^{'})} = \Theta\left(s\binom{n^2/2s}{2} n^{2(v_G-2)}\right)$.

In our context, let $P(n, M)$ resp. $P(n, p)$ be the probability of the occurence of the non π-disjoint pair of copies of G, $(H^{'}, H^{''})$ in $\mathcal{G}_{n,M}$ resp. $\mathcal{G}_{n,p}$. The property of occurence of the non π-disjoint pair of copies of G, $(H^{'}, H^{''})$ in $\mathcal{G}_{n,\tilde{p}}$ resp. $\mathcal{G}_{n,\tilde{M}}$ is montononically increasing for any $\tilde{p} \in [0, 1]$ and any $\tilde{M} \in [\binom{n}{2}]$. In case 1, $(H^{'} \cap H^{''})$ is isomorphic to a graph K of G. So, for a $\tilde{p} \in [0, 1]$, we have $P(n, \tilde{p}) = \tilde{p}^{2e(G) - e(K)}$. By (4.6), we get

$$\left(\frac{(1-\epsilon)M}{\binom{n}{2}}\right)^{2e_G - e_K} - e^{-\mathcal{O}(M)} \leqslant P(n, M) \leqslant \left(\frac{(1+\epsilon)M}{\binom{n}{2}}\right)^{2e_G - e_K} + e^{-\mathcal{O}(M)} \tag{4.21}$$

For large value of M, the term $e^{-\mathcal{O}(M)}$ tends exponentially fast to 0, and is negligible. Hence we can choose constants $c_1, c_2 > 0$ such that:

$$c_1 p^{2e_G - e_K} \leqslant P(n, M) \leqslant c_2 p^{2e_G - e_K} \tag{4.22}$$

Similarly, in case 2, for a $\tilde{p} \in [0, 1]$, we have $P(n, \tilde{p}) = \tilde{p}^{2e(G)}$. By (4.6), we get

$$\left(\frac{(1-\epsilon)M}{\binom{n}{2}}\right)^{2e_G} - e^{-\mathcal{O}(M)} \leqslant P(n, M) \leqslant \left(\frac{(1+\epsilon)M}{\binom{n}{2}}\right)^{2e_G} + e^{-\mathcal{O}(M)} \tag{4.23}$$

For large value of M, the term $e^{-\mathcal{O}(M)}$ tends exponentially fast to 0, and is negligible. Hence we can choose constants $c_3, c_4 > 0$ such that:

$$c_3 p^{2e_G} \leqslant P(n, M) \leqslant c_4 p^{2e_G} \tag{4.24}$$

As G is strictly K_2 balanced, $\mathbb{E}[X_G] = o(\mathbb{E}[X_L])$ for any proper subgraph L of G (see Remark 3.17 and Section 3.2 [JŁR11]). Also note when $M = \Theta(n^{2 - \frac{1}{m(G)}})$ then $\mathbb{E}[X_G] = \Theta(1)$. Hence the expected number of non π-disjoint pairs is:

$$
\begin{aligned}
\mathbb{E}\left[Y_{\pi,G}\right] &= \sum_{K=H'\cap H''\neq\varnothing} N_{1,K}\, p^{2e_G-e_K} + \sum_{H'\cap H''=\varnothing} p^{2e_G} \;\; (K = H^{'} \cap H^{''} \neq \varnothing \text{ in all first sums}) \\
&= \sum_{K=H'\cap H''} N_{1,K}\, \mathcal{O}(p^{2e_G-e_K}) + N_2\, \mathcal{O}(p^{2e_G}) \text{ (by 4.22 and 4.24)} \\
&= \sum_{K=H'\cap H''} \mathcal{O}(n^{2v_G-v_K} p^{2e_G-e_K}) + \mathcal{O}\left(s\binom{n^2/2s}{2} n^{2(v_G-2)} p^{2e_G}\right) \\
&= \sum_{K=H'\cap H''} \mathcal{O}\left(\frac{\mathbb{E}\left[X_G\right]^2}{\mathbb{E}\left[X_K\right]}\right) + \mathcal{O}\left(s\binom{n^2/2s}{2} n^{2(v_G-2)} p^{2e_G}\right) \text{ (by Lemma 4.5)} \\
&= \sum_{K=H'\cap H''} \underbrace{\mathcal{O}\left(\frac{\mathbb{E}\left[X_G\right]^2}{\mathbb{E}\left[X_K\right]}\right)}_{\text{is negligible given } \mathbb{E}[X_G] = o(\mathbb{E}[X_K])} + \mathcal{O}\left(s\binom{n^2/2s}{2} n^{2(v_G-2)} p^{2e_G}\right) \qquad (4.25) \\
&= \mathcal{O}\left(s\binom{n^2/2s}{2} n^{2(v_G-2)} p^{2e_G}\right) \\
&= \mathcal{O}\left(s\binom{n^2/2s}{2} n^{2(v_G-2)} p^{2e_G}\right) \\
&= \mathcal{O}\left(\frac{n^4}{s} n^{2(v_G-2)} p^{2e_G}\right) \\
&= \mathcal{O}\left(\frac{n^{2v_G} p^{2e_G}}{s}\right) \\
&= \mathcal{O}\left(\frac{(\mathbb{E}\left[X_G\right])^2}{s}\right) \text{ (by Lemma 4.5)} \qquad (4.26)
\end{aligned}
$$

The term $\sum_{K=H'\cap H''} \mathcal{O}\left(\frac{\mathbb{E}[X_G]^2}{\mathbb{E}[X_K]}\right)$ in inequality (4.25) is negligible given $\mathbb{E}\left[X_G\right] = o(\mathbb{E}\left[X_K\right])$ and the sum is finite as K is fixed since H is fixed.

Hence, for sufficiently large c, using $s = c\mathbb{E}\left[X_G\right]$, we have

$$
\begin{aligned}
\mathbb{E}\left[D_{\pi,G}\right] &\geqslant \mathbb{E}\left[X_G\right] - \mathbb{E}\left[Y_{\pi,G}\right] \\
&\geqslant \mathbb{E}\left[X_G\right]\left(1 - \mathcal{O}\left(\frac{\mathbb{E}\left[X_G\right]}{s}\right)\right) \text{ (by 4.26)} \\
&= \mathbb{E}\left[X_G\right]\left(1 - \mathcal{O}\left(\frac{1}{c}\right)\right) \\
&\geqslant c_5\mathbb{E}\left[X_G\right] \text{ (for some constant } c_5 > 0) \qquad (4.27)
\end{aligned}
$$

□

Proof. (of Theorem 4.4)

We apply McDiarmid's generalized bounded differences inequality (Theorem 2.10) and get

$$\begin{aligned}
\mathbb{P}(\mathcal{G}_{n,M} \not\ni G) &= \mathbb{P}(X_G = 0) \\
&\leqslant \mathbb{P}(D_{\pi,G} = 0) \\
&\leqslant \mathbb{P}(|D_{\pi,G} - \mathbb{E}[D_{\pi,G}]| \geqslant \mathbb{E}[D_{\pi,G}]) \\
&\leqslant 2\exp\Big(\frac{-2(\mathbb{E}[D_{\pi,G}])^2}{\sum_{k=1}^{M} \hat{r}_k^2}\Big) \\
&\leqslant 2\exp\Big(\frac{-2(\mathbb{E}[D_{\pi,G}])^2}{M}\Big) \text{ (by Lemma 4.6)} \\
&\leqslant 2\exp\Big(-\frac{c_2(\mathbb{E}[X_G])^2}{M}\Big) \text{ (by Lemma 4.7)} \\
&\leqslant 2\exp\Big(-\frac{c_3 n^{2v_G} p^{2e_G}}{M}\Big) \text{ (for some constant } c_3 > 0 \text{ with Lemma 4.5)} \\
&\leqslant 2\exp\Big(-c_4 n^{2v_G - 4e_G} M^{2e_G - 1}\Big) \text{ (for some constant } c_4 > 0) \\
&= 2\exp\Big(-c_4 n^{2v_G - 4 - 4e_G + 4} M^{2e_G - 1}\Big) \\
&= 2\exp\Big(-c_4 n^{2(e_G - 1)(\frac{v_G - 2}{e_G - 1} - 2)} M^{2e_G - 1}\Big) \\
&\leqslant 2\exp\Big(-c_4 n^{2(e_G - 1)(\frac{1}{m(G)} - 2)} M^{2e_G - 1}\Big) \text{ (as } m(G) \geqslant \tfrac{e_G - 1}{v_G - 2}) \\
&= 2\exp\Big(-c_4 M\Big(\frac{M}{n^{(2 - 1/m(G))}}\Big)^{2(e_G - 1)}\Big).
\end{aligned}$$

□

Finally, we can prove Theorem 4.3:

Proof of Theorem 4.3:

Let's fix $M = n^{2-1/m(G)}$. In this case Theorem 4.4 implies Theorem 4.3.

□

Chapter 5

Derandomizing Maker's Strategy for the Maker-Breaker Subgraph Game

In this chaper we study the Maker-Breaker subgraph game $\mathsf{G}(G, n, b)$, where $b, n \in \mathbb{N}$ and present a deterministic Maker strategy by derandomizing the random strategy of Bednarska and Łuczack [BŁ00]. The game is played on the complete graph K_n on n vertices by two players, Maker and Breaker. In each round of the game, Maker chooses an edge of K_n, which has not been claimed previously and breaker responds by selecting at most b edges from K_n. The variable b denotes the *bias* of the game. The game ends if each of $\binom{n}{2}$ edges of K_n is claimed by either of the players. If the subgraph constructed during the game by the Maker contains a copy of G, then he wins, otherwise he looses and Breaker wins the game.

5.1 Previous Work

The game $\mathsf{G}(G, n, b)$ is a special case of positional games on graphs with G as a fixed graph. Games where the size of G depends upon n, like spanning trees, big stars, Hamiltonian cycles etc. have been extensively studied by Beck [Bec81; Bec94; Bec85; Bec82]. The game where $G = K_3$ was introduced by Erdős and Chvátal [CE78]. They showed that for $b \leqslant \sqrt{2n}$ Maker has a winning strategy, while for $b \geqslant 2\sqrt{n}$ Breaker has a winning strategy. For 40 years, there was no essential progress to close the gap between $\sqrt{2}\sqrt{n}$ and $2\sqrt{n}$. Recently, Glazik and Srivastav [GS18] gave a deterministic breaker strategy for $b \geqslant \sqrt{\frac{8}{3} + \epsilon}\sqrt{n}$, where $\epsilon > 0$ is small and fixed, introducing a new potential function approach, and almost matching the lower bound of Maker's win of $\sqrt{2}\sqrt{n} = \sqrt{\frac{8}{4}}\sqrt{n}$.

Beck [Bec94; Bec85; Bec82] also explored the relation of the bias of positional games and threshold properties of random graphs. Bednarska and Łuczack [BŁ00] in a breakthrough work used the results on random graphs to show that for the game $\mathsf{G}(G, n, b)$, a random strategy for the maker is asymptotically optimal in n, but to find the right matching constants is a major open problem in this area. They modified the above game so that Maker cannot see Breaker moves but Breaker has all the information about the moves of the Maker. They proved that even with this weak strategy of Maker, where Maker chooses his edges uniformly at random among all edges which have not been claimed by him so far, Maker wins with a probability of at least $\frac{1}{3}$. The main result in [BŁ00] is the following theorem:

Theorem 5.1. For every graph G which contains at least 3 non-isolated vertices, there exist positive constants c_0, n_0 such that for every $n \geqslant n_0$ and $b \leqslant c_0 n^{1/m(G)}$, $n, b \in \mathbb{N}$, Maker has a winning strategy for the game $\mathbf{G}(G, n, b)$.

The above result by Bednarska and Łuczack is based on the following theorem mentioned in Janson, Łuczak and Ruckinski [JŁR11].

Theorem 5.2. For every graph G containing a cycle, there exists a constant $c_1' > 0$ and $n_1' \in \mathbb{N}$ such that for every $n \geqslant n_1'$, $n \in \mathbb{N}$, and $M = n^{2-\frac{1}{m(G)}}$, we have

$$\mathbb{P}(\mathcal{G}_{n,M} \nsupseteq G) \leqslant e^{-c_1' M} \tag{5.1}$$

5.2 Proof Strategy and Main Theorem

We briefly describe the proof strategy adopted by Bednarska and Łuczack [BŁ00] to prove the existence of a random winning strategy for Maker for $b \leqslant c_0 n^{\frac{1}{m(G)}}$.

1. Assume that Maker has no information about the moves of the Breaker, while Breaker has full information about moves of the Maker. Let $e_1, e_2, \cdots, e_{k-1}$ be the edges chosen by the Maker in first $k-1$ rounds. In the k-th move, Maker chooses edges uniformly at random from the set $E_k = \{E(K_n) \setminus \{e_1, e_2, \cdots, e_{k-1}\}\}$ where $E(K_n)$ is the set of edges for K_n.
2. Maker may choose a Breaker edge, in which case Maker looses his chance, and this instance is referred as *failure*.
3. Let S_d be the set of edge-disjoint copies of G in $\mathcal{G}_{n,M}$ where $M = 2n^{2-\frac{1}{m(G)}}$. Let D be the event, "$|S_d| \geqslant \delta' M$", where $\delta' > 0$. In Lemma 4 [BŁ00], it was established that $\mathbb{P}(D) \geqslant \frac{2}{3}$.
4. For the random strategy to be successful, it must be shown that the number of failures f is upper bounded by $\delta' M$. Let us denote this event by F. In fact, it was shown that $\mathbb{P}(F) \geqslant \frac{2}{3}$. Hence $\mathbb{P}(D \cap F) \geqslant \frac{1}{3}$.

Now let us now move to our derandomized strategy.

Since $\mathbb{P}(D \cap F) \geqslant \frac{1}{3}$, Maker's random strategy is successful with positive probability, and we would like to derandomize it. But the critical point here is that in the proof of Bednarska and Łuczack [BŁ00], there is no derandomized version of Theorem 5.2. Our approach will be to use the new proof of Theorem 5.2 for graphs which are strictly K_2-balanced which we estbalished in chapter 4 with the generalized BDI, and then invoke the derandomized BDI (Theorem 3.6). Here are the steps:

1. The π-disjointness property implies edge-disjointness. Let S_π be the set of π-disjoint copies of G in $\mathcal{G}_{n,M}$ where $M = 2n^{2-\frac{1}{m(G)}}$. Let D' be an event, "$|S_\pi| \geqslant \delta' M$" where $\delta' > 0$. We analyze the event D' using generalized bounded differences inequality (2.10) and construct pessimistic estimators $(V_l)_l$ as defined in the chapter 3.

2. Let us define $F^{'}$ as the event "$f \leqslant \frac{3}{4}\delta^{'}M$" where f is the number of failures. We analyze the event $F^{'}$ using generalized bounded differences inequality (2.10) and construct pessimistic estimators $(U_l)_l$ as defined in chapter 3.

3. Using Proposition 3.10, we define the pessimistic estimators for the event $D^{'} \cap F^{'}$. Using the DERAND algorithm, we get our desired derandomized Maker winning strategy.

Let $Z_1, Z_2, \cdots, Z_M$ be the random variables corresponding to the M moves of the Maker. In the k-th iteration, Z_k is chosen from the set $E_k := \{E(K_n) \setminus \{e_1, e_2, \cdots, e_{k-1}\}\}$ uniformly at random, where $e_1, e_2, \cdots, e_{k-1}$ are the edges already chosen in the first $k-1$ iterations. So, the $E_1, E_2, \cdots, E_M$ are the edge sets depending on the random variables $Z_1, Z_2, \cdots, Z_M$. The Z_i's are dependent random variables, of course. Note that every graph generated in this way has exactly n nodes and M edges, and appears with the *same* probability. Therefore, the graph generated by the Maker moves is a random graph in the $\mathcal{G}_{n,M}$ model.

We want to show that the Maker graph contains a copy of G with positive probability, and will derandomize this probabilistic statement. Let $\mathcal{A}$ be the event that the Maker graph contains a copy of G. Actually, we consider a much stronger event in the upcoming analysis, namely that the number of edge-disjoint copies of G in the Maker graph is at least αM for some suitable constant $\alpha > 0$. Let $\mathcal{B}_\alpha$ be this event. Then, $\mathcal{B}_\alpha$ enforces $\mathcal{A}$. We proceed to show that

$$\mathbb{P}(\mathcal{B}_\alpha) \geqslant \frac{1}{3}, \tag{5.2}$$

for some α, and thereafter enter derandomization.

Now as we want to apply the derandomized BDI , we consider two other events enforcing $\mathcal{B}_\alpha$. The first event $D^{'}$ is "$D_{\pi,G} \geqslant \delta^{'}M$" where constant δ', $\delta' > 0$, was fixed in the proof by Bednarska and Łuczack [BŁ00]. The second event $F^{'}$ is "$f \leqslant \frac{3}{4}\delta^{'}M$".

Remark 5.3. Now, if the number of failures f is at most $\frac{3}{4}\delta^{'}M$, then among at least $\delta^{'}M$ edge-disjoint copies of G in $\mathcal{G}_{n,M}$ at most $\frac{3}{4}$-th of them are affected by failures, and $\frac{1}{4}\delta^{'}M$ edge-disjoint copies of G with Maker edges, which were never claimed by Breaker remain. Thus $\mathcal{B}_\alpha$ hold with $\alpha = \frac{\delta^{'}}{4}$, enforcing the desired event $\mathcal{A}$, which represents Maker's win. In the subesquent corollary 5.7, we will prove that $D^{'} \cap F^{'}$ enforces $\mathcal{B}_{\delta'/4}$ with probability at least $\frac{1}{3}$.

First, we count the number of edge-disjoint copies of G with the BDI.

