Order Number 9411127

Probabilistic methods in computer science and combinatorics

Narayanan, Babu Ozhur, Ph.D.

New York University, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Probabilistic Methods in Computer Science and Combinatorics

Babu O. Narayanan

August 1993

A dissertation in the Department of Computer Science submitted to the faculty of the Graduate School of Arts and Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy at New York University.

Approved: Ravi B. Boppana

Ravi B. Boppana

Research Advisor

Abstract

Over the last few years, the Probabilistic method has become an important tool in Computer Science and Combinatorics. This thesis deals with three applications of the Probabilistic method. The first problem concerns a model of imperfect randomness: the *slightly-random source* of Santha and Vazirani. In a slightly-random source with bias ϵ, the conditional probability that the next bit output is 1, given complete knowledge of the previous bits output, is between $\frac{1}{2} - \epsilon$ and $\frac{1}{2} + \epsilon$. We show that, for every fixed $\epsilon < \frac{1}{2}$, and for most sets, the probability of hitting that set using a slightly-random source is bounded away from 0. The second problem arises in parallel and distributed computing. A set of n processors is trying to collectively flip a coin, using a protocol that tolerates a large number of faulty processors. We demonstrate the existence of perfect-information protocols that are immune to sets of ϵn faulty processors, for every fixed $\epsilon < \frac{1}{2}$. Finally, we consider a problem in Ramsey theory. Let an adversary color the edges of the Binomial random graph with r colors, the edge probability being $\frac{c}{\sqrt{n}}$, where c is a large enough constant. We show that, almost surely, a constant fraction of the triangles in the graph will be monochromatic.

Acknowledgements

I would like to express my profound gratitude to my advisor, Prof. Ravi B. Boppana. His patience, understanding, and encouragement have been most invaluable during my time as a graduate student. I thank him for everything that he has taught me and done for me. I would also like to express my sincere thanks to Prof. Joel Spencer. My interest in the field of Probabilistic methods was kindled during an independent study with him in the Spring of 1992. The Ramsey theoretic result in the thesis is joint work with him. I have benefited a great deal from the stimulating discussions I have had with him, and I hope some of his enthusiasm has rubbed off on me. My thanks are also due to Prof. Bud Mishra, Prof. Alan Siegel and Prof. Chee Yap for reading my thesis, helpful comments and discussions. I am also grateful to the rest of the Faculty and the Administrative staff, Tamar Arnon and Anina Karmen-Meade in particular, for all their help and guidance.

My thanks are also due to all my teachers at the Indian Institute of Technology, Madras for the fine undergraduate education bestowed on me. In particular, I would like to thank my advisor, Prof. C. Pandu Rangan.

I would like to thank my parents for all their love and constant encouragement. Without their support, this would not have been possible. I also thank the rest of my family for their support. Their sacrifices and encouragement will always be

remembered.

I would also like to thank my colleagues and friends who have made the past five years truly enjoyable. I also thank all my classmates from the IITMCS88 batch for their support, friendship and the great times we have had together.

Contents

Chapter 1

Introduction

The Probabilistic method, pioneered by Paul Erdős, has recently become a significant tool in the areas of Combinatorics and Theoretical Computer Science. The essence of the Probabilistic method can be described as follows: We want to prove the existence of a combinatorial structure that satisfies certain properties. To do so, we create an appropriate probability space and show that a randomly chosen element has the desired properties with strictly positive probability. Here is an example. Given a graph $G = (V, E)$ with n vertices and e edges, we want to show the existence of a bipartite subgraph $H = (V_1, V_2, E')$ with at least $\frac{e}{2}$ edges. For each vertex v in V, independently and randomly put it in V_1 or V_2. By linearity of expectation, the average number of cross edges between V_1 and V_2 is equal to $e/2$. Therefore, the probability that a random partitioning of the vertices results in $\frac{e}{2}$ cross edges is strictly positive. This implies the existence of a bipartite subgraph with at least $\frac{e}{2}$ edges. A full exposition of the probabilistic method, replete with

eye-opening examples can be found in [ASE92,Spe87].

The use of the Probabilistic method in computer science is a relatively recent phenomenon. But it has caught on rapidly, and is now used almost everywhere. Randomized algorithms have always been considered important, especially for their practical value. Some wonderful examples include the algorithm for primality testing by Rabin [Rab76] and the fundamental work on polynomial identities by J. Schwartz [Sch80]. The use of random restrictions is a powerful technique to provide lower bounds for various problems. It is important to point out that the probabilistic method does not provide an explicit construction of the combinatorial structure we are searching for. Sometimes, we can show the existence of an efficient algorithm but finding an explicit construction is very difficult. So, it would be nice to be able to actually construct one in polynomial time. Derandomizing such proofs has received a lot of attention of late; see [ASE92,Spe92] for more on this.