Theorem 5.4. For every fixed strictly K_2-balanced graph G containing a cycle let S_d be the set of edge-disjoint copies of G in $\mathcal{G}_{n,M}$. Then there exists $\tilde{n} \in \mathbb{N}$ such that for every $n \geqslant \tilde{n}$,

$n \in \mathbb{N}$, a constant $\delta^{'}$, $\delta^{'} > 0$ and $M = 2n^{2-1/m(G)}$, we have

$$\mathbb{P}(|S_d| \geqslant \delta^{'} M) \geqslant \frac{2}{3} \tag{5.3}$$

Proof. Let $p := \frac{M}{\binom{n}{2}}$. By Lemma 4.5,

$$\begin{aligned}
\mathbb{E}\left[X_{G}\right] &\geqslant c_1 n^{v_G - 2e_G} M^{e_G} \text{ (for some constant } c_1 > 0) \\
&= c_1 n^{v_G - 2 - 2e_G + 2} M^{e_G} \\
&= c_1 n^{(e_G - 1)(\frac{v_G - 2}{e_G - 1} - 2)} M^{e_G} \\
&= c_1 n^{(e_G - 1)(\frac{1}{m(G)} - 2)} M^{e_G} \left(m(G) = \frac{e_G - 1}{v_G - 2} \right) \\
&= c_1 \frac{1}{(n^{2 - \frac{1}{m(G)}})^{e_G - 1}} M^{e_G} \\
&= c_2 M \text{ (for some constant } c_2 > 0)
\end{aligned} \tag{5.4}$$

Let $D^{'}$ be the event $D_{\pi,G} \geqslant \delta' M$. Now we apply McDiarmid's generalized bounded differences inequality (Theorem 2.10) and get:

$$\begin{aligned}
\mathbb{P}(D_{\pi,G} \leqslant \delta' M) &= \mathbb{P}(D_{\pi,G} - \mathbb{E}\left[D_{\pi,G}\right] \leqslant \delta' M - \mathbb{E}\left[D_{\pi,G}\right]) \\
&\leqslant \mathbb{P}(D_{\pi,G} - \mathbb{E}\left[D_{\pi,G}\right] \leqslant \delta' M - \tilde{c}\mathbb{E}\left[X_G\right]) \text{ (by Lemma 4.7)} \\
&\leqslant \mathbb{P}(D_{\pi,G} - \mathbb{E}\left[D_{\pi,G}\right] \leqslant \delta' M - c_3 M) \text{ (by 5.4)} \\
&\leqslant \exp\Big(-\frac{c_4 M^2}{\sum_{k=1}^{M} \hat{r}_k^2}\Big) \text{ (by Theorem 2.10)} &\quad (5.5) \\
&= \exp\Big(-c_4 M\Big) \text{ (by Lemma 4.6)} &\quad (5.6) \\
&\leqslant \frac{1}{3}, &\quad (5.7)
\end{aligned}$$

where $\tilde{c}, c_3, c_4 > 0$ are suitable constants where $c_3 > \delta'$.

Since $|S_d| \geqslant D_{\pi,G}$, we are done.

□

Remark 5.5. For $b \leqslant \frac{1}{8}\delta' n^{1/m(G)}$, $M = 2n^{2-\frac{1}{m(G)}}$ and sufficiently large n, after M rounds at most $\frac{\delta^{'}}{2}\binom{n}{2}$ edges are either claimed by Maker or Breaker.

Recall that f is the number of failures and $F^{'}$ is the event "$f \leqslant \frac{3}{4}\delta^{'} M$" for $\delta^{'} > 0$. We now show that the probability of event $F^{'}$ is at least $\frac{2}{3}$. It was established in [BŁ00] that the probability of failure is at most $\delta^{'}/2$. Let $X_1, X_2, \cdots, X_M$ be the 0-1 random variables which represent whether the i-th round by the Maker resulted in a failure, $i \in [M]$. Let B_i be the set of edges acquired by Breaker in the first $i-1$ rounds, $i \in [M]$. So,

$$X_i = \begin{cases} 1 & \text{if } Z_i \in B_i \\ 0 & \text{otherwise,} \end{cases}$$

$i \in [M]$. Clearly, $X_1, X_2, \cdots, X_M$ are dependent random variables and $f = X_1 + X_2 + \cdots + X_M$. Since, the probability of failure in an round is at most $\delta^{'}/2$, we have $\mathbb{E}[f] \leqslant \delta^{'} M/2$.

By the generalized bounded differences inequality (Theorem 2.10), we get

Lemma 5.6. Under the assumptions of Theorem 5.4, given $b \leqslant \frac{1}{8}\delta^{'} n^{1/m(G)}$ and constant $\delta^{'}$, $\delta^{'} > 0$, we have $\mathbb{P}(F^{'c}) \leqslant \frac{1}{3}$.

Proof. Let $x_1, x_2, \cdots, x_{k-1}$ be arbitrary but fixed values of the random variables $X_1, X_2, \cdots, X_{k-1}$. Let $x := (x_1, x_2, \cdots, x_{k-1})$. Set

$$ran_k(x) := \sup_{e, e^{'} \in E_k} \left| \mathbb{E}[f \mid x, X_{k_e}] - \mathbb{E}\left[f \mid x, X_{k_{e^{'}}}\right] \right| \tag{5.8}$$

where $X_k = X_{k_e}, X_{k_{e^{'}}}$ if $Z_k = e$, $Z_k = e^{'}$ respectively. ran_k is the range function w.r.t f as defined in the context of the generalized BDI (chapter 1, 2.10). Further, let for $k = 1, 2, \cdots, M$

$$\hat{r}_k := \max_x ran_k(x), \text{ and } \bar{\text{r}}^2 := \sum_{k=1}^{M} \hat{r}_k^2. \tag{5.9}$$

Let f resp. $f^{'}$ be the failures given $x = (x_1, x_2, \cdots, x_{k-1}, x_k, x_{k+1}, \cdots, x_M)$ resp. $x^{'} = (x_1, x_2, \cdots, x_{k-1}, x_k^{'}, x_{k+1}, \cdots, x_M)$ where x be some arbitrary but now fixed values of random variables $X_1, X_2, \cdots, X_M$ where $X_k = X_{k_e}$ and $x^{'}$ where $X_k = X_{k_{e^{'}}}$. Hence $|f - f^{'}| = |X_{k_e} - X_{k_{e^{'}}}| \leqslant 1$. Now as shown in the proof of Claim 2 of Lemma 4.6, we can establish that $ran_k(x) \leqslant 1$. Hence $\bar{\text{r}}^2 \leqslant M$. So,

$$\begin{aligned} \mathbb{P}(f \geqslant \frac{3}{4}\delta^{'} M) &= \mathbb{P}(f - \mathbb{E}[f] \geqslant \frac{3}{4}\delta^{'} M - \mathbb{E}[f]) \\ &= \mathbb{P}(f - \mathbb{E}[f] \geqslant \frac{3}{4}\delta^{'} M - \frac{1}{2}\delta^{'} M) \\ &\leqslant \mathbb{P}\left(f - \mathbb{E}[f] \geqslant \underbrace{\frac{\delta^{'} M}{4}}_{=:\text{t}} \right) &\quad (5.10) \\ &\leqslant e^{-\frac{2t^2}{\bar{\text{r}}^2}} &\quad (5.11) \end{aligned}$$

$$\leqslant e^{-\frac{(\delta')^2 M}{8}} \tag{5.12}$$

$$\leqslant \frac{1}{3}, \text{ for sufficiently large } M. \tag{5.13}$$

□

Corollary 5.7. Under the assumptions of Theorem 5.4, given $b \leqslant \frac{1}{8}\delta' n^{1/m(G)}$ and constant δ', $\delta' > 0$, we have $\mathbb{P}(D' \cap F') \geqslant \frac{1}{3}$, and $\mathbb{P}(\mathcal{A}) \geqslant \frac{1}{3}$, so with probability at least $\frac{1}{3}$ the Maker graph without failure edges contains a copy of G.

Proof. By (5.13), Theorem 5.4 and Remark 5.3 we are done. □

Thus we can construct a derandomized winning strategy for Maker using the derandomization framework developed for the generalized BDI.

5.3 The Derandomized Maker Strategy

We construct pessimistic estimators $\mathcal{V} = (V_l)_l$, $l \in \{0, 1, 2, \cdots, M\}$, for the event D' and $\mathcal{U} = (U_l)_l$, $l \in \{0, 1, 2, \cdots, M\}$, for the event F' according to the Definition 3.17 in chapter 3. Using Proposition 3.10, we can then define the pessimistic estimators $\mathcal{W} = (W_l)_l$, where $W_l = U_l + V_l$, $l \in \{0, 1, 2, \cdots, M\}$, for the event $D' \cap F'$. Maker plays according to the following algorithm. Maker chooses an edge in each iteration which mimimizes the value of the pessimistic estimator $(W_l)_l$.

Algorithm 1: DERANDOMIZED MAKER STRATEGY

Input : The event $D' \cap F'$ representing Maker's win and pessimistic estimators $\mathcal{W} = (W_l)_l$ for the event $D' \cap F'$ and $M := 2n^{2-\frac{1}{m(G)}}$

▻ For $l = 1, \cdots, M$, do : If the edges corresponding $z_1, \cdots, z_{l-1}$ are fixed, choose Z_l as the minimizer of the function $\omega \to W_l(z_1, \ldots, z_{l-1}, \omega)$.

Output: A vector $z \in D' \cap F'$

Theorem 5.8. Let G be a fixed and strictly K_2-balanced graph with at least 3 non-isolated vertices and a cycle.

(i) There exist positive constants $\tilde{c}_0, n_0$ such that for every $n \geqslant n_0$ and $b \leqslant \tilde{c}_0 n^{1/m(G)}$, Maker wins the game $\mathbf{G}(G, n, b)$ playing according to Algorithm 1.

(ii) In each round of the game, Maker needs $\mathcal{O}(2^{n^{2-\frac{1}{m(G)}}\log(n)})$ time to compute the pessimistic estimator.

Proof. We have two events $F^{'}$ and $D^{'}$. Let, in view of Theorem 3.6, $F^{'}$ be the first event and $D^{'}$ be the second event. So, $m = 2$. Further, since the A_k's are subsets of $E(K_n)$, $N \leqslant \frac{n^2}{2}$. The sample space in Theorem 3.6 is of the form $\Omega = \prod_{k=1}^{n'} A_k$ for some $n' \in \mathbb{N}$. In our context, the game is played for M rounds, so $n' = M$. For the first event $F^{'}$, according to the notation of Theorem 3.6 as $\hat{r}_k \leqslant 1$, $d_{1k} = 1$ for all $k \in [M]$, $D_{11} = d_{11}^2 + \cdots + d_{1M}^2 = M$, and by (5.10), $\lambda_1 \leqslant M$. Furthermore, f is sum of M indicator random variables so $f \leqslant M$ as well. Since f is linear and has M terms, we can compute $\mathbb{E}\left[f \mid x_1, x_2, \cdots, x_l\right]$, where $X_i = x_i$ for $i \in [l]$, in $\mathcal{O}(M)$ time, for any choice of $l \leqslant M$, $z_1, \cdots, z_l$ and corresponding $x_1, \cdots, x_l$.

We now consider the second event $D^{'}$. According to the bound used in the generalized BDI estimation (5.5), $\lambda_2 = \Theta(M)$. The parameter D_{21} as defined in the context of the derandomized generalized BDI (Theorem 3.6), is $D_{21} = \hat{r}_1^2 + \cdots + \hat{r}_M^2$, and by (5.6), $D_{21} = M$. Furthermore,

$$\|D_{\pi,G}\|_{\infty} = \max_{H \in \mathcal{G}_{n,M}} D_{\pi,G}(H) = \mathcal{O}(M) \tag{5.14}$$

Here is an argument for (5.14). Let us fix an $H \in \mathcal{G}_{n,M}$. The partition π has at most s partition sets, where $s = c\mathbb{E}\left[X_G\right]$. Since two π-disjoint copies of partition set cannot share an edge in a partition set, there are at most s π-disjoint copies of G in H.

By (4.10) and by argument similar to given in (5.4), we have

$$\mathbb{E}\left[X_G\right] = \Theta(n^{v_G} p^{e_G}) = \mathcal{O}(M).$$

So,

$$D_{\pi,G}(H) = \mathcal{O}(M), \text{ and } \|D_{\pi,G}\|_{\infty} = \mathcal{O}(M).$$

Let Q_0 be the running time to compute the conditional expectations $\mathbb{E}\left[D_{\pi,G} \mid z_1, z_2, \cdots, z_l\right]$, $l \in [M]$ and set $Q = \max(\mathcal{O}(M), Q_0)$. Then the overall running time according to Theorem 3.6 is

$$\mathcal{O}\Big(mn'N\Big[\max_{i=1,2} P_i \frac{\lambda_i}{D_{i1}} + \log(\frac{mn'}{\delta}) + Q\Big]\Big). \tag{5.15}$$

Further, from the above

$$\frac{\lambda_1}{D_{11}} \leqslant \frac{M}{D_{11}} = \frac{M}{M} = 1.$$

And,

$$\frac{\lambda_2}{D_{21}} = \frac{\theta(M)}{M} = \mathcal{O}(1).$$

Thus, (5.15) becomes

$$\mathcal{O}\Big(n^4\Big[n^{2-\frac{1}{m(G)}} + \log(n^2) + Q\Big]\Big) = \mathcal{O}\Big(n^{6-\frac{1}{m(G)}} + n^4 Q\Big) \tag{5.16}$$

We now compute Q_0, and hence Q.

Correctness : By corollary 5.7, and assuming Q being calculated, the generalized BDI delivers a vector $z \in F^{'} \cap D^{'}$. By remark 5.3, this is Maker's win.

Running Time : We now upper bound the computation time time Q_0 for the conditional expectation in Theorem 3.6. This needs some work for the event $D^{'}$.

Claim: The conditional expectations $\mathbb{E}\left[D_{\pi,G} \,|\, z_1, z_2, \cdots, z_l\right]$, can be computed for each $l \in [M]$ in $\mathcal{O}(2^{\frac{1}{m(G)}n^{2-\frac{1}{m(G)}}\log(n)})$ time. Thus, $Q_0 = \mathcal{O}(2^{\frac{1}{m(G)}n^{2-\frac{1}{m(G)}}\log(n)})$. So $Q = \mathcal{O}(2^{\frac{1}{m(G)}n^{2-\frac{1}{m(G)}}\log(n)})$, and this is also the overall running time according to (5.16).

We define a *configuration* as a set of M Maker edges from K_n. Let us first compute the number of configurations C after l rounds of the game. Set $u := \binom{n}{2} - l$. Then u is the number of edges left for Maker after l rounds, including failures. As l out of the M edges are already chosen by Maker, he can choose rest $M - l$ edges from u in upcoming rounds, and there are $\binom{u}{(M-l)}$ such choices. So,

$$C = \binom{u}{M-l} = \frac{u!}{(u-M+l)!M-l!} \tag{5.17}$$

We can easily see that the value of C is monotonically decreasing in l. So

$$\begin{aligned}
C = \frac{u!}{u-(M-l)!(M-l)!} &\leqslant \frac{\binom{n}{2}!}{\left(\binom{n}{2}-M\right)!M!} \\
&\leqslant \left(\frac{n^2}{2}\right)^M \Big/ M! \\
&\leqslant \left(\frac{n^2}{2}\right)^M \frac{e^M}{eM^M} \text{ (by Stirling's formula)} \\
&\leqslant \left(\frac{2n^2}{M}\right)^M \\
&\leqslant (n^{\frac{1}{m(G)}})^M \\
&= \mathcal{O}\left(2^{\frac{1}{m(G)}n^{2-\frac{1}{m(G)}}\log(n)}\right)
\end{aligned} \tag{5.18}$$

Now let's choose an arbitrary, but now fixed configuration. We count copies of G upto isomorphism in the subgraph formed by the configuration. We can do this in $\mathcal{O}(n^{v_G}e_G^2)$ time but since G is fixed, the running time is $\mathcal{O}(n^{v_G})$.

Let L_G be the set of copies of G in the aforementioned configuration computed so far. Recall from Chapter 4 that $\pi = (P_1, P_2, \cdots, P_s)$ be an arbitrary parition of the edges of the complete graph K_n. For each graph $K \in L_G$, let $\mathcal{P}_K$ be the set of partition sets of π, which contain some edge of K. Let us consider a graph Γ with vertices as elements of L_G, and an edge $\{K, K^{'}\}$, $K, K^{'} \in L_G$ exists, if $\mathcal{P}_K \cap \mathcal{P}_{K'} \neq \varnothing$. The size of a maximum independent set in Γ is equal $D_{\pi,_G}$. We can find the maximum independent set in $\mathcal{O}(2^{n^{2-\frac{1}{m(G)}}})$ time.

So,

$$
\begin{aligned}
Q_0 &\leqslant 2^{\frac{1}{m(G)} n^{2-\frac{1}{m(G)}} \log(n)} . n^{v_G} . 2^{n^{2-\frac{1}{m(G)}}} \\
&= 2^{\frac{1}{m(G)} n^{2-\frac{1}{m(G)}} \log(n) + v_G \log(n) + n^{2-\frac{1}{m(G)}}} \\
&= 2^{\frac{1}{m(G)} n^{2-\frac{1}{m(G)}} \log(n)(1+o(1))}
\end{aligned}
$$

leading to $Q = \max(M, Q_0) = 2^{\frac{1}{m(G)} n^{2-\frac{1}{m(G)}} \log(n)(1+o(1))}$ and with (5.16) the overall running time is $2^{\frac{1}{m(G)} n^{2-\frac{1}{m(G)}} \log(n)(1+o(1))}$.

□

Remark 5.9. In the proof of Theorem 5.8, i.e. the computation of the pessimistic estimator, we saw that the superexponential time complexity stems from three hard problems: Counting subgraphs, the subgraph isomorphism problem and the maximum independent set problem.

Open Question: Is there a *polynomial-time deterministic* Maker strategy for the subgraph problem which is asymptotically optimal for $b = \mathcal{O}(n^{\frac{1}{m(G)}})$?

Chapter 6

Derandomized Hybrid Algorithm for the Set-Multicover Problem

Set cover is an important problem in field of combinatorial optimization which led to the development of fundamental techniques in the field of approximation algorithms. We will study generalized version SET b-MULTICOVER, of the set cover problem. The problem can be naturally expressed in the context of hypergraphs. A hypergraph $\mathcal{H} = (V, \mathcal{E})$ consists of a finite set V and a set $\mathcal{E}$ of subsets of V. We call elements of V as vertices and the elements of $\mathcal{E}$ (hyper-)edges. We fix $n := |V|$, $m := |\mathcal{E}|$.

The *edge size* is the cardinality of the edge. Let l be the maximum edge size and let Δ be the maximum vertex degree, where the degree of a vertex is the number of edges containing that vertex. For $v \in V$, let $\mathcal{E}(v)$ be the set of hyperedges containing v. For $b \in \mathbb{N}$, a set b-multicover in $\mathcal{H}$ is a set of edges $C \subseteq \mathcal{E}$ such that every vertex in V belongs to at least b edges in C. SET b-MULTICOVER is the problem of finding a set b-multicover of minimum cardinality. Define $\delta := \Delta - b + 1$.

6.1 Previous Work

For a minimization problem, a polynomial-time algorithm A has an approximation ratio $\alpha \geqslant 1$, or is an α-approximation, if for all problem instances I it holds that the set b-multicover $A(I)$ returned by A satisfies $|A(I)|/\mathrm{Opt} \leqslant \alpha$ where Opt is the optimum value of the problem instance. Hall and Hochbaum [HH86] gave a δ-approximation algorithm with running time of $\mathcal{O}(m \cdot \max\{n, m\})$. Peleg, Schechtman and Wool [PSW97], broke the δ barrier and gave an approximation algorithm with an approximation ratio of $\delta \cdot (1 - (\frac{c}{n})^{\frac{1}{\delta}})$, where $c > 0$ is a constant bounded from below roughly by $2^{-2^{50}}$ and hence the approximation ratio tends to δ as n grows. The authors conjectured that unless $\mathcal{P} = \mathcal{NP}$, there does not exist a polynomial time approximation algorithm with a *constant* approximation ratio smaller than δ.

El Ouali et al. [EMS14] settled this conjecture for $l \in \max\{(nb)^{\frac{1}{5}}, n^{\frac{1}{4}}\}$. They presented a hybrid randomized algorithm, combining LP-based randomized rounding and a greedy repairing, if the randomized solution is infeasible. It was shown that the algorithm achieves an approximation ratio of $\delta \cdot \left(1 - \frac{11(\Delta - b)}{72l}\right)$ with constant probability for $l \in \max\{(nb)^{\frac{1}{5}}, n^{\frac{1}{4}}\}$.

They improve upon results of Peleg et al. [PSW97]: For any l satisfying $l = \mathcal{O}\left((\Delta - b) \cdot n^{\frac{1}{\delta}}\right)$ approximation ratio is at most the ratio of Peleg et al., and it is better, smaller the l is. In the important case of l being a constant, the ratio is $\delta \cdot (1 - c)$, $c \in (0, 1)$ a constant *independent* of n.

El Ouali et al. use two different methods to analyze the algorithm and hence get two bounds on l. They use independent bounded differences inequality and Angluin-Valiant bound for analysis of the algorithm where $l \leqslant (nb)^{\frac{1}{5}}$. Hence we use the derandomization framework described in the Theorem 3.14 to get deterministic counterpart of randomized algorithm.

6.2 Linear Program and the Hybrid Randomized Algorithm

We will define the linear program for the SET b-MULTICOVER. Let $A = (a_{ij})_{(i,j)\in[n]\times[m]}$ be its incidence matrix of hypergraph $\mathcal{H}$, where $a_{ij} = 1$ if vertex i is contained in edge j, and $a_{ij} = 0$ otherwise. Then the (relaxed) linear programming formulation of SET b-MULTICOVER is the following:

$$
\begin{aligned}
(b - LP) \qquad & \min \sum_{j=1}^{m} x_j \\
& \sum_{j=1}^{m} a_{ij} x_j \geqslant b \quad \text{for all } i \in [n] \\
& x_j \in [0, 1] \quad \text{for all } j \in [m].
\end{aligned}
$$

Let Opt* be the value of an optimal solution to b−LP.

We first briefly describe the algorithm of El Ouali et al.

Algorithm 2: SET b-MULTICOVER

Input : $b \in \mathbb{N}$, $\epsilon \in (0,1)$, a hypergraph $\mathcal{H} = (V, \mathcal{E})$ with maximum degree Δ

Output: A set b-multicover C

1. Initialize $C := \varnothing$ and set $\lambda := (1-\epsilon)\delta$.
2. Obtain an optimal solution $x^* \in [0,1]^m$ by solving the LP relaxation of set multicover.
3. Set $S_0 := \{E_j \in \mathcal{E} \mid x_j^* = 0\}$, $S_{\geqslant} := \{E_j \in \mathcal{E} \mid x_j^* \geqslant \frac{1}{\lambda}\}$ and $S_< := \{E_j \in \mathcal{E} \mid 0 \neq x_j^* < \frac{1}{\lambda}\}$.
4. Take all edges of $S_{\geqslant}$ into the cover C and set $\mathcal{E} = \mathcal{E} \backslash S_0$.
5. (Randomized Rounding) For all edges $E_j \in S_<$ include the edge E_j in the cover C, independently for all such E_j, with probability λx_j^*.
6. (Repairing) Repair the cover C (if necessary) as follows: Include arbitrary edges from $S_<$, incident to vertices not covered by b edges, to C until all vertices are covered by at least b edges.
7. Return the cover C.

The following lemma shows that all vertices are almost covered after step 4 of Algorithm 2.

Lemma 6.1. (Lemma 2.1, [PSW97]) *Let $0 < \epsilon \leqslant \frac{1}{4}$ and $b, d, \Delta \in \mathbb{N}$ with $2 \leqslant b \leqslant d-1 \leqslant \Delta - 1$. Let $\lambda = (1-\epsilon)\delta$ and let $x_j \in [0,1]$, $j \in [d]$, such that $\sum_{j=1}^{d} x_j \geqslant b$. Then at least $b-1$ of the x_j fulfill the inequality $x_j \geqslant \frac{1}{\lambda}$.*

For the derandomization we need some more details of the analysis of the randomized algorithm, in particular we need to identify the events of interest. Let $X_1, ..., X_m$ be $\{0,1\}$-random variables defined as follows:

$$X_j = \begin{cases} 1 & \text{if the edge } E_j \text{ was picked into the set cover before repairing} \\ 0 & \text{otherwise.} \end{cases}$$

Note that the $X_1, ..., X_m$ are independent random variables. For all $i \in [n]$ we define the $\{0,1\}$-random variables Z_i as follows:

$$Z_i = \begin{cases} 1 & \text{if the vertex } v_i \text{ is covered by at least } b \text{ edges before repairing} \\ 0 & \text{otherwise.} \end{cases}$$

Then $Y := \sum_{j=1}^{m} X_j$ is the cardinality of set cover after randomized rounding and $W := \sum_{i=1}^{n} Z_i$ is the number of vertices which are covered at least b times after this step. By Lemma 6.1, there exists atleast $b-1$ edges in $S_{\geqslant}$ that contains v for all $v \in V$.

Thus the number of additional edges taken into the cover in the repairing step is at most $n - W$. So the size of the cover can be estimated as below:

$$|C| \leqslant Y + n - W. \tag{6.1}$$

Let us denote by $t(LP)$ the time to solve the LP relaxation in step 2 of the Algorithm 2. We know that $t(LP)$ is polynomial as we can use any polynomial time LP-solver (e.g. Khachiyan [Kha79] or Karmakar [Kar84]).

6.3 The Derandomized Algorithm

We now state the derandomized version of the randomized set b-multicover algorithm (Theorem 4, [EMS14]).

Theorem 6.2. Let $b \in \mathbb{N}_{\geqslant 2}$ and let $\mathcal{H}$ be a hypergraph with maximum vertex degree Δ and maximum edge size l, $3 \leqslant l \leqslant \underbrace{2^{-1/4} \cdot 3^{-1}}_{\approx 0.28} \cdot (nb)^{\frac{1}{5}}$. Let Opt* be the value of an optimal solution of relaxed linear program. Then there exists a deterministic $\mathcal{O}(t(LP) + m^2 n)$ time algorithm that returns a set b-multicover of size

$$|C| \leqslant \left(1 - \frac{11(\Delta - b)}{72l}\right) \delta \cdot \mathrm{Opt}^*. \tag{6.2}$$

Proof. Given $3 \leqslant l \leqslant \underbrace{2^{-1/4} \cdot 3^{-1}}_{\approx 0.28} \cdot (nb)^{\frac{1}{5}}$ and $\beta = \frac{\sqrt{2}l}{\sqrt{nb}}$, let us now define the event E, which takes care of the feasibility of the solution as:

$$W > n(1 - \epsilon^2) - \alpha, \tag{6.3}$$

where $\alpha := \sqrt{\delta \cdot \sum_{k=1}^{m} |E_k|^2}$ and $\mathbb{E}[W] \geqslant n(1 - \epsilon^2)$. The event F, which relates to the quality of the solution is defined as:

$$Y < \lambda(1 + \beta)\mathrm{Opt}^*. \tag{6.4}$$

We will first analyze the event E^c. Consider the function $f(X_1, \ldots, X_m) := \sum_{i=1}^{n} Z_i$. Then we have for any two vectors $x = (x_1, \ldots, x_k, \ldots, x_m)$ and $x' = (x_1, \ldots, x'_k, \ldots, x_m)$ that only differ in the k-th coordinate

$$|f(x_1, \ldots, x_k, \ldots, x_m) - f(x_1, \ldots, x'_k, \ldots, x_m)| \leqslant |E_k|$$

for all $k \in [m]$.

In [EMS14] it is shown that the BDI (Theorem 3.5) for event E and the Angluin-Valiant inequality (Theorem 2.3) for event F give $\mathbb{P}(E^c \cup F^c) \leqslant e^{-2} + e^{-4}$, hence $\mathbb{P}(E \cap F) \geqslant 0.84$. It is shown in [EMS14] that a vector from $E \cap F$ satisfies the approximation in (6.2). We wish to apply Theorem 3.14 to the event $E \cap F$ leading to the derandomized construction of vector in $E \cap F$. For application of Theorem 3.14 we have to first fix P and Q. Since $\max_{\{0,1\}^n} \sum_{i=1}^{n} Z_i \leqslant n$, so we can set $P = n$. For computing $\mathbb{E}(\sum_{i=1}^{n} Z_i | X_1, \ldots, X_l)$ the linearity of expectation reduce the problem to the computation of $\mathbb{E}(Z_i | X_1, \ldots, X_l)$. Let

$S_l := S_{\geqslant} \cup \{E_1, E_2, \cdots, E_k\}$ where $E_1, E_2, \cdots, E_k$ denote the chosen hyperedges, once X_i, $1 \leqslant i \leqslant l$, are fixed. Let $\mathcal{E}(v_i)$ denote the set of hyperedges which contain vertex i. We have

$$\begin{aligned}
\mathbb{E}\Big(\sum_{i=1}^{n} Z_i \,|\, X_1, \ldots, X_l\Big) &= \sum_{i=1}^{n} \mathbb{E}(Z_i \,|\, X_1, \ldots, X_l) \\
&= \sum_{i=1}^{n} \Big(1 - \mathbb{P}(Z_i = 0 \,|\, X_1, \ldots, X_l)\Big) \\
&= \sum_{i=1}^{n} \Big(1 - \prod_{E_j \in (\mathcal{E}(v_i) \setminus S_l)} (1 - \lambda x_j^*)\Big)
\end{aligned}$$

Hence we need $\mathcal{O}(nm)$ pre-computation time (to construct $\mathcal{E}(v_i)$ sets for every $i \in [n]$) and at most $\mathcal{O}(n\delta) = \mathcal{O}(nm)$ (as $m > \delta$) steps to compute the conditional expectation, so we can set $Q := nm$. Also $\frac{\lambda}{D_{11}} = \frac{\alpha}{\sum_{k=1}^{m} |E_k|^2} = \frac{\sqrt{\delta \cdot \sum_{k=1}^{m} |E_k|^2}}{\sum_{k=1}^{m} |E_k|^2} \leqslant 1$.

By Theorem 3.14, the running time to construct $x \in E \cap F$ is $\mathcal{O}(m(n + \ln(m) + mn) + m^2 \ln(m)) = \mathcal{O}(m^2 n)$.

$\square$

Chapter 7

Derandomized Algorithm for the Multidimensional-Bin Packing Problem

Bin packing is another extensively studied problem which had rich implications on the field of approximation algorithms. In the *classical bin packing* problem, assume we have n items, $\{i_1, i_2, \cdots, i_n\}$, where the size of each item $i_k \in (0, 1]$ for all $k \in [n]$ and bins have capacity 1. A solution is optimal if it minimizes the number of bins which are required to pack all the items. The problem was proved to be $\mathcal{NP}$-Hard by Garey and Johnson in 1979 [GJ79]. Bin packing problem can be naturally extended to higher dimensions, namely *vector bin packing* problem and *geometric bin packing* problem.

In 2-dimensional geometric bin packing problem, we are given a collection of rectangular items to be packed into minimum unit-sized squares. This variant and other high dimensional variants have applications in cutting stock, vehicle loading, pellet packing and other logistics, robotics related problems. In the d-dimensional vector bin packing problem, each bin and item has d dimensions and we need to partition the items such that we can pack them in minimum number of bins. For example, we can think of each job as an item with CPU, RAM, disk, network requirements etc. as its d dimensions. The goal is to assign all the the jobs to minimum number of computing devices, which can be considered as d-dimensional bins with bounded amount of d resources required by the jobs.

7.1 Bansal's Algorithm

Bansal et al. [BCS09] gave the *Round and Approx* framework (R&A) and used it to construct algorithms for two-dimensional geometric bin packing problem and vector bin packing problem. In their paper, Bansal et al. derandomized the R&A framework using the potential function approach. We give an alternative approach using our derandomization framework for independent bounded differences inequality.

We first describe the randomized R&A framework of Bansal et al. The basic idea is to formulate the bin packing problem as a set covering problem. Please note bin packing problem considered below can either be vector bin packing problem or geometric bin packing problem. Let us define a *configuration* C as a subset of items I, which can be packed into a bin. Let us denote the set of configurations as $\mathcal{C} \subseteq 2^I$. Note that $\mathcal{C}$ can be exponentially large. We safely assume that the relaxation of the below mentioned LP problem can be solved with in $(1+\kappa)$ accuracy where $\kappa > 0$. The set covering formulation is to choose the minimum number of configurations which covers all the items :

$$\min\Big\{\sum_{C\in\mathcal{C}} x_C : \sum_{C\ni i} x_C \geqslant 1\,(i\in I), x_C \in \{0,1\}\,(C\in\mathcal{C})\Big\} \tag{7.1}$$

They construct an approximate solution of the set covering problem (7.1) by performing the following steps, where $\alpha > 0$ is a parameter to be specified later.

1. **Solving LP:** Solve the linear programming relaxation of the integer linear program (7.1), possibly approximately in case $\mathcal{C}$ is exponentially large in the input size. Let x^* be the near-optimal solution of the LP relaxation and $z^* = \sum_{C\in\mathcal{C}} x^*_C$.
2. **Randomized Rounding:** Define a binary vector $x^r \in \mathbb{R}^{|\mathcal{C}|}$ where initially $x^r_C = 0,$ for all $C \in \mathcal{C}$. We will iterate $\lceil \alpha z^* \rceil$ times where for each iteration, we select one configuration C' independently at random, letting each $C \in \mathcal{C}$ be selected with probability $\frac{x^*_C}{z^*}$. Put $x^r_{C'} := 1$.
3. **Solving the residual instance:** Consider the set of items $S \subseteq I$ that are not covered by x^r, namely $i \in S$ if and only if $\sum_{C\ni i} x^r_C = 0$, and the associated optimization problem for the residual instance of the problem is:

$$\min\Big\{\sum_{C\in\mathcal{C}} x_C : \sum_{C\ni i} x_C \geqslant 1\,(i\in S), x_C \in \{0,1\}\,(C\in\mathcal{C})\Big\} \tag{7.2}$$

   Apply an approximation algorithm to the residual instance (7.2) yielding solution x^a.
4. **Combining the solution:** Return the solution $x^h := x^r + x^a$.

This framework was analyzed using Theorem 3.5 and hence we can derandomize its step 2 using Theorem 3.7.

We will fix definitions before we state the result of Bansal et al. Given a deterministic approximation algorithm, we say that it has *asymptotic approximation guarantee* ρ if there exists a constant λ such that the value of the solution found by the algorithm is atmost $\rho\,\mathrm{OPT}(I)+\lambda$ for each instance I. The $\mathrm{OPT}(I)$ denotes the optimal value of the given problem instance I. If $\lambda = 0$, then the algorithm has *absolute approximation guarantee* of ρ. Given a randomized approximation algorithm, we say that it has *asymptotic approximation guarantee* ρ if there exists a constant λ such that the value of the solution found by the algorithm is atmost $\rho\,\mathrm{OPT}(I) + \lambda$ for each instance I with a probability tending to 1 as $\mathrm{OPT}(I)$ tends to infinity.

We also introduce the concept of *subset oblivious algorithms*. We let $\mathrm{OPT}(S)$ and $\mathcal{A}(S)$ respectively, the optimal value of solution of (7.2) and the value of the heuristic solution by the approximation algorithm.

Definition 7.1. An asymptotic ρ-approximation algorithm for the problem (7.1) is called *subset oblivious*, if for any fixed $\varepsilon > 0$, there exist constants k, β, δ (possibly depending upon ε) such that, for every instance I of (7.1), there exists vectors $v^1, \ldots, v^k \in \mathbb{R}^{|I|}$ with the following properties:

1. $\sum_{i \in C} v_i^j \leqslant \beta$, for each $C \in \mathcal{C}$ and $j = 1, 2, \ldots, k$.
2. $\mathrm{OPT}(I) \geqslant \max_{j=1}^{k} \sum_{i \in I} v_i^j$
3. $\mathcal{A}(S) \leqslant \rho \max_{j=1}^{k} \sum_{i \in S} v_i^j + \varepsilon\,\mathrm{OPT}(I) + \delta$, for each $S \subseteq I$.

This means that quality of approximation algorithm can be expressed in terms of the above stated set of vectors $v^1, \ldots, v^k$. Essentially Bansal et al. established that if have an ρ-approximation algorithm which satisfies subset oblivious property then R&A frameowork gives an approximation algorithm with asymptotic approximation guarantee of $\ln(\rho) + 1$.