This thesis highlights three applications of the probabilistic method. The first involves an important issue in the field of randomized algorithms; namely, how can we obtain true randomness out of the imperfect random sources available to us? The second is an application of the probabilistic method to show the existence of a good algorithm. A set of processors are trying to achieve a goal. We want to provide a protocol that will work even if many of the processors are faulty. We show that a random protocol, drawn from an appropriate probability space, is very fault-tolerant. Finally, we focus on a problem in pure combinatorics. This is a problem that arises in Ramsey theory. The following is a brief description of the

three results.

1.1 Slightly-Random Sources

Randomized algorithms, which often provide simpler, faster ways to solve problems, usually assume the existence of a source of true randomness. There is a witness set S, a subset of $\{0,1\}^n$, that the algorithm is trying to hit. Unfortunately, physical sources of randomness, such as Zener diodes, are not truly random. Santha and Vazirani [SV86] devised a notion of randomness that approximates the behavior of physical sources. A *slightly-random source* (with *bias* $0 \leq \epsilon \leq \frac{1}{2}$) is a sequence $x = (x_1, x_2, \ldots, x_n)$ of random bits such that the conditional probability that $x_i = 1$, given the outcomes of the first $i-1$ bits, is always between $\frac{1}{2}-\epsilon$ and $\frac{1}{2}+\epsilon$. Intuitively, an adversary, who knows the complete history of previous coin flips, gets to choose the bias of each coin so that the sequence generated will avoid the witness set S. Define the *ϵ-biased probability* $\Pr_\epsilon(S)$ of S by

$$\Pr_\epsilon(S) = \min_x \Pr[x \in S],$$

where x ranges over all slightly-random sources with bias ϵ. For example, $\Pr_0(S) = |S|/2^n$, whereas $\Pr_{1/2}(S) = 0$ (unless $S = \{0,1\}^n$). The ϵ-biased probability measures the minimum odds of hitting S when our adversary, who knows S, gets to decide the bias of each coin flip. Slightly-random sources would be ideal if, for some fixed $\epsilon > 0$, the ϵ-biased probability of a witness set S were always within a constant factor of $|S|/2^n$. Unfortunately, this is not always possible, as observed by Alon and

Rabin [AR89]. Instead, we aim for a bound that applies to *almost every* witness set. Alon and Rabin [AR89] showed that for $\epsilon < \frac{1}{2}(\sqrt{2}-1) \approx .207$, the ϵ-biased probability of almost every witness set is bounded away from 0. They posed the open question of whether the same conclusion is valid for *every* $\epsilon < \frac{1}{2}$. We answer it in the affirmative. See also [BN93a].

Theorem 2.10: For every $\epsilon < \frac{1}{2}$, there is a constant $c_\epsilon > 0$ such that for almost every witness set S, the ϵ-biased probability of hitting S, $\Pr_\epsilon(S)$ is at least c_ϵ. □

The Alon-Rabin result is obtained by a second moment method. For our result, we need to estimate the higher moments of the underlying random process and this makes the proof more interesting.

1.2 Collective Coin Flipping

Closely related to the biased coin problem is the problem of collective coin flipping. In the area of parallel and distributed computing, often a set of processors have to produce the same random bit in order to perform some task. We could simply designate some processor as a leader, and make it generate a random bit that all the other processors would accept. However, the processor could be faulty, so we would like to design a protocol that would work even with many faulty processors. We will permit only broadcast messages and therefore treat the problem as a perfect-information game. Also, we want the protocol to be robust against large coalitions of faulty/dishonest processors. Ben-Or and Linial [BL85,BL89]

formalized the notion of coin flipping protocols as perfect information games. A perfect-information coin flipping protocol for a set of processors N is a rooted tree T. Every interior vertex v is labelled by the name of one processor. Also associated with v is a probability distribution D_v on its children. The leaves are labelled 0 or 1. The protocol starts at the root vertex r with the corresponding processor choosing one of r's children according to the distribution D_r; the protocol proceeds to the chosen vertex and repeats. Finally a leaf is reached; the value at the leaf is said to be the outcome of the protocol. Let $\Pr(T = 1)$ be the probability that the outcome of the protocol is 1. Let $S \subseteq N$ be a coalition of faulty/dishonest processors. For $i \in \{0, 1\}$, let $\Pr_S(T = i)$ be the minimum probability that the outcome of the protocol is i when the coalition plays the optimal strategy. Faulty processors need not use the probability distribution. A protocol T is immune to t cheaters if $\Pr_S(T = 0)$ and $\Pr_S(T = 1)$ are bounded away from 0, as n approaches infinity, for every coalition S of size t or less. One way to solve the coin flipping problem is to first elect a leader and then let the leader flip a coin. Saks [Sak89] noted that no coin-flipping protocol for n players could be immune to $\lceil n/2 \rceil$ cheaters. Alon and Naor [AN93] show the existence of a protocol that is immune to ϵn cheaters, for every $\epsilon < \frac{1}{3}$. They do this by solving the leader election problem and their proof is probabilistic. They asked if there exists a protocol that is immune to ϵn cheaters, for every $\epsilon < \frac{1}{2}$. We improve the analysis of their protocol to answer the question in the affirmative. See also [BN93b].