The main result of Bansal et al. is:

Theorem 7.2. (Theorem 2, [BCS09]) If the method R&A uses a μ-approximation algorithm to find an optimal solution of the LP relaxation for $\mu < \rho$,i.e. $z^* \leqslant \mu\ \mathrm{OPT}(I)$, and there exists an asymptotic ρ-approximation subset oblivious algorithm for problem (7.1) then, for any constant $\gamma > 0$, the cost of the final heuristic solution produced by R&A using that algorithm in step 3 with $\alpha := \ln \rho - \ln \mu$ is at most

$$(\mu(\ln \rho - \ln \mu + 1) + \varepsilon)\,\mathrm{OPT}(I) + \delta + \gamma z^* + 1 \tag{7.3}$$

with probability at least $1 - ke^{-2(\gamma z^*)^2/(\beta^2 \lceil z^* \ln \rho \rceil)}$, i.e. method R&A is a randomized asymptotic $(\ln \rho + \varepsilon + 1)$-approximation algorithm for problem (7.1) in case when $\mu = 1$ or $\mu = 1 + \delta$ for δ arbitrarily close to 0.

Now, we wish to apply our derandomization framework. For this purpose, more details of the analysis, in particular the definition of the events we are interested in, are necessary.

Let $C_1, \ldots, C_m \in \mathcal{C}$ be the configurations associated with nonzero components of the relaxed LP solution x^*. Set $\lceil \alpha z^* \rceil := c$. Let us define c random variables $Z_1, \ldots, Z_c$ where for each $i \in [c]$, $Z_i \in \{1, 2, \ldots, m\}$. The random variable Z_i represents the configuration selected in i^{th} randomized rounding step. At the end of the c iterations of the randomized rounding step, we would possibly have a set S of items still uncovered by the selected configurations. We choose atmost $\lceil \alpha z^* \rceil$ edges in the randomized rounding step hence cost of x^r is $\lceil \alpha z^* \rceil \leqslant \alpha\mu\,\mathrm{OPT}(I) + 1 \leqslant \mu \ln(\rho/\mu)\,\mathrm{OPT}(I) + 1$.

Now we calculate the cost of x^a. Consider the random variable $\sum_{i \in S} v_i^j$ for $j = 1, \ldots, k$. This is the same quantity which occurs on the right hand side of the inequality for $\mathcal{A}(S)$ given in the Definition 7.1 for subset-oblivious algorithms. Let us first approximate the expectation of the random variable $\sum_{i \in S} v_i^j$:

$$\mathbb{E}\Big(\sum_{i \in S} v_i^j\Big) = \sum_{i \in I} v_i^j \mathbb{P}(i \in S) = \sum_{i \in I} v_i^j \Big(1 - \sum_{C \ni i} x_C^*/z^*\Big)^{\lceil \alpha z^* \rceil} \leqslant e^{-\alpha} \sum_{i \in I} v_i^j \tag{7.4}$$

where the above inequality holds because $\sum_{C \ni i} x_C^* \geqslant 1$ for $i \in I$ and $(1 - 1/a)^{\lceil \alpha a \rceil} \leqslant (1 - 1/a)^{\alpha a} \leqslant e^{-\alpha}$ for $a > 0$.

By the structure of the algorithm, the random variable $\sum_{i \in S} v_i^j$ is a function of $Z_1, Z_2, \ldots, Z_c$. Let us define functions f_j by

$$f_j(Z_1, Z_2, \ldots, Z_c) := \sum_{i \in S} v_i^j \tag{7.5}$$

for all $j = 1, 2, \ldots, k$.

In our context, changing value of Z_i means that we choose a different configuration C' in place of configuration C. Let S' be the resulting set of uncovered items. Then

$$\Big|\sum_{i \in S} v_i^j - \sum_{i \in S'} v_i^j\Big| \leqslant \max\Big(\sum_{C \backslash C'} v_i^j, \sum_{C' \backslash C} v_i^j\Big) \leqslant \beta \tag{7.6}$$

The last inequality follows from the first property of subset oblivious algorithms.

Let us define a event E_j for $j = 1, 2, \cdots, k$ by:

$$\sum_{i \in S} v_i^j - \mathbb{E}(\sum_{i \in S} v_i^j) \leqslant \gamma z^* \tag{7.7}$$

Applying Theorem 3.5, we get the following upper bound on probability for E_j^c

$$\mathbb{P}\Big[\sum_{i \in S} v_i^j - \mathbb{E}(\sum_{i \in S} v_i^j) \geqslant \gamma z^*\Big] \leqslant e^{-2(\gamma z^*)^2/(\beta^2 \lceil \alpha z^* \rceil)} \tag{7.8}$$

Based on the above bound, we substitute $\max_{j=1}^k \sum_{i \in S} v_i^j$ by $\mathbb{E}\left[\sum_{i \in S} v_i^j\right] + \gamma z^*$ in the third condition of subset obliviousness to calculate the cost of x^a. By union bound, we get the probability of the event $\bigcup_j E_j^c$ and hence we establish the statement of Theorem 7.2.

7.2 The Dearandomized Algorithm

We now prove the derandomized counterpart of Theorem 7.2.

Theorem 7.3. If the method R&A uses a μ-approximation algorithm to find an optimal solution of the LP relaxation for $\mu < \rho$,i.e. $z^* \leqslant \mu \ \mathrm{OPT}(I)$, and there exists an asymptotic ρ-approximation subset oblivious algorithm for problem (7.1), then, for any constant $\gamma > 0$, we can construct a solution deterministically using the algorithm in step 3 with $\alpha := \ln\rho - \ln\mu$ with the final cost at most

$$(\mu(\ln\rho - \ln\mu + 1) + \varepsilon)\,\mathrm{OPT}(I) + \delta + \gamma z^* + 1, \tag{7.9}$$

i.e. method R&A is an asymptotically $(\ln\rho + \varepsilon + 1)$-approximation algorithm for problem (7.1) in case when $\mu = 1$ or $\mu = 1 + \delta$ for δ arbitarily close to 0.

Proof. We invoke the derandomized BDI (Theorem 3.7). Let E be the event $E := \bigcap_{j=1}^{k} E_j$. It is proved in [BCS09] that the probability of E is lower bounded by some α for $\alpha > 0$. For the application of Theorem 3.7, we have to first fix P and Q. We can choose $P = |I|$ as $\sum_{i\in S} v_i^j \leqslant \sum_{i\in I} v_i^j \leqslant \mathrm{OPT}(I) \leqslant |I|$. Estimating the value of Q is not straightforward. Q is the time to compute $\mathbb{E}(f_j \mid Z_1, Z_2, \cdots, Z_l)$. We choose l not necessarily distinct configurations. Let U be the union of the l selected configurations. Hence

$$\mathbb{E}\Big(\sum_{i\in S} v_i^j \mid Z_1, Z_2, \cdots, Z_l\Big) = \sum_{i\in I\setminus U} v_i^j \mathbb{P}(i \in S) = \sum_{i\in I\setminus U} v_i^j \Big(1 - \sum_{C\ni i} x_C^*/z^*\Big)^{c-l} \tag{7.10}$$

The above expression can be computed in $\mathcal{O}(|I|(m+c))$ time. Hence $Q = \mathcal{O}(|I|(m+c))$. With $\max_i \frac{\lambda_i}{D_{i1}} = \frac{\gamma z^*}{c\beta^2}$, the running time of the derandomization step according to Theorem 3.7 is $\mathcal{O}(mkc[|I|\frac{\gamma z^*}{c\beta^2} + \log\frac{kc}{\beta} + |I|(m+c)])$, and we have derandomized the step 2 of R&A framework i.e. randomized rounding step.

$\square$

Note that Bansal et al. proved that for the 1-dimensional bin packing problem the aysmptotic polynomial time approximation algorithm of Vega and Lueker [Fer81] with minor adaptation is subset oblivious. Based on this modified algorithm, Bansal et al. designed a subset oblivious algorithm for the d-dimensional bin packing problem with asymptotic approximation guarantee arbitarily close to d. They improved the known asymptotic approximation guarantee of d-dimensional bin packing from $\mathcal{O}(\ln(d))$ to $\ln(d) + 1$. Details of the impact of R&A method can be found in the survey article by Christensen et al. [CKP+17]. We conclude this section with the remark that our derandomization approach is applicable to all results using R&A method as the framework.

Chapter 8

Derandomized Approximation of Constrained Hypergraph Coloring (CHC)

The *constrained hypergraph coloring problem*(CHC) problem was introduced by Ahuja and Srivastav [AS02] as a multicolor generalization of the property B hypergraph coloring problem. It also models special cases of multidimensional bin packing (MDBP) problem and the resource constrained scheduling (RCS) problem.

Definition 8.1. (Constrained hypergraph coloring problem(CHC), [AS02]) Let $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \in \{0,1\}^{m \times n}$ be the edge-vertex incidence matrix for hypergraph $\mathcal{H} = (V, \mathcal{E})$ with $V = \{1, 2, \cdots, n\}, \mathcal{E} = \{E_1, \ldots, E_m\}$, and $l = \max\limits_{1 \leqslant i \leqslant m} |E_i|$. Let $b = (b_1, b_2, \ldots, b_m)^t$ be a vector. The problem is to partition the vertex set into a minimum number of sets such that there are at most b_i vertices in E_i of any partition set for all $i \in [m]$. We may color the vertices of each partition set with one color and call the set *color class*.

CHC reduces to extensively studied well known problems based on the underlying combinatorial structure and value of b_i. For a simple graph and $b_i = 1$ for all $i \in [m]$, CHC reduces to the graph coloring problem. The hypergraph $\mathcal{H}$ is said to be c-colorable iff there is a function $V \to \{1, 2, \cdots, c\}$ such that no edge is monochromatic. Hypergraph 2-colorability is the famous property B [Erd63]. For a hypergraph $\mathcal{H}$ with property B and $b_i = |E_i| - 1$ for all $i \in [m]$, CHC is equivalent to the problem of finding a non-monochromatic 2-coloring of $\mathcal{H}$.

In resource constrained scheduling problem(RCS) we are given m resources and n jobs, where for the i^{th} resource b_i units are available at each time slot, $i \in [m]$. Each job requires at least one resource and the aim is to partition the set of jobs such that resource constraints are not violated for each partition and the time taken to process the jobs is minimum. We now define the hypergraph $\mathcal{H}$ in this scenario. We consider V as the index set of jobs and E_i for all $i \in [m]$, is the hyperedge with length n where the j^{th} element of E_i is 1 if job j requires resource i and 0 otherwise. In the special case of the multidimensional bin backing problem(MDBP), we are given n items which are m-dimensional integral vectors $y_1, y_2, \cdots, y_n \in \{0,1\}^m$ and an unlimited number of m-dimensional bins. We assume that the i^{th} coordinate of each bin is at most b_i for all $i \in [m]$. We aim to assign the items to a minimum of bins such that in each bin the sum of the i-th coordinate of items in this bin is at most b_i, $i \in [m]$. Again, this is a partitioning problem and "packing" simply means to partition the set of items.

In 1997, Srivastav, Stangier [SS97] gave an randomized polynomial time algorithm for the RCS problem with a makespan of at most $\lceil (1+\epsilon)\, C_{opt} \rceil$ where $\epsilon > 0$ is arbitrary provided that $b_i \geqslant \frac{3(1+\epsilon)}{\epsilon^2} \log(8Cm)$ and $m \geqslant \frac{3(1+\epsilon)}{\epsilon^2} \log(8C)$. Let C_{opt} and C resp. denote the optimal integer solution and optimal fractional solution respectively of the linear programming relaxation of the RCS problem. We present a deterministic polynomial time algorithm with a makespan of at most $\lceil (1+\epsilon)\, C_{opt} \rceil$ provided that $b_i \geqslant \frac{1+\epsilon}{\epsilon}\sqrt{\frac{l \ln(c_1 mn)}{2}}$ for all $i \in [m]$, where $c_1 > 1$ is a constant. Our bound on b_i is better if $l \leqslant \frac{18 \log^2(8mn)}{\epsilon^2}$ ($c_1 \in [1, 8]$). We claim similar improvements to MDBP results presented in [SS97].

Ahuja and Srivastav [AS02] gave a deterministic algorithm for CHC with at most $\lceil (1+\epsilon)\, C_{opt} \rceil$ color classes for any $\epsilon \in (0, 1)$, provided that $b_i = \Omega(\epsilon^{-2}(1+\epsilon)\log(D))$, where D can be understood as the dependency of the graph used in the Lovász-Local-Lemma invoked in the analysis of their algorithm. From [AS02], it is known that $D = \mathcal{O}((Cm)^2)$. Our bound on b_i is better, if $l \leqslant \frac{c_3 \log^2(gCm)}{\epsilon^2 \log c_1 mn} \leqslant \frac{c_3 \log(vmn)}{\epsilon^2}$, given constants $c_3, g > 0$, and $v = \max(g, c_1)$.

8.1 CHC as an Integer Linear Program

Since the randomized algorithm for the CHC problem uses the randomized rounding scheme, let us state an integer linear programming formulation of CHC (CHC − ILP). Consider vectors $x^{(k)} = (x_{1,k}, x_{2,k}, \ldots, x_{n,k})^t \in \mathbb{R}^n$, $k \in \{1, \cdots, T\}$ for some $T \in \mathbb{N}$, $T \leqslant n$. The integer linear program is the following:

$$
\begin{aligned}
(CHC - \text{ILP}) \qquad & \min T \\
& Ax^{(k)} \leqslant b \quad \forall k \in \{1, 2, \cdots, T\} && (8.1)\\
& \sum_{k=1}^{T} x_{j,k} = 1 \quad \forall j \in \{1, 2, \cdots, n\} && (8.2)\\
& x_{j,k} = 0 \quad \forall j \in \{1, 2, \cdots, n\} \text{ and } k > T && (8.3)\\
& x_{j,k} \in \{0, 1\} \quad \forall j, k. && (8.4)
\end{aligned}
$$

Let C_{opt} be the minimum value for T for CHC − ILP.

Proposition 8.2. CHC − ILP is equivalent to CHC.

Proof. We start with a solution of CHC − ILP with T color classes. For $i \in [m]$ and $k \in [T]$ condition (8.1) is $(Ax^{(k)})_i \leqslant b_i$, and says that each color class has at most b_i elements in each E_i for all $i \in [m]$.

(8.2) ensures that each vertex of the hypergraph is in one of the color classes. Hence all the conditions of the CHC problem are satisfied by the $CHC-\mathrm{ILP}$. On the other hand, given a solution of the CHC problem with T color classes, we get a solution of $CHC-\mathrm{ILP}$ as follows. We define the variables $x_{j,k}$ as 1 if vertex j is in color class k, and 0 otherwise, for all $j \in [n], k \in [T]$. Then $Ax^{(k)} \leqslant b$ for all $k \in [T]$. Since the other conditions of $CHC-\mathrm{ILP}$ are satisfied by definition, we established the equivalence of CHC and $CHC-\mathrm{ILP}$. □

Let us relax the integrality condition $x_{j,k} \in \{0,1\}$ to $x_{j,k} \in [0,1]$ for all j,k. We call this program $CHC-\mathrm{LP}$. We can find an optimal T for $CHC-\mathrm{LP}$ by binary search solving at most $\log(T)$ linear programs starting with $T = n$. Hence C can be computed in polynomial time, if we use standard polynomial time LP algorithms. Let an optimal solution for $CHC-\mathrm{LP}$ corresponding to C colors be $\tilde{x}_{jk}$, $0 \leqslant \tilde{x}_{jk} \leqslant 1$. A possible randomized rounding procedure is to represent each vertex j independently by a C-faced dice with face probabilities $\tilde{x}_{jk}$, where k represents the choice of the color for the vertex j for all $k = 1, 2, \ldots, C$ and $j = 1, 2, \ldots, n$. Unfortunately, there is a high probability of constraints getting violated, when we follow this rounding procedure. We therefore modify the first constraint and tighten the right hand side of the first constraint as follows. Let $\epsilon > 0$ and consider the tighter LP:

$$
\begin{aligned}
(CHC-\mathrm{LPT}) \qquad & \min T \\
& Ax^{(k)} \leqslant \frac{b}{1+\epsilon} \quad \forall k \in \{1, 2, \cdots, T\} \\
& \sum_{k=1}^{T} x_{j,k} = 1 \quad \forall j \in \{1, 2, \cdots, n\} \\
& x_{j,k} = 0 \quad \forall j \in \{1, 2, \cdots, n\} \text{ and } k > T \\
& x_{j,k} \in [0,1] \quad \forall j, k.
\end{aligned}
$$

We construct a new fractional solution with value $\lceil (1+\epsilon) C \rceil$.

Let $\delta = \frac{1}{1+\epsilon}$, $\alpha = \epsilon\delta / \lceil \epsilon C \rceil$ and set

$I := \{1, 2, \ldots, \lceil (1+\epsilon) C \rceil\}$,

$I_1 := \{1, \ldots, C\}$ and $I_1^\epsilon = \{C+1, \ldots, C + \lceil \epsilon C \rceil\}$. So $I = I_1 \cup I_1^\epsilon$.

The new fractional assignments $x^*_{j,k}$ are:

$$
x^*_{j,k} := \begin{cases} \delta \tilde{x}_{jk} & \text{for } k \in I_1 \\ \alpha & \text{for } k \in I_1^\epsilon \end{cases} \tag{8.5}
$$

Proposition 8.3. The $(x^*_{j,k})_{\substack{1 \leqslant j \leqslant n, \\ 1 \leqslant k \leqslant T}}$ satisfy the constraints of $(CHC-\mathrm{LPT})$ with $\lceil (1+\epsilon) C \rceil$

colors.

Proof. We have
$$\sum_{k\in I} x^*_{j,k} = \sum_{k\in I_1} x^*_{j,k} + \sum_{k\in I_1^\epsilon} x^*_{j,k} = \delta + \alpha\lceil\epsilon C\rceil = 1$$

Consider an arbitrary, but fixed $i \in [m], k \in [T]$. We have for $k \in I$
$$(Ax^{*(k)})_i = \sum_{j=1}^{n} a_{ij}x^*_{j,k} = \delta\sum_{j=1}^{n} a_{ij}\tilde{x}_{jk} \leqslant \delta b_i$$

For $k \in I_1^\epsilon$, we have
$$\begin{aligned}
(Ax^{*(k)})_i = \sum_{j=1}^{n} a_{ij}x^*_{j,k} = \sum_{j=1}^{n} \alpha a_{ij} &= \sum_{j=1}^{n} \alpha a_{ij} \sum_{k\in I_1} \tilde{x}_{jk} \\
&= \sum_{k\in I_1} \alpha \sum_{j=1}^{n} a_{ij}\tilde{x}_{jk} \\
&\leqslant \sum_{k\in I_1} \alpha b_i \\
&\leqslant \alpha C b_i \\
&\leqslant \delta b_i
\end{aligned}$$

By definition of our solution, we used $\lceil (1+\epsilon)\, C\,\rceil$ colors. □

8.2 The Randomized Algorithm and its Analysis

We give the detailed description of the algorithm.

Algorithm 3: Constrained Hypergraph Coloring Problem

Input : A hypergraph $\mathcal{H} = (V, \mathcal{E})$ with maximum edge cardinality l, $n := |V|, m := |\mathcal{E}|$.

Output : Coloring of the vertices of $\mathcal{H}$ with $\lceil (1+\epsilon)\, C\,\rceil$ colors.

1. **LP-Relaxation:**
   Solve the relaxed version of CHC - ILP: Starting with $T = n$ consider at most $\ln(n)$ times to find smallest value of T for which the LP is feasible. Let $(\tilde{x}_{jk})_{j\in[n]\,\&\,k\in\{1,2,\ldots,C\}}$ be the optimal solutions associated to C, the smallest value of T.
2. **Extending the Solution:** We extend the solution by modifying the solution $(\tilde{x}_{jk})$ to get a solution $(x^*_{j,k})$ with $\lceil (1+\epsilon)\, C\,\rceil$ color classes as in (8.5).
3. **Rounding the Solution:** We carry out randomized rounding for each j, with $(x^*_{j,k})$ being the probability for rounding the k-th color to 1 for $k \in \lceil (1+\epsilon)\, C\,\rceil$.
4. Return $\lceil (1+\epsilon)\, C\,\rceil$ and all corresponding color class vectors.

We proceed to the analysis of the randomized algorithm.

Theorem 8.4. Let $\epsilon > 0$. Algorithm 3 outputs in randomized polynomial time a solution for the CHC problem with $\lceil (1+\epsilon) C \rceil$ colors with probability at least $1 - \frac{1}{c_1}$, $c_1 > 0$ a constant, provided that $b_i \geqslant \frac{1+\epsilon}{\epsilon}\sqrt{\frac{l \ln(c_1 mn)}{2}}$ for all $i \in [m]$.

Proof. Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables taking values in I, where for each $k \in I$, the distribution of X_j is defined by

$$\mathbb{P}(X_j = k) = x^*_{jk}, j \in [n] \tag{8.6}$$

For $k \in I$ and $j \in [n]$, let $X_{j,k}$ be the 0-1 random variable, which is 1, if $X_j = k$ and zero otherwise. For $i \in [m]$ and $k \in I$, we define events E_{ik} such that for each color class vector and each edge, the following constraint is not violated:

$$\sum_{j=1}^{n} a_{ij} X_{j,k} \leqslant b_i \text{ for all } i, k \tag{8.7}$$

Let us define $f_{ik}(X_{1,k}, X_{2,k}, \ldots, X_{n,k}) = \sum_{j=1}^{n} a_{ij} X_{j,k}$ for all i, k. We first bound the expectation of f_{ik} from above.

$$\mathbb{E}(f_{ik}) = \mathbb{E}\Big(\sum_{j=1}^{n} a_{ij} X_{j,k}\Big) = \sum_{j=1}^{n} a_{ij} \mathbb{E}(X_{j,k}) = \sum_{j=1}^{n} a_{ij} x^*_{j,k} \leqslant \frac{1}{1+\epsilon} b_i \text{ for all } i, k \tag{8.8}$$

Note that for all i, j, k, $|f_{ik}(X) - f_{ik}(X')| = a_{ij}$, where $X = (X_{1,k}, \cdots, X_{j-1,k}, X_{j,k}, X_{j+1,k}, \cdots, X_{n,k})$ and $X' = (X_{1,k}, \cdots, X_{j-1,k}, X'_{j,k}, X_{j+1,k}, \cdots, X_{n,k})$ differ only on the j-th coordinate. Since $a_{ij} \in \{0, 1\}$ for all i, j,

$$\sum_{j=1}^{n} a_{ij}^2 = \sum_{j=1}^{n} a_{ij} = |E_i| \tag{8.9}$$

Let $i \in [m]$ and $k \in I$. We now apply the independent bounded differences inequality (Theorem 2.11) for $t = \frac{\epsilon}{1+\epsilon} b_i$:

$$\begin{aligned} \mathbb{P}(E^c_{ik}) = \mathbb{P}\Big(f_{ik} \geqslant b_i\Big) &\leqslant \mathbb{P}\Big(f_{ik} - \mathbb{E}(f_{ik}) \geqslant \frac{\epsilon b_i}{1+\epsilon}\Big) \quad \text{(by 8.8)} \\ &\leqslant \exp\Big(-\frac{2\epsilon^2 b_i^2}{(1+\epsilon)^2 |E_i|}\Big) \quad \text{(Theorem 2.11)} \\ &\leqslant \exp\Big(-\frac{2\epsilon^2 b_i^2}{(1+\epsilon)^2 l}\Big) \quad (\text{as } |E_i| \leqslant l) \end{aligned} \tag{8.10}$$

Using $b_i \geqslant \frac{1+\epsilon}{\epsilon}\sqrt{\frac{l\ln(c_1 mn)}{2}}$ and the fact that $|I| \leqslant n$, (8.10) and the union bound give

$$\begin{aligned}
\mathbb{P}\Big(\bigcup_{i=1,k=1}^{i=m,k=|I|} E_{ik}^c\Big) &\leqslant \sum_{i=1,k=1}^{i=m,k=|I|} \mathbb{P}(E_{ik}^c) \\
&\leqslant \sum_{i=1,k=1}^{i=m,k=n} \mathbb{P}(E_{ik}^c) \\
&\leqslant \sum_{i=1}^{m}\sum_{k=1}^{n} \exp\Big(-\frac{2\epsilon^2 b_i^2}{(1+\epsilon)^2 l}\Big) \text{ (by 8.10)} \\
&\leqslant \frac{1}{c_1} \text{ (lower bound on } b_i \text{ used)}
\end{aligned} \tag{8.11}$$

The randomized polynomial running time stems from the fact that there are polynomial time LP solvers [Kha79; Kar84].

□

8.3 The Derandomized CHC Algorithm

In this section, we derandomize the randomized algorithm given in the above section. Let us set $c := \lceil (1+\epsilon)\, C \,\rceil$.

For the values of b_i in Theorem 8.4, we have by (8.11)

$$\sum_{i=1}^{m}\sum_{k=1}^{c} e^{\left(-\frac{2\epsilon^2 b_i^2}{(1+\epsilon)^2 l}\right)} \leqslant \frac{1}{c_1} \tag{8.12}$$

with constant $c_1 > 0$ as there.

This means that $\mathbb{P}(\bigcap_{i=1,k=1}^{i=m,k=c} E_{ik}) \geqslant 1 - \frac{1}{c_1}$. Hence, we can apply Theorem 3.7 to the event $\bigcap_{i=1,k=1}^{i=m,k=c} E_{ik}$. For application of Theorem 3.7, we have to fix P and Q. As

$$\max_{i,k} f_{ik} = \max_{i,k} \sum_{j=1}^{n} a_{ij} X_{j,k} \leqslant n, \tag{8.13}$$

we can set $P = n$. Since

$$\mathbb{E}\left[f_{ik} \mid X_{1,k}, \cdots, X_{1,l}\right] = \mathbb{E}\left[\sum_{j=1}^{n} a_{ij} X_{j,k} \mid X_{1,k}, \cdots, X_{1,l}\right] = \sum_{j=1}^{n} \mathbb{E}\left[a_{ij} X_{j,k} \mid X_{1,k}, \cdots, X_{1,l}\right] \tag{8.14}$$

we can set the computation time Q for the conditional expectations as $Q = \mathcal{O}(n)$.

Theorem 8.5. Let $\epsilon \in (0,1)$ and $b := (b_1, b_2, \cdots, b_m)^t$ where $b_i \geqslant \frac{1+\epsilon}{\epsilon}\sqrt{\frac{l \ln(c_1 mn)}{2}}$ for all $i \in [m]$.Then a solution of the CHC problem with respect to b with atmost $c := \lceil (1+\epsilon) C \rceil$ colors can be computed in $\mathcal{O}(cmn[n + \log \frac{cmn}{1-\frac{1}{c_1}}])$ time.

Proof. Observe that we have cn random varibles $X_{j,k}$ for $j \in [m]$, $k \in [c]$. We also have $\max_i \frac{\lambda_i}{D_{i1}} = \max_i \frac{\epsilon b_i}{(1+\epsilon)|E_i|} = \frac{\epsilon}{1+\epsilon} \leqslant 1$ as $b_i \leqslant |E_i|$. Now by Theorem 3.7, we get the desired solution of the CHC problem in $\mathcal{O}(cmn[n + \log \frac{cmn}{1-\frac{1}{c_1}}])$ time. □

Chapter 9

Concentration Bounds with Partially Dependent Random Variables

Svante Janson [Jan04] extended Hoeffding's bound (Theorem 2.1) for sums of independent random variables to obtain concentration bounds for sums of dependent random variables with a defined dependency structure. The method is based on breaking the sum of non-independent random variables into sums of independent random variables. Janson applied the framework to U-Statistics, random strings and random graphs. In this chapter, we extend the Alon-Spencer [AS04] concentration bound for sums of independent random variables, which generalizes Hoeffding's bound, to obtain concentration bounds for the sum of dependent random variables with similar dependency structure as defined in [Jan04]. Let us first define the dependency structure as given in [Jan04].

In Hoeffding's bound, we have the following situation: Y_i, with $i \in \mathcal{I}$ where $\mathcal{I}$ is some index set, are mutually independent random variables, and one considers large deviation for the sum

$$Z := \sum_{i \in \mathcal{I}} Y_i \tag{9.1}$$

We will now investigate the case where the random variables Y_i are not mutually independent, but there is some amount of independence for partial sums. A typical example would be

$$Z = \sum_{(i_1,i_2) \in \mathcal{I}} W_{i_1} W_{i_2} \tag{9.2}$$

where $W_1, W_2, \cdots, W_n$ be n independent random variables and $\mathcal{I}$ is the set of all pairs (i_1, i_2) with $1 \leqslant i_1 < i_2 \leqslant n$. We can easily see that each summand is dependent on $n-2$ summands and independent of $\frac{(n-2)(n-3)}{2}$ summands as for fixed i_1, i_2, we have $\frac{(n-2)(n-3)}{2}$ random variables which are independent of $W_{i_1} W_{i_2}$.

We now define the dependency structure, which forms the basis of our work. We are given an finite index set $\mathcal{I}$ and random variables Y_i, $i \in \mathcal{I}$

- ▻ A subset $\mathcal{I}'$ of $\mathcal{I}$ is *independent* if the corresponding random variables $\{Y_i \mid i \in \mathcal{I}'\}$ are independent.

- ▻ A family $\{\mathcal{I}_j\}_j$ of subsets of $\mathcal{I}$ is a *cover* of $\mathcal{I}$ if $\bigcup_j \mathcal{I}_j = \mathcal{I}$.
- ▻ A family $\{(\mathcal{I}_j, w_j)\}_j$ of pairs $(\mathcal{I}_j, w_j)$, where $\mathcal{I}_j \subseteq \mathcal{I}$ and $w_j \in [0,1]$, is called a *fractional cover* of $\mathcal{I}$, if $\sum_j w_j \mathbb{1}_{\mathcal{I}_j} \geqslant \mathbb{1}_{\mathcal{I}}$, i.e. $\sum_{j\,:\,i\in\mathcal{I}_j} w_j \geqslant 1$ for each $i \in \mathcal{I}$. A fractional cover is said to be *exact* if $\sum_j w_j \mathbb{1}_{\mathcal{I}_j} = \mathbb{1}_{\mathcal{I}}$.
- ▻ A (fractional) cover is called *proper* if each set $\mathcal{I}_j$ in it is independent.
- ▻ $\chi(\mathcal{I})$ is the size of a smallest cover of $\mathcal{I}$, i.e. the smallest m such that $\mathcal{I}$ is union of m independent subsets. If $\mathcal{I}$ itself is independent, then $\chi(\mathcal{I}) = m = 1$.
- ▻ $\chi^*(\mathcal{I})$ is the minimum of $\sum_j w_j$ over all proper fractional covers $\{(I_j, w_j)\}_j$.

Let us first define the *dependency graph* for Y_i, $i \in \mathcal{I}$. The dependency graph G has vertex set $\mathcal{I}$ and edge set such that if vertices $i, j \in \mathcal{I}$ does not have an edge implies Y_i and Y_j are independent. So, if $\mathcal{I}' \subset \mathcal{I}$ is an independent set of G, then random variables Y_i, $i \in \mathcal{I}'$ are independent. Let $\chi(G)$ and $\chi^*(G)$ be the chromatic number and fractional chromatic number resp. of graph G. Hence, $\chi(\mathcal{I}) \leqslant \chi(G)$ and $\chi^*(\mathcal{I}) \leqslant \chi^*(G)$. Given $\Delta(G)$ as the maximum degree of graph G, and since the computation of $\chi(\mathcal{I})$ and $\chi(G)$ is $\mathcal{NP}$-hard, we may use $\chi(G) \leqslant \Delta(G) + 1$ [Bro41] in the bounds, and work with $\Delta(G)$ without serious loss.

We now state lemma from [Jan04], which we will use as well.

Lemma 9.1. (Janson [Jan04] Lemma 3.1) If $\{(\mathcal{I}_j, w_j)\}_j$ is an exact fractional cover of $\mathcal{I}$, and c_i, $i \in \mathcal{I}$, be any numbers, then

$$\sum_{i\in\mathcal{I}} c_i = \sum_j w_j r_j \tag{9.3}$$

where $r_j := \sum_{i\in\mathcal{I}_j} c_i$. In particular, $|\mathcal{I}| = \sum_j w_j |\mathcal{I}_j|$.

Let us assume $|\mathcal{I}| = n$ and $Y_i = 1 - q_i$ with probability q_i and $Y_i = -q_i$ with probability $1 - q_i$, where $q_i \in [0,1]$ for $i \in [n]$. Let $q := \frac{\sum_i q_i}{n}$.

In the above setting, Alon and Spencer improved upon Hoeffding's bound (Theorem 2.1) by replacing n by qn. We will first state the Alon-Spencer bound [AS04] (Theorem A.1.11 and A.1.13).

Theorem 9.2. (Alon, Spencer [AS04]) Let $\mathcal{I}$ be a set with $|\mathcal{I}| = n$, and let the independent random variables Y_i, $i \in \mathcal{I}$ satisfy $\mathbb{P}(Y_i = 1 - q_i) = q_i$ and $\mathbb{P}(Y_i = -q_i) = 1 - q_i$, where $q_i \in [0,1]$. Let $q := \frac{\sum_i q_i}{n}$ and $Z := \sum_{i\in\mathcal{I}} Y_i$. Then for any $a > 0$,

1. $\mathbb{P}(Z \geqslant a) \leqslant e^{-\frac{a^2}{2nq} + \frac{a^3}{2(nq)^2}}$.
2. $\mathbb{P}(Z < -a) \leqslant e^{-\frac{a^2}{2nq}}$.

We now proceed to our main result of this chapter, which generalizes the Alon-Spencer bound to sums of partially dependent random variables, where $\chi^*(\mathcal{I})$ resp. $\chi(\mathcal{I})$ enters the Alon-Spencer bounds in a very natural way.

Theorem 9.3. Let $\mathcal{I}$ be a set with $|\mathcal{I}| = n$, and let the random variables Y_i, $i \in \mathcal{I}$ satisfying $\mathbb{P}(Y_i = 1-q_i) = q_i$ and $\mathbb{P}(Y_i = -q_i) = 1-q_i$, where $q_i \in [0,1]$. Let $q := \frac{\sum_i q_i}{n}$ and $Z := \sum_{i\in\mathcal{I}} Y_i$. Then for any $a > 0$,

1. $\mathbb{P}(Z \geqslant a) \leqslant e^{-\frac{a^2}{2nq\chi^*(\mathcal{I})^2}+\frac{a^3}{2(nq)^2\chi^*(\mathcal{I})^3}}$.

2. $\mathbb{P}(Z < -a) \leqslant e^{-\frac{a^2}{2nq\chi^*(\mathcal{I})}}$.

Proof. Let $\{(\mathcal{I}_j, w_j)\}_j$ be an exact proper fractional cover of $\mathcal{I}$ and $Z_j := \sum_{i\in\mathcal{I}_j} Y_i$. By Lemma 9.1, $Z = \sum_j w_j Z_j$. Note that by definition, each Z_j is a sum of independent random variables Y_i, $i \in \mathcal{I}_j$. Let $p_j > 0$ such that $\sum_j p_j = 1$. We will choose the p_j's later in an appropriate way. By Jensen's inequality we have

$$\begin{aligned}
\exp(uZ) &= \exp\Big(\sum_j uw_jZ_j\Big)\\
&\leqslant \exp\Big(\sum_j p_j\frac{uw_jZ_j}{p_j}\Big)\\
&\leqslant \sum_j p_j \exp\Big(\frac{uw_jZ_j}{p_j}\Big) \text{ (Jensen's inequality)}
\end{aligned}$$

Note we have broken down the sum into j summands, where each of them is expressed in terms of independent random variables. Now computing the expection of $\exp(uZ)$, we get

$$\begin{aligned}
\mathbb{E}(e^{uZ}) &\leqslant \sum_j p_j \mathbb{E}\Big(e^{\frac{uw_jZ_j}{p_j}}\Big)\\
&= \sum_j p_j \prod_{i\in\mathcal{I}_j} \mathbb{E}\Big(e^{\frac{uw_jY_i}{p_j}}\Big) \text{ (Z_j is sum of independent random variables)}\\
&= \sum_j p_j \prod_{i\in\mathcal{I}_j} \Big(e^{\frac{uw_j(1-q_i)}{p_j}}q_i + e^{-\frac{uw_jq_i}{p_j}}(1-q_i)\Big)\\
&= \sum_j p_j \prod_{i\in\mathcal{I}_j} e^{-\frac{uw_jq_i}{p_j}}\Big(e^{\frac{uw_j}{p_j}}q_i + 1 - q_i\Big)\\
&= \sum_j p_j e^{-\frac{uw_jQ_j|\mathcal{I}_j|}{p_j}} \prod_{i\in\mathcal{I}_j} \Big(e^{\frac{uw_j}{p_j}}q_i + 1 - q_i\Big) \text{ (with } Q_j = \frac{\sum_{i\in\mathcal{I}_j} q_i}{|\mathcal{I}_j|}\text{)}
\end{aligned}$$

By geometric mean/arithmetic mean inequality, we get

$$\mathbb{E}(e^{uZ}) \leqslant \sum_j p_j e^{-\frac{uw_jQ_j|\mathcal{I}_j|}{p_j}} \Big(e^{\frac{uw_j}{p_j}} Q_j + 1 - Q_j\Big)^{|\mathcal{I}_j|} \tag{9.4}$$

Applying Markov's inequality, we get

$$\mathbb{P}(Z \geqslant a) = \mathbb{P}(e^{uZ} \geqslant e^{ua}) \leqslant \frac{\mathbb{E}(e^{uZ})}{e^{ua}} \tag{9.5}$$

Now by (9.4) and (9.5), we get

$$\mathbb{P}(Z \geqslant a) \leqslant \sum_j p_j e^{-\frac{uw_jQ_j|\mathcal{I}_j|}{p_j}} \Big(e^{\frac{uw_j}{p_j}} Q_j + 1 - Q_j\Big)^{|\mathcal{I}_j|} e^{-ua} \tag{9.6}$$

We minimize the RHS of (9.6) with respect to u. By definition, we know $\chi^*(\mathcal{I}) = \sum_j w_j$. Further, we may choose p_j, namely set $p_j := \frac{w_j}{\chi^*(\mathcal{I})}$. To condense the below calculations, we denote $\chi^*(\mathcal{I})$ by χ^*. By elementary calculus, we may find the optimal solution, but which is too cumbersome to give a analytic formula in a, n, q and χ^*. Using an suboptimal u satisfying the equation $e^{u\chi^*} = 1 + \frac{a}{nq\chi^*}$ and the fact that $(1 + \frac{a}{n})^n \leqslant e^a$ in (9.6) gives

$$\begin{aligned}
\mathbb{P}(Z \geqslant a) &\leqslant \sum_j \frac{w_j}{\chi^*} e^{-\frac{u\chi^* w_jQ_j|\mathcal{I}_j|}{p_j\chi^*}} \Big(e^{\frac{u\chi^* w_j}{p_j\chi^*}} Q_j + 1 - Q_j\Big)^{|\mathcal{I}_j|} e^{-ua} \\
&\leqslant \sum_j \frac{w_j}{\chi^*} \Big(1 + \frac{a}{nq\chi^*}\Big)^{-Q_j|\mathcal{I}_j|} \Big(e^{u\chi^*} Q_j + 1 - Q_j\Big)^{|\mathcal{I}_j|} e^{-ua} \text{ (given } w_j = p_j/\chi^*) \\
&= \sum_j \frac{w_j}{\chi^*} \Big(1 + \frac{a}{nq\chi^*}\Big)^{-Q_j|\mathcal{I}_j|} \Big((1 + \frac{a}{nq\chi^*})Q_j + 1 - Q_j\Big)^{|\mathcal{I}_j|} e^{-ua} \\
&= \sum_j \frac{w_j}{\chi^*} \Big(1 + \frac{a}{nq\chi^*}\Big)^{-Q_j|\mathcal{I}_j|} \Big(1 + \frac{aQ_j}{nq\chi^*}\Big)^{|\mathcal{I}_j|} e^{u\chi^*(-\frac{a}{\chi^*})} \\
&= \sum_j \frac{w_j}{\chi^*} \Big(1 + \frac{a}{nq\chi^*}\Big)^{-Q_j|\mathcal{I}_j|} \Big(1 + \frac{aQ_j}{nq\chi^*}\Big)^{n.\frac{|\mathcal{I}_j|}{n}} e^{u\chi^*(-\frac{a}{\chi^*})} \\
&\leqslant \sum_j \frac{w_j}{\chi^*} \Big(1 + \frac{a}{nq\chi^*}\Big)^{-Q_j|\mathcal{I}_j|} e^{\frac{aQ_j|\mathcal{I}_j|}{nq\chi^*}} e^{u\chi^*(-\frac{a}{\chi^*})} \text{ (as } (1 + \frac{a}{n})^n \leqslant e^a \text{)} \\
&\leqslant \sum_j \frac{w_j}{\chi^*}\ e^{-(\frac{a}{\chi^*} + Q_j|\mathcal{I}_j|)\ln(1 + \frac{a}{nq\chi^*}) + \frac{aQ_j|\mathcal{I}_j|}{nq\chi^*}}
\end{aligned} \tag{9.7}$$

With $k := \frac{a}{nq\chi^*}$, the inequality

$$\ln(1+k) \geqslant k - \frac{k^2}{2} \tag{9.8}$$

is valid for all $k \geqslant 0$. (9.8) applied to (9.7) gives

$$\begin{aligned}
\mathbb{P}(Z \geqslant a) &\leqslant \sum_j \frac{w_j}{\chi^*} e^{-(\frac{a}{\chi^*}+Q_j|\mathcal{I}_j|)(\frac{a}{nq\chi^*}-\frac{a^2}{2n^2q^2(\chi^*)^2})+\frac{aQ_j|\mathcal{I}_j|}{nq\chi^*}} \\
&\leqslant \sum_j \frac{w_j}{\chi^*} e^{-\frac{a^2}{nq(\chi^*)^2}+\frac{a^3}{2n^2q^2(\chi^*)^3}-\frac{aQ_j|\mathcal{I}_j|}{nq\chi^*}+\frac{a^2Q_j|\mathcal{I}_j|}{2n^2q^2(\chi^*)^2}+\frac{aQ_j|\mathcal{I}_j|}{nq\chi^*}} \\
&\leqslant \sum_j \frac{w_j}{\chi^*} e^{-\frac{a^2}{2nq(\chi^*)^2}+\frac{a^3}{2n^2q^2(\chi^*)^3}} \text{ (as } Q_j|\mathcal{I}_j| \leqslant nq) \qquad (9.9) \\
&= e^{-\frac{a^2}{2nq(\chi^*)^2}+\frac{a^3}{2(nq)^2(\chi^*)^3}} \text{ (since } \sum_j w_j = \chi^* \text{)} \qquad (9.10)
\end{aligned}$$

Proof for lower tail

We use the same terminology and set up as above. By application of Markov's inequality we get

$$\begin{aligned}
\mathbb{P}(Z \leqslant -a) &= \mathbb{P}(-Z \geqslant a) \\
&\leqslant \mathbb{P}(e^{-uZ} \geqslant e^{ua}) \\
&\leqslant \frac{\mathbb{E}(e^{-uZ})}{e^{ua}} \qquad (9.11)
\end{aligned}$$

We first estimate the expectation $\mathbb{E}(e^{-uZ})$. As above, with Lemma 9.1, $Z = \sum_j w_j Z_j$. So

$$\begin{aligned}
\exp(-uZ) &= \exp\Big(-\sum_j uw_jZ_j\Big) \\
&= \exp\Big(-\sum_j p_j \frac{uw_jZ_j}{p_j}\Big) \\
&\leqslant \sum_j p_j \exp\Big(-\frac{uw_jZ_j}{p_j}\Big) \text{ (by Jensen's inequality)}
\end{aligned}$$

We further upper bound the expectation of $\exp(-uZ)$

$$\mathbb{E}(e^{-uZ}) \leqslant \sum_j p_j \mathbb{E}\Big(e^{-\frac{uw_jZ_j}{p_j}}\Big)$$

$$
\begin{aligned}
&= \sum_j p_j \prod_{i \in \mathcal{I}_j} \mathbb{E}\Big(e^{-\frac{uw_j Y_i}{p_j}}\Big) \text{ (Z_j is sum of independent random variables)} \\
&= \sum_j p_j \prod_{i \in \mathcal{I}_j} \Big(e^{-\frac{uw_j(1-q_i)}{p_j}} q_i + e^{\frac{uw_j q_i}{p_j}}(1-q_i)\Big) \\
&= \sum_j p_j \prod_{i \in \mathcal{I}_j} e^{\frac{uw_j q_i}{p_j}} \Big(e^{-\frac{uw_j}{p_j}} q_i + 1 - q_i\Big)
\end{aligned}
$$

With $Q_j = \frac{\sum_{i \in \mathcal{I}_j} q_i}{|\mathcal{I}_j|}$ and by the geometric mean/arithmetic mean inequality we get

$$
\mathbb{E}(e^{-uZ}) \leqslant \sum_j p_j e^{\frac{uw_j Q_j |\mathcal{I}_j|}{p_j}} \Big(e^{-\frac{uw_j}{p_j}} Q_j + 1 - Q_j\Big)^{|\mathcal{I}_j|} \tag{9.12}
$$

Using the inequality

$$
1 + \lambda \leqslant e^{\lambda}
$$

for any λ, we get

$$
\Big(e^{-\frac{uw_j}{p_j}} Q_j + 1 - Q_j\Big) = 1 + (e^{-\frac{uw_j}{p_j}} - 1)Q_j \leqslant e^{\left(e^{-\frac{uw_j}{p_j}} - 1\right)Q_j} \tag{9.13}
$$

We now employ the inequality,

$$
e^{-\lambda} \leqslant 1 - \lambda + \frac{\lambda^2}{2}
$$

for $\lambda > 0$, so

$$
e^{\left(e^{-\frac{uw_j}{p_j}} - 1\right)Q_j} \leqslant e^{\left(-\frac{uw_j}{p_j} + \frac{u^2 w_j^2}{2p_j^2}\right)Q_j} \tag{9.14}
$$

We apply (9.13) and (9.14) to (9.12)

$$
\begin{aligned}
\mathbb{E}(e^{-uZ}) &\leqslant \sum_j p_j e^{\frac{uw_j Q_j |\mathcal{I}_j|}{p_j}} e^{\left(e^{-\frac{uw_j}{p_j}} - 1\right)Q_j |\mathcal{I}_j|} \text{ (by 9.13)} \\
&\leqslant \sum_j p_j e^{\frac{uw_j Q_j |\mathcal{I}_j|}{p_j}} e^{\left(-\frac{uw_j}{p_j} + \frac{u^2 w_j^2}{2p_j^2}\right)Q_j |\mathcal{I}_j|} \text{ (by 9.14)}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_j p_j e^{\frac{uw_jQ_j|\mathcal{I}_j|}{p_j}} e^{-\frac{uw_jQ_j|\mathcal{I}_j|}{p_j}} e^{\frac{u^2w_j^2Q_j|\mathcal{I}_j|}{2p_j^2}} \\
&= \sum_j p_j e^{\frac{u^2w_j^2Q_j|\mathcal{I}_j|}{2p_j^2}} \qquad (9.15)
\end{aligned}
$$

Now we choose $p_j := \frac{w_j\sqrt{Q_j|\mathcal{I}_j|}}{T}$, where $T = \sum_j w_j\sqrt{Q_j|\mathcal{I}_j|}$. Substituting the p_j's in expression (9.15) we get

$$
\begin{aligned}
\mathbb{E}(e^{-uZ}) &\leqslant \sum_j p_j e^{\frac{u^2w_j^2Q_j|\mathcal{I}_j|T^2}{2w_j^2Q_j|\mathcal{I}_j|}} \\
&= \sum_j p_j e^{\frac{u^2T^2}{2}} \\
&\leqslant e^{\frac{u^2T^2}{2}} \text{ (as } \sum_j p_j = 1) \qquad (9.16)
\end{aligned}
$$

(9.16) applied to (9.11) gives

$$
\mathbb{P}(Z \leqslant -a) \leqslant e^{\frac{u^2T^2}{2}-ua} \qquad (9.17)
$$

Note that u is a free parameter. We set $u = \frac{a}{T^2}$ to minimize the RHS of (9.17) leading to

$$
\begin{aligned}
\mathbb{P}(Z \leqslant -a) &\leqslant e^{\frac{a^2.T^2}{2.T^4}-\frac{a.a}{T^2}} \\
&= e^{-\frac{a^2}{2T^2}} \qquad (9.18)
\end{aligned}
$$

As defined earlier $\chi^*(\mathcal{I}) = \sum_j w_j$. By the Cauchy-Schwarz inequality and Lemma 9.1, we get

$$
\begin{aligned}
T^2 &= \Big(\sum_j w_j\sqrt{Q_j|\mathcal{I}_j|}\Big)^2 \\
&= (\sum_j \sqrt{w_j}\sqrt{w_jQ_j|\mathcal{I}_j|}\Big)^2 \\
&\leqslant \underbrace{\sum_j w_j}_{\chi^*(\mathcal{I})} \sum_j w_jQ_j|\mathcal{I}_j| \\
&= \chi^*(\mathcal{I}) \sum_j w_j \sum_{i\in\mathcal{I}_j} q_i \text{ (as } Q_j = \sum_{i\in\mathcal{I}_j}\frac{q_i}{|\mathcal{I}_j|})
\end{aligned}
$$

$$= \chi^*(\mathcal{I}) \sum_{i \in \mathcal{I}} q_i \text{ (by Lemma 9.1)}$$
$$= nq\chi^*(\mathcal{I}) \text{ (since } q = \sum_{i \in \mathcal{I}} \frac{q_i}{n}) \tag{9.19}$$

So (9.18) reduces to,

$$\mathbb{P}(Z \leqslant -a) \leqslant e^{-\frac{a^2}{2nq\chi^*(\mathcal{I})}} \tag{9.20}$$

$\square$

Remark 9.4. Let $\Delta(\mathcal{D})$ be maximum degree of graph $\mathcal{D}$ and $\Delta_1(\mathcal{D}) := \Delta(\mathcal{D}) + 1$. We know by Brooks theorem [Bro41]that $\chi(\mathcal{D}) \leqslant \Delta_1(\mathcal{D})$. Thus

$$\chi^*(\mathcal{I}) \leqslant \chi^*(\mathcal{D}) \leqslant \chi(\mathcal{D}) \leqslant \Delta_1(\mathcal{D}) \tag{9.21}$$

Hence we can replace $\chi^*(\mathcal{I})$ by any of the quantity $\chi^*(\mathcal{D}), \chi(\mathcal{D}), \Delta_1(\mathcal{D})$ in Theorem 9.3 without much detoriation to the sharpness of the bound.

Remark 9.5. In case the optimal exact fractional cover of $\mathcal{I}$ is given, then we can use the $\chi^*(\mathcal{I})$ in our result. Since the calculation of optimal exact fractional cover can be modelled as a set cover problem so finding the exact fractional cover is a intractable problem. By definition, the exact fractional cover is calculated for the indpendent sets and even enumerating all the independent sets of $\mathcal{I}$ has exponential complexity as number of independent sets $\geqslant 2^{\alpha(D)}$, where $\alpha(D)$ is the independence number of D.

Chapter 10

Derandomizing the Alon-Spencer Concentration Inequality

In Chapter 9, we stated the Alon-Spencer bound (Theorem 9.2). In this chapter, we derandomize the Alon-Spencer bound.

Let Y_j, $j \in [n]$, be independent random variables defined by $Y_j = 1 - q_j$ with probability q_j and $Y_j = -q_j$ with probability $1 - q_j$, where $q_j \in [0, 1]$ for all j. Let $q := \frac{\sum_j q_j}{n}$.

Definition 10.1. For $i \in [m]$, let ψ_i denote the random variable $\psi_i := \sum_{j=1}^n a_{ij} Y_j$ where $a_{ij} \in \{0, 1\}$ for all i, j. Let $a_i > 0$ be rational numbers and for $i \in [m]$, let $E_i^{(+)}$ denote the event $\psi_i \leqslant a_i$ and let $E_i^{(-)}$ denote the event $\psi_i \geqslant -a_i$. Furthermore set $E = \bigcap_{i=1}^m E_i$ where E_i is either $E_i^{(+)}$ or $E_i^{(-)}$. Let $n_i = \sum_{j=1}^n a_{ij}$ for all $i \in [m]$.

For each event E_i, let $f(E_i^c)$ be the upper bound on $\mathbb{P}(E_i^c)$ as in Theorem 9.2, so $f(E_i^c) = e^{-\frac{a^2}{2n_i q} + \frac{a^3}{2(n_i q)^2}}$ or $f(E_i^c) = e^{-\frac{a^2}{2n_i q}}$. We assume that for some $0 < \delta < 1$

$$\sum_{i=1}^m f(E_i^c) \leqslant 1 - \delta, \tag{10.1}$$

so using the union bound, we get

$$\mathbb{P}(\bigcap_{i=1}^m E_i) \geqslant 1 - \mathbb{P}(\bigcup_{i=1}^m E_i^c) \geqslant 1 - \sum_{i=1}^m \mathbb{P}(E_i^c) \geqslant \delta > 0, \text{ thus } \bigcap_{i=1}^m E_i \neq \varnothing. \tag{10.2}$$

The problem in derandomization is to construct a vector $x \in \bigcap_{i=1}^m E_i$ in polynomial time.

10.1 Pessimistic Estimators

We first consider the event $E_i = E_i^+$ and define the weak pessmistic estimator for the event E_i by going through the proof of the Alon-Spencer bound [AS04], Corollary A.1.8/Theorem A.1.11.

Definition 10.2. For each $i \in [m]$, let V_i be a family of functions $(V_{il})_l$, $l \in [n] \cup \{0\}$ defined as follows. For $y_1, \ldots, y_l$ with $y_j \in \{-q_j, 1-q_j\}$ with $j \in \{1, 2, \ldots, l\}$, $u_i > 0$ for $i \in [m]$, set

(i) $V_{il}(y_1, \ldots, y_l) = e^{-u_i a_i} e^{\sum_{j=1}^{l} u_i a_{ij} y_j} \prod_{j=l+1}^{n} \left(q_j e^{u_i a_{ij}(1-q_j)} + (1-q_j) e^{-u_i a_{ij} q_j} \right)$

(ii) $V_{i0} = e^{-\frac{a^2}{2n_i q} + \frac{a^3}{2(n_i q)^2}}$

We now prove that the family V_i is a weak pessimistic estimator for the event E_i for $i \in [m]$.

Lemma 10.3. Under the assumption (10.2), the family $V_i = (V_{il})_l$, $l \in [n] \cup \{0\}$, is a weak pessimistic estimator for the event E_i.

Proof. We need to check conditions (a)-(c) of the definition of weak pessmistic estimator (Definition 3.8).

Let $l \in [n] \cup \{0\}$ be arbitrary, but fixed and let $y_1, \ldots, y_l$ with $y_j \in \{-q_j, 1-q_j\}$ for $j \in \{1, 2, \ldots, l\}$. We condition on $Y_j = y_j$ for $l \in \{1, 2, \ldots, n\}$.

Condition (a): $\mathbb{P}(E_i^c \mid y_1, \ldots, y_l) \leqslant V_{il}(y_1, \ldots, y_l)$. Here is the proof. Let $u_i > 0$ the parameter as in Definition 10.2.

$$
\begin{aligned}
\mathbb{P}(E_i^c \mid y_1, \ldots, y_l) &= \mathbb{P}(\psi_i \geqslant a_i \mid y_1, \ldots, y_l) \\
&= \mathbb{P}\Big(\sum_{j=1}^{n} a_{ij} Y_j \geqslant a_i \mid y_1, \ldots, y_l \Big) \\
&\leqslant \mathbb{E}\left[e^{u_i \sum_{j=1}^{n} a_{ij} Y_j} \geqslant e^{u_i a_i} \mid y_1, \ldots, y_l \right] e^{-u_i a_i} \text{ (by Markov's inequality)} \\
&= \mathbb{E}\left[e^{\sum_{j=1}^{l} u_i a_{ij} y_j + \sum_{j=l+1}^{n} u_i a_{ij} Y_j} \right] e^{-u_i a_i} \\
&= e^{-u_i a_i} e^{\sum_{j=1}^{l} u_i a_{ij} y_j} \mathbb{E}\left[e^{\sum_{j=l+1}^{n} u_i a_{ij} Y_j} \right] \\
&= e^{-u_i a_i} e^{\sum_{j=1}^{l} u_i a_{ij} y_j} \prod_{j=l+1}^{n} \left(q_j e^{u_i a_{ij}(1-q_j)} + (1-q_j) e^{-u_i a_{ij} q_j} \right) \\
&= V_{il}
\end{aligned}
$$

Condition (b):

From Definition 10.2, we know that

$$
V_{i,l+1}(y_1, \ldots, y_l, y_{l+1}) = e^{-u_i a_i} e^{\sum_{j=1}^{l+1} u_i a_{ij} y_j} \prod_{j=l+2}^{n} \left(q_j e^{u_i a_{ij}(1-q_j)} + (1-q_j) e^{-u_i a_{ij} q_j} \right)
$$

To shorten the notation we define

$$V_{i,l+1}(1-q_{l+1}) := V_{i,l+1}(y_1,\dots,y_l,1-q_{l+1}) \tag{10.3}$$

Further, we define

$$\begin{aligned}V_{i,l+1}(-q_{l+1}) &:= V_{i,l+1}(y_1,\dots,y_l,-q_{l+1})\\ &= e^{-u_i a_i} e^{\sum_{j=1}^{l} u_i a_{ij} y_j - u_i a_{i,l+1} q_{l+1}} \prod_{j=l+2}^{n} \left(q_j e^{u_i a_{ij}(1-q_j)} + (1-q_j) e^{-u_i a_{ij} q_j}\right)\end{aligned} \tag{10.4}$$

Now let us define $V_{i,l+1}^{con}$ as the convex combination of $V_{i,l+1}(1-q_{l+1})$ and $V_{i,l+1}(-q_{l+1})$, so

$$V_{i,l+1}^{con} = q_{l+1} V_{i,l+1}(1-q_{l+1}) + (1-q_{l+1}) V_{i,l+1}(-q_{l+1}) \tag{10.5}$$

Now we have

$$\begin{aligned}V_{i,l+1}^{con} &= e^{-u_i a_i}\left(q_{l+1} e^{\sum_{j=1}^{l} u_i a_{ij} y_j + u_i a_{i,l+1}(1-q_{l+1})} + (1-q_{l+1}) e^{\sum_{j=1}^{l} u_i a_{ij} y_j - u_i a_{i,l+1} q_{l+1}}\right)\\ &\quad \prod_{j=l+2}^{n} \left(q_j e^{u_i a_{ij}(1-q_j)} + (1-q_j) e^{-u_i a_{ij} q_j}\right)\\ &= e^{-u_i a_i} e^{\sum_{j=1}^{l} u_i a_{ij} y_j}\left(q_{l+1} e^{u_i a_{i,l+1}(1-q_{l+1})} + (1-q_{l+1}) e^{-u_i a_{i,l+1} q_{l+1}}\right)\\ &\quad \prod_{j=l+2}^{n} \left(q_j e^{u_i a_{ij}(1-q_j)} + (1-q_j) e^{-u_i a_{ij} q_j}\right)\\ &= e^{-u_i a_i} e^{\sum_{j=1}^{l} u_i a_{ij} y_j} \prod_{j=l+1}^{n} \left(q_j e^{u_i a_{ij}(1-q_j)} + (1-q_j) e^{-u_i a_{ij} q_j}\right)\\ &= V_{il}(y_1,\dots,y_l)\end{aligned}$$

Thus $\min\left(V_{i,l+1}(1-q_{l+1}), V_{i,l+1}(-q_{l+1})\right) \leqslant V_{il}(y_1,\dots,y_l)$.

Condition (c): By Definition 10.2, $V_{i0} = e^{-\frac{a^2}{2n_i q} + \frac{a^3}{2(n_i q)^2}}$ and we have

$$V_{i0} \leqslant 1-\delta.$$

□

For sake of completeness, we now define the pessimistic estimators for the event $E_i = E_i^{(-)}$ by passing through the proof of Theorem A.1.13 [AS04].

Definition 10.4. For each $i \in [m]$, let V_i be a family of functions $(V_{il})_l$, $l \in [n] \cup \{0\}$, defined as follows. For $y_1,\dots,y_l \in \{-q_l, 1-q_l\}$, $v_i > 0$ for $i \in [m]$, set

(i) $V_{il}(y_1, \ldots, y_l) = e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j} \prod_{j=l+1}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j) e^{v_i a_{ij} q_j} \right)$

(ii) $V_{i0} = e^{-\frac{a^2}{2n_i q}}$

Lemma 10.5. The family V_i is a weak pessmistic estimator for the event $E_i = E_i^{(-)}$.

Proof. We need to check conditions (a)-(c) of the definition of weak pessmistic estimator (Definition 3.8).

Let $l \in [n] \cup \{0\}$ be arbitrary, but fixed and let $y_1, \ldots, y_l$ with $y_j \in \{-q_j, 1-q_j\}$ for $j \in \{1, 2, \ldots, l\}$. We condition on $Y_j = y_j$ for $l \in \{1, 2, \ldots, n\}$.

Condition (a): $\mathbb{P}(E_i^c \mid y_1, \ldots, y_l) \leqslant V_{il}(y_1, \ldots, y_l)$. Here is the proof. Let $v_i > 0$ the parameter as in Definition 10.4.

$$
\begin{aligned}
\mathbb{P}(E_i^c \mid y_1, \ldots, y_l) &= \mathbb{P}(\psi_i \leqslant -a_i \mid y_1, \ldots, y_l) \\
&= \mathbb{P}\Big(-\sum_{j=1}^{n} a_{ij} Y_j \geqslant a_i \mid y_1, \ldots, y_l \Big) \\
&\leqslant \mathbb{E}\left[e^{-v_i \sum_{j=1}^{n} a_{ij} Y_j} \geqslant e^{v_i a_i} \mid y_1, \ldots, y_l \right] e^{-v_i a_i} \text{ (by Markov's inequality)} \\
&= \mathbb{E}\left[e^{-\sum_{j=1}^{l} v_i a_{ij} y_j - \sum_{j=l+1}^{n} v_i a_{ij} Y_j} \right] e^{-v_i a_i} \\
&= e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j} \mathbb{E}\left[e^{\sum_{j=l+1}^{n} -v_i a_{ij} Y_j} \right] \\
&= e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j} \prod_{j=l+1}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j) e^{v_i a_{ij} q_j} \right) \\
&= V_{il}
\end{aligned}
$$

Condition (b):

From Definition 10.2, we know that

$$
V_{i,l+1}(y_1, \ldots, y_l, y_{l+1}) = e^{-v_i a_i} e^{\sum_{j=1}^{l+1} -v_i a_{ij} y_j} \prod_{j=l+2}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j) e^{v_i a_{ij} q_j} \right)
$$

To shorten the notation we define

$$
V_{i,l+1}(1-q_{l+1}) := V_{i,l+1}(y_1, \ldots, y_l, 1-q_{l+1}) \tag{10.6}
$$

Further, we define

$$\begin{aligned} V_{i,l+1}(-q_{l+1}) &:= V_{i,l+1}(y_1, \ldots, y_l, -q_{l+1}) \\ &= e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j + v_i a_{i,l+1} q_{l+1}} \prod_{j=l+2}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j) e^{v_i a_{ij} q_j} \right) \end{aligned} \tag{10.7}$$

Now let us define $V_{i,l+1}^{con}$ as the convex combination of $V_{i,l+1}(1-q_{l+1})$ and $V_{i,l+1}(-q_{l+1})$, so

$$V_{i,l+1}^{con} = q_{l+1} V_{i,l+1}(1-q_{l+1}) + (1-q_{l+1}) V_{i,l+1}(-q_{l+1}) \tag{10.8}$$

Now we have

$$\begin{aligned} V_{i,l+1}^{con} &= e^{-v_i a_i} \left(q_{l+1} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j - v_i a_{i,l+1}(1-q_{l+1})} + (1-q_{l+1}) e^{\sum_{j=1}^{l} -v_i a_{ij} y_j + v_i a_{i,l+1} q_{l+1}} \right) \\ &\quad \prod_{j=l+2}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j) e^{v_i a_{ij} q_j} \right) \\ &= e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j} \left(q_{l+1} e^{-v_i a_{i,l+1}(1-q_{l+1})} + (1-q_{l+1}) e^{v_i a_{i,l+1} q_{l+1}} \right) \\ &\quad \prod_{j=l+2}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j) e^{v_i a_{ij} q_j} \right) \\ &= e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j} \prod_{j=l+1}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j) e^{v_i a_{ij} q_j} \right) \\ &= V_{il}(y_1, \ldots, y_l) \end{aligned}$$

Thus $\min\left(V_{i,l+1}(1-q_{l+1}), V_{i,l+1}(-q_{l+1})\right) \leqslant V_{il}(y_1, \ldots, y_l)$.

Condition (c): By Definition 10.2, $V_{i0} = e^{-\frac{a^2}{2n_i q}}$ and we have

$$V_{i0} \leqslant 1 - \delta.$$

□

Let $E_1, E_2, \ldots, E_m$ be events, where either $E_i = E_i^{(+)}$ or $E_i = E_i^{(-)}$, $i \in [m]$. Let $E = \bigcap_{i=1}^m E_i$. We assume that E satisfies (10.1) of Definition 10.1, so $\sum_{i=1}^m f(E_i^c) \leqslant 1-\delta$ for some $0 < \delta < 1$. Thus, by (10.2), $E \neq \varnothing$.

Definition 10.6. Let $\mathcal{V} = \bigoplus_{i=1}^m V_i$ be the family of functions, serving as the candidate for a weak pessimistic estimator for the event $E = \bigcap_{i=1}^m E_i$.

Indeed,

Lemma 10.7. $\mathcal{V}$ is a weak pessimistic estimator for the event $E = \bigcap_{i=1}^m E_i$.

Proof. We apply Proposition 3.10, and must show $\sum_{i=1}^m V_{i0} \leqslant 1-\delta$ and the convexity of $\mathcal{V}$.

By Definition 10.2/10.4, $V_{i0} = e^{-\frac{a^2}{2n_iq}+\frac{a^3}{2(n_iq)^2}}$ for $E_i = E_i^{(+)}$ and $V_{i0} = e^{-\frac{a^2}{2n_iq}}$ for $E_i = E_i^{(-)}$. By (10.1) of Definition 10.1, we have $\sum_{i=1}^m V_{i0} \leqslant 1-\delta$ for some $0<\delta<1$.

For the convexity of $\mathcal{V}$, we must show for any $l \in [n]$ and $y_j \in \{q_j, 1-q_j\}$ for $j \in [l-1]$ that there are real numbers μ_j, j=1,2, with $\mu_1+\mu_2=1$ such that

$$\mu_1 \sum_{i=1}^m V_{il}(y_1,\dots,y_{l-1},-q_l) + \mu_2 \sum_{i=1}^m V_{il}(y_1,\dots,y_{l-1},1-q_l) \leqslant \sum_{i=1}^m V_{i,l-1}(y_1,\dots,y_{l-1}) \quad (10.9)$$

Let $\mu_1 = 1-q_l$ and $\mu_2 = q_l$. From Lemma 10.3 / 10.5, we have

$$(1-q_l)V_{il}(y_1,\dots,y_{l-1},-q_l) + q_l V_{il}(y_1,\dots,y_{l-1},1-q_l) \leqslant V_{i,l-1}(y_1,\dots,y_{l-1})$$

.

Summing all terms in the above expression for $i \in [m]$, we get requisite condition (10.9).

□

We are ready to state the main theorem, which we call the algorithmic or derandomized form of the Alon-Spencer bound.

Theorem 10.8. Let $0<\delta<1$ and $E_1, E_2, \cdots, E_m$ be the events estimated by the Alon-Spencer bound which satisfy (10.1). We further assume that the probabilities of the underlying random setting in Definition 10.1 stay away from zero, i.e. for all $j \in [n]$, $q_j = \Omega(n^{-k})$ for some constant $k>0$. Then a vector $x \in \bigcap_{i=1}^m E_i$ can be constructed in $\mathcal{O}(mn^2(n\ln(n)+\ln(\frac{mn}{\delta})))$.

Proof. We approximate the functions of the weak pessimistic estimator $\mathcal{V}$ for the event E by suitable Taylor polynomials and then define the pessimistic estimator for the event E. Let's start with an event of the form $E_i = E_i^{(+)}$. The functions V_{il}, $l \in [n]$, $i \in [m]$ have the form

$$\begin{aligned} V_{il}(y_1,y_2,\cdots,y_l) &= e^{-u_ia_i} e^{\sum_{j=1}^l u_ia_{ij}y_j} \prod_{j=l+1}^n \left(q_j e^{u_ia_{ij}(1-q_j)} + (1-q_j)e^{-u_ia_{ij}q_j} \right) \\ &= e^{-u_ia_i} e^{\sum_{j=1}^l u_ia_{ij}y_j} \prod_{j=l+1}^n \mathbb{E}\left[e^{u_ia_{ij}Y_j} \right] \end{aligned}$$

By Theorem A.1.9 [AS04], the optimal choice of u_i satisfies $e^{u_i} = b_i := 1 + \frac{a_i}{\sum_{j=1}^n a_{ij}q_j}$. Set

$$Q_{il} := \prod_{j\,\in\,\{[n]\,\cup\,\{0\}\}} \mathbb{E}\left[e^{\ln(b_i)c_{ij}}\right] \tag{10.10}$$

where

$$c_{ij} = \begin{cases} -a_i & : j = 0 \\ a_{ij}y_j & : j = 1, 2, \cdots, l \\ a_{ij}Y_j & : j = l+1, 2, \cdots, n \end{cases}$$

So,

$$V_{il}(y_1, \ldots, y_l) = Q_{il} \tag{10.11}$$

Note that Y_j is a random variable, so c_{ij} is also a random variable for $j = l+1, l+2, \cdots, n$. Set $\gamma = \frac{\delta}{6mn}$. Let $T(c_{ij}d_i)$ be the N-th degree Taylor polynomial of the exponential function $e^{\ln(b_i)c_{ij}}$, where d_i be a rational number approximating $\ln(b_i)$ for $i \in [m]$. Now we establish the value for N using Lemma 2.8 (i) of [SS96] such that the following approximation

$$|\prod_{j\,\in\,\{[n]\,\cup\,\{0\}\}} e^{\ln(b_i)c_{ij}} - \prod_{j\,\in\,\{[n]\,\cup\,\{0\}\}} T(c_{ij}d_i| \leqslant \gamma \tag{10.12}$$

uniformly holds for all c_{ij} depending upon $y_1, y_2, \cdots, y_l$.

Clearly, $|c_{ij}| \leqslant 1$ for $j \in \{1, 2, \cdots, l\}$ and $|c_{i0}| \leqslant n$. Hence $n + \sum_{j=1}^{l} |c_{ij}| \leqslant 2n = P$. Also $b_i = 1 + \frac{a_i}{\sum_{j=1}^{n} a_{ij}q_j} \leqslant 1 + \frac{n_i}{\sum_{j=1}^{n} a_{ij}q_j} \leqslant \mathcal{O}(n^k) = Q$ by our assumption $q_j = \Omega(n^{-k})$, for all $j \in [n]$ and $k > 0$ constant. By Lemma 2.8 (i) of [SS96], with $N = \mathcal{O}\Big(n\ln(n) + \ln(\frac{n+1}{\gamma})\Big)$ a rational number d_i approximating $\ln(b_i)$ and the numbers $T(c_{ij}d_i)$ for all $i \in [m], j \in \{[n] \cup \{0\}\}$ can be computed in $\mathcal{O}\Big(n\ln(n) + \ln(\frac{1}{\gamma})\Big)$ time so that (10.12) holds.

Given the independence of random variables Y_j, $j \in [n]$, taking expectation on (10.12), we get

$$\left|\prod_{j\,\in\,\{[n]\,\cup\,\{0\}\}} \mathbb{E}\left[e^{\ln(b_i)c_{ij}}\right] - \prod_{j\,\in\,\{[n]\,\cup\,\{0\}\}} \mathbb{E}\left[T(c_{ij}d_i)\right]\right| \leqslant \gamma \tag{10.13}$$

Since the Y_j, $j \in \{l+1, \ldots, n\}$ takes only two values, each factor of the product $\prod_{j\,\in\,\{[n]\,\cup\,\{0\}\}} \mathbb{E}\left[T(c_{ij}d_i)\right]$ can be calculated in $\mathcal{O}\Big(n\ln(n) + \ln(\frac{1}{\gamma})\Big)$ time.

Now let us consider the event $E_i = E_i^{(-)}$. For sake of completeness, we repeat the arguments.

$$V_{il}(y_1, \ldots, y_l) = e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j} \prod_{j=l+1}^{n} \left(q_j e^{-v_i a_{ij}(1-q_j)} + (1-q_j)e^{v_i a_{ij} q_j}\right)$$

$$= e^{-v_i a_i} e^{\sum_{j=1}^{l} -v_i a_{ij} y_j} \prod_{j=l+1}^{n} \mathbb{E}\left[e^{-v_i a_{ij} Y_j}\right]$$

By Theorem A.1.13 [AS04], the optimal choice of v_i is $v_i = f_i := \frac{a_i}{\sum_{j=1}^{n} a_{ij} q_j}$. Set

$$Q_{il} := \prod_{j \in \{[n] \cup \{0\}\}} \mathbb{E}\left[e^{f_i c_{ij}}\right] \tag{10.14}$$

where

$$c_{ij} = \begin{cases} -a_i & : j = 0 \\ -a_{ij} y_j & : j = 1, 2, \cdots, l \\ -a_{ij} Y_j & : j = l+1, 2, \cdots, n \end{cases}$$

So,

$$V_{il}(y_1, \ldots, y_l) = Q_{il} \tag{10.15}$$

By our assumption, $f_i = \frac{a_i}{\sum_{j=1}^{n} a_{ij} q_i} = \mathcal{O}(n^k)$. Clearly, $|c_{ij}| \leqslant 1$ for $j \in \{1, 2, \cdots, l\}$ and $|c_{i0}| \leqslant n$. Hence $n + \sum_{j=1}^{l} |c_{ij}| \leqslant 2n = P$. By Lemma 2.8 (ii) of [SS96], $N = \mathcal{O}\Big(n + \ln(\frac{n+1}{\gamma})\Big)$ and the computation of the N-th degree Taylor polynomial $T(f_i c_{ij})$ of the exponential function $e^{f_i c_{ij}}$ can be done in $\mathcal{O}\Big(n + \ln(\frac{1}{\gamma})\Big)$ time.

Set $T_{il} := \prod_{j \in \{[n] \cup \{0\}\}} \mathbb{E}\left[T(c_{ij} d_i)\right]$ or $\prod_{j \in \{[n] \cup \{0\}\}} \mathbb{E}\left[T(c_{ij} f_i)\right]$.

Hence, for all $i \in [m]$, $l \in [n]$,

$$|V_{il} - T_{il}| \leqslant \gamma \text{ and } |V_l - T_l| \leqslant m\gamma = \gamma^{'} \tag{10.16}$$

We invoke Proposition 3.12 with $\gamma^{'}$ as the approximation error. Now $\gamma^{'} = m\gamma = m.\frac{\delta}{6mn} = \frac{\delta}{6n} < \frac{\delta}{4n+1}$, so $\gamma^{'}$ satisfies the assumption of Proposition 3.12. By Proposition 3.12,

$$U_l = T_l + (2n - l)\gamma^{'}, l \in \{[n] \cup 0\}$$

is a pessimistic estimator for the event E.

We now fix the computation time for the l^{th} round of sequential derandomization. For each Q_{il}, there are at most $n+1$ Taylor approximations. So the computation time for all m summands will be $\mathcal{O}(mn(n\ln(n) + \ln(\frac{mn}{\delta})))$. For $n+1$ rounds, the total computation time is $\mathcal{O}(mn^2(n\ln(n) + \ln(\frac{mn}{\delta})))$.

□

Remark 10.9. The above proof assumes mutually independent random variables. In case we assume partial dependency structure as we defined in chapter 9, proving convexity of the pessimistic estimators remains an open problem.

Declaration

I declare that:

- apart from my supervisor's guidance - the content and design of the thesis is all my own work and only using the sources listed in the thesis.
- I have not submitted the thesis either partially or wholly as part of a doctoral examination procedure to another examining body and neither it has been published or submitted for publication.
- the thesis has been prepared subject to the Rules of Good Scientific Practice of the German Research Foundation.
- no academic degree has ever been withdrawn for myself.

Regards

Mayank

Mayank

Bibliography

[AS02] Nitin Ahuja and Anand Srivastav. "On constrained hypergraph coloring and scheduling". In: *Approximation Algorithms for Combinatorial Optimization, 5th International Workshop, APPROX 2002, Rome, Italy, September 17-21, 2002, Proceedings*. 2002, pp. 14–25. DOI: 10.1007/3-540-45753-4_4. URL: http://dx.doi.org/10.1007/3-540-45753-4_4.

[AS04] N. Alon and J.H. Spencer. "Appendix a: bounding of large deviations". In: *The Probabilistic Method*. John Wiley & Sons, Ltd, 2004, pp. 263–273. ISBN: 9780471722151. DOI: 10.1002/0471722154.app1. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/0471722154.app1. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/0471722154.app1.

[Azu67] Kazuoki Azuma. "Weighted sums of certain dependent random variables". In: *Tohoku Math. J. (2)* 19.3 (1967), pp. 357–367. DOI: 10.2748/tmj/1178243286. URL: http://dx.doi.org/10.2748/tmj/1178243286.

[Ban10] Nikhil Bansal. "Constructive algorithms for discrepancy minimization". In: *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*. 2010, pp. 3–10.

[BB96] Robert B. Burckel and Heinz Bauer. *Probability theory*. De Gruyter, 1996.

[BCS09] Nikhil Bansal, Alberto Caprara, and Maxim Sviridenko. "A new approximation method for set covering problems, with applications to multidimensional bin packing". In: *SIAM J. Comput.* 39.4 (Oct. 2009), pp. 1256–1278. ISSN: 0097-5397. DOI: 10.1137/080736831. URL: http://dx.doi.org/10.1137/080736831.

[Bec81] József Beck. "Van der waerden and ramsey type games". English (US). In: *Combinatorica* 1.2 (June 1981), pp. 103–116. ISSN: 0209-9683. DOI: https://doi.org/10.1007/BF02579267.

[Bec82] József Beck. "Remarks on positional games. i". In: *Acta Mathematica Academiae Scientiarum Hungarica* 40 (1982), pp. 65–71.

[Bec85] József Beck. "Random graphs and positional games on the complete graph". In: *Random Graphs '83*. Ed. by Michal Karoński and Andrzej Ruciński. Vol. 118. North-Holland Mathematics Studies. North-Holland, 1985, pp. 7–13. DOI: https://doi.org/10.1016/S0304-0208(08)73609-0. URL: http://www.sciencedirect.com/science/article/pii/S0304020808736090.

[Bec91] József Beck. "An algorithmic approach to the lovász local lemma. i". In: *Random Structures & Algorithms* 2.4 (1991), pp. 343–365. ISSN: 1098-2418. DOI: 10.1002/rsa.3240020402. URL: http://dx.doi.org/10.1002/rsa.3240020402.

[Bec94] József Beck. "Deterministic graph games and a probabilistic intuition". In: *Combinatorics, Probability and Computing* 3.1 (1994), pp. 13–26. DOI: 10.1017/S0963548300000936.

[BF81] József Beck and Tibor Fiala. "Integer-making theorems". In: *Discrete Applied Mathematics* 3.1 (1981), pp. 1–8. ISSN: 0166-218X. DOI: http://dx.doi.org/10.1016/0166-218X(81)90022-6. URL: http://www.sciencedirect.com/science/article/pii/0166218X81900226.

[BŁ00] Małgorzata Bednarska and Tomasz Łuczak. "Biased positional games for which random strategies are nearly optimal". In: *Combinatorica* 20.4 (Apr. 2000), pp. 477–488. ISSN: 1439-6912. DOI: 10.1007/s004930070002. URL: https://doi.org/10.1007/s004930070002.

[Bre76] Richard P. Brent. "Fast multiple-precision evaluation of elementary functions". In: *J. ACM* 23.2 (Apr. 1976), pp. 242–251. ISSN: 0004-5411. DOI: 10.1145/321941.321944. URL: http://doi.acm.org/10.1145/321941.321944.

[Bro41] R. L. Brooks. "On colouring the nodes of a network". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 37.2 (1941), pp. 194–197. DOI: 10.1017/S030500410002168X.

[BS11] Nikhil Bansal and Joel Spencer. "Deterministic discrepancy minimization". In: *Algorithms - ESA 2011 - 19th Annual European Symposium, Saarbrücken, Germany, September 5-9, 2011. Proceedings*. 2011, pp. 408–420.

[CE78] V. Chvátal and P. Erdős. "Biased positional games". In: *Algorithmic Aspects of Combinatorics*. Ed. by B. Alspach, P. Hell, and D.J. Miller. Vol. 2. Annals of Discrete Mathematics. Elsevier, 1978, pp. 221–229. DOI: https://doi.org/10.1016/S0167-5060(08)70335-2. URL: http://www.sciencedirect.com/science/article/pii/S0167506008703352.

[Che52] Herman Chernoff. "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations". In: *The Annals of Mathematical Statistics* 23.4 (1952), pp. 493–507. URL: http://www.jstor.org/stable/2236576.

[CKP+17] Henrik I. Christensen, Arindam Khan, Sebastian Pokutta, and Prasad Tetali. "Approximation and online algorithms for multidimensional bin packing: a survey". In: *Computer Science Review* 24 (2017), pp. 63–79. ISSN: 1574-0137. DOI: https://doi.org/10.1016/j.cosrev.2016.12.001. URL: http://www.sciencedirect.com/science/article/pii/S1574013716301356.

[EMS14] Mourad El Ouali, Peter Munstermann, and Anand Srivastav. “Randomized approximation for the set multicover problem in hypergraphs”. English. In: *Algorithmica* (2014), pp. 1–15. ISSN: 0178-4617. DOI: 10.1007/s00453-014-9962-9. URL: http://dx.doi.org/10.1007/s00453-014-9962-9.

[ER59] Paul Erdős and A. Rényi. “On random graphs i”. In: *Publ. Math. (Debrecen)* 6 (1959), p. 290.

[Erd63] Paul Erdős. “On a combinatorial problem”. In: *Nordisk Matematisk Tidskrift* 11.1 (1963), pp. 5–10. ISSN: 00291412. URL: http://www.jstor.org/stable/24524364.

[ES73] P Erdős and J.L Selfridge. “On a combinatorial game”. In: *Journal of Combinatorial Theory, Series A* 14.3 (1973), pp. 298–301.

[Fer81] G. S. Fernandez de la Vega W.and Lueker. “Bin packing can be solved within $1 + \epsilon$ in linear time”. In: *Combinatorica* 1.4 (Dec. 1981), pp. 349–355. ISSN: 1439-6912. DOI: 10.1007/BF02579456. URL: https://doi.org/10.1007/BF02579456.

[FK15] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2015. DOI: 10.1017/CBO9781316339831.

[Gil59] E. N. Gilbert. “Random graphs”. In: *Ann. Math. Statist.* 30.4 (Dec. 1959), pp. 1141–1144. DOI: 10.1214/aoms/1177706098. URL: https://doi.org/10.1214/aoms/1177706098.

[GJ79] Michael R. Garey and David S. Johnson. *Computers and intractability: a guide to the theory of np-completeness*. New York, NY, USA: W. H. Freeman & Co., 1979. ISBN: 0716710447.

[GS18] Christian Glazik and Anand Srivastav. “A new bound for the maker-breaker triangle game”. In: 2018.

[HH86] Nicholas G. Hall and Dorit S. Hochbaum. “A fast approximation algorithm for the multicovering problem”. In: *Discrete Applied Mathematics* 15.1 (1986), pp. 35–40. ISSN: 0166-218X. DOI: https://doi.org/10.1016/0166-218X(86)90016-8. URL: http://www.sciencedirect.com/science/article/pii/0166218X86900168.

[Hoe63] Wassily Hoeffding. “Probability inequalities for sums of bounded random variables”. English. In: *Journal of the American Statistical Association* 58.301 (1963), pp. 13–30. URL: http://www.jstor.org/stable/2282952.

[Jan04] Svante Janson. “Large deviations for sums of partly dependent random variables”. In: *Random Structures & Algorithms* 24.3 (2004), pp. 234–248. DOI: 10.1002/rsa.20008. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20008. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20008.

[JŁR11] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. "Small subgraphs". In: *Random Graphs*. John Wiley & Sons, Ltd, 2011. Chap. 3, pp. 53–80. ISBN: 9781118032718. DOI: 10.1002/9781118032718.ch3. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118032718.ch3. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118032718.ch3.

[JŁR90] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. "An exponential bound for the probability of nonexistence of a specified subgraph in a random graph". In: *Random Graphs '87*. Wiley, 1990, pp. 73–87.

[Kar84] N. Karmarkar. "A new polynomial-time algorithm for linear programming". In: *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*. STOC '84. New York, NY, USA: ACM, 1984, pp. 302–311. ISBN: 0-89791-133-4. DOI: 10.1145/800057.808695. URL: http://doi.acm.org/10.1145/800057.808695.

[Kha79] L. G. Khachiyan. "A polynomial algorithm in linear programming". In: *Doklady Akademii Nauk SSSR* 244 (1979), pp. 1093–1096.

[KLL13] Mirosław Kowaluk, Andrzej Lingas, and Eva-Marta Lundell. "Counting and detecting small subgraphs via equations". In: *SIAM Journal on Discrete Mathematics* 27 (Apr. 2013). DOI: 10.1137/110859798.

[Mat96] Jiří Matoušek. "Derandomization in computational geometry". In: *J. Algorithms* 20.3 (May 1996), pp. 545–580. ISSN: 0196-6774. DOI: 10.1006/jagm.1996.0027. URL: http://dx.doi.org/10.1006/jagm.1996.0027.

[McD89] C. McDiarmid. "On the method of bounded differences". In: *Surveys in Combinatorics*. London Mathematical Society Lecture Note Series 141. Cambridge University Press, Aug. 1989, pp. 148–188.

[McD98] Colin McDiarmid. "Concentration". English. In: *Probabilistic Methods for Algorithmic Discrete Mathematics*. Ed. by Michel Habib, Colin McDiarmid, Jorge Ramirez-Alfonsin, and Bruce Reed. Vol. 16. Algorithms and Combinatorics. Springer Berlin Heidelberg, 1998, pp. 195–248. ISBN: 978-3-642-08426-3.

[MR99] Sanjeev Mahajan and H. Ramesh. "Derandomizing approximation algorithms based on semidefinite programming". In: *SIAM J. Comput.* 28.5 (May 1999), pp. 1641–1663. ISSN: 0097-5397.

[MT10] Robin A. Moser and Gábor Tardos. "A constructive proof of the general lovász local lemma". In: *J. ACM* 57.2 (Feb. 2010), 11:1–11:15. ISSN: 0004-5411. DOI: 10.1145/1667053.1667060. URL: http://doi.acm.org/10.1145/1667053.1667060.

[MU17] Michael Mitzenmacher and Eli Upfal. *Probability and computing: randomization and probabilistic techniques in algorithms and data analysis*. 2nd. USA: Cambridge University Press, 2017. ISBN: 110715488X.

[Nik30] Otton Nikodym. "Sur une généralisation des intégrales de m. j. radon". fre. In: *Fundamenta Mathematicae* 15.1 (1930), pp. 131–179. URL: http://eudml.org/doc/212339.

[PSW97] D. Peleg, G. Schechtman, and A. Wool. "Randomized approximation of bounded multicovering problems". English. In: *Algorithmica* 18.1 (1997), pp. 44–66. ISSN: 0178-4617.

[Rag88] Prabhakar Raghavan. "Probabilistic construction of deterministic algorithms: approximating packing integer programs". In: *J. Comput. Syst. Sci.* 37.2 (1988), pp. 130–143.

[Spe77] Joel Spencer. "Balancing games". In: *Journal of Combinatorial Theory, Series B* 23.1 (1977), pp. 68–74.

[Spe94] Joel Spencer. *Ten lectures on the probabilistic method (cbms-nsf regional conference series in applied mathematics)*. Society for Industrial and Applied Mathematics, 1994. ISBN: 0898713250.

[Sri99] Aravind Srinivasan. "Improved approximation guarantees for packing and covering integer programs". In: *SIAM J. Comput.* 29.2 (Oct. 1999), pp. 648–670. ISSN: 0097-5397.

[SS93] Anand Srivastav and Peter Stangier. "On quadratic lattice approximations". In: *Algorithms and Computation, 4th International Symposium, ISAAC '93, Hong Kong, December 15-17, 1993, Proceedings*. Ed. by Kam-Wing Ng, Prabhakar Raghavan, N. V. Balasubramanian, and Francis Y. L. Chin. Vol. 762. Lecture Notes in Computer Science. Springer, 1993, pp. 176–184. DOI: 10.1007/3-540-57568-5_247. URL: https://doi.org/10.1007/3-540-57568-5%5C_247.

[SS96] Anand Srivastav and Peter Stangier. "Algorithmic chernoff-hoeffding inequalities in integer programming". In: *Random Structures & Algorithms* 8.1 (1996), pp. 27–58. ISSN: 1098-2418.

[SS97] Anand Srivastav and Peter Stangier. "Tight approximations for resource constrained scheduling and bin packing". In: *Discrete Applied Mathematics* 79.1–3 (1997), pp. 223–245. ISSN: 0166-218X. DOI: http://dx.doi.org/10.1016/S0166-218X(97)00045-0. URL: http://www.sciencedirect.com/science/article/pii/S0166218X97000450.