Theorem 3.1: For every $\epsilon < \frac{1}{2}$, there exists coin flipping protocols and leader

election protocols that are immune to ϵn cheaters. □

1.3 A Ramsey-theoretic Result

The power of the probabilistic method in combinatorics was demonstrated first by Erdős [Erd47] and later by Erdős and Rényi [ER60] when they laid the foundation for the theory of random graphs. Since then, numerous new combinatorial results have been proved using this beautiful technique and elegant proofs provided for classical theorems [ASE92]. Also, the theory of random graphs has developed into a rich field with many exciting problems. Ramsey theory [GRS90] is one fascinating area where the theory of random graphs comes into play quite naturally. A very exciting recent result is a random Ramsey result of Rödl and Ruciński [RR94]. They show: For every $r \geq 2$, there is a $C > 0$ such that if $p > C/\sqrt{n}$ then almost surely every r-coloring of the edges of the random graph $G(n,p)$ results in a monochromatic triangle. Their proof involves a very clever application of the Szemerédi regularity lemma and is full of nice probabilistic lemmas. What about the number of monochromatic triangles? If we are guaranteed one monochromatic triangle, are we guaranteed many? Here, we show that indeed, almost surely, a fraction of the triangles will be monochromatic.

Theorem 4.2: For every $r \geq 2$, there is a $C > 0$ such that if $p > C/\sqrt{n}$, then almost surely, every r-coloring of the edges of the random graph $G(n,p)$ results in $\Omega(n^{3/2})$ monochromatic triangles. □

The organization of the thesis is as follows. Chapter 2 discusses the problems and results on slightly-random sources. Chapter 3 deals with perfect information coin flipping and leader election protocols. In Chapter 4, we discuss the Ramsey-theoretic result concerning monochromatic triangles. Chapter 5 mentions the open problems and possible areas for future research.

Chapter 2

Slightly-Random Sources

2.1 Introduction

Several applications, such as randomized algorithms [Rab76], cryptographic protocols [Blu82,GM82] and stochastic simulation experiments [KG], assume the existence of a source of truly random bits. However, the available physical sources are imperfect. We would therefore like to be able to model these imperfect sources and, hopefully, be able to extract truly random bits from them. One simple model of an imperfect source is a coin with an unknown but fixed bias. Von Neumann [von63] proposed a simple algorithm to generate truly random, independent bits from such a coin. Later Blum [Blu84] considered the problem of generalizing this simple model to a finite state Markov process. He showed that the obvious generalization of Von Neumann's idea does not work in this case. However, and surprisingly, he showed that changing the order in which the bits are output yields completely random, in-

dependent bits. A much more general model of randomness was devised by Santha and Vazirani [SV86]. They call their model a slightly-random source. A slightly random source (with bias $0 \leq \epsilon \leq \frac{1}{2}$) is a sequence $x = (x_1, \ldots, x_n)$ of random bits such that the conditional probability that $x_i = 1$, given the outcome of the first $i-1$ bits, is always between $\frac{1}{2} - \epsilon$ and $\frac{1}{2} + \epsilon$. The intuition is that the bias of next coin flip is decided by an adversary who has complete knowledge of the history of the process. Murry [Mur70] explains that this models the known practical sources of randomness such as the zener diode, in which the frequency of 0's and 1's *drifts* over a period of time. There are other models of weakly random sources such as the 'probability-bounded' source of Chor and Goldreich [CG88] and the δ-source [Zuc91]. For an extensive review of these and other sources, see [BLS87,Vaz86]. We will be concerned with the slightly-random source only.

2.2 Previous Work

Santha and Vazirani [SV86] devised the notion of a slightly-random source. Then, they asked the following natural question: How can we generate better random bits from the bits of a slightly-random source? They noted that there is no algorithm to extract a sequence of absolutely unbiased coin flips from the bits of a slightly-random source. They next turned to the idea of generating strings that look random, called quasi-random strings.

Definition 2.1 *A functional statistical test is a function* $f : \{0,1\}^* \to [0,1]$, *where* $[0,1]$ *denotes the unit interval.*

We are given a source, which for every length n, generates n-length strings $x \in \{0,1\}^n$ with some probability $p(x)$. Let $\mu_n(f) = \frac{1}{2^n}\sum_{|x|=n} f(x)$ be the average value of f on random strings of length n. Let $\mu_n^*(f) = \sum_{|x|=n} p(x)f(x)$, be the average value of f on strings of length n generated by the slightly-random source.

Definition 2.2 *A quasi-random source is a source such that for every* $t > 0$, *for* n *sufficiently large, and for every functional statistical test* f, *we have* $|\mu_n(f) - \mu_n^*(f)| < \frac{1}{n^t}$.

Note that f need not even be computable. Now, Santha and Vazirani [SV86] show that quasi-random sources have very strong properties that enable them to replace truly random sources. We state one of their theorems without proof.

Theorem 2.3 *Let* $G : \{0,1\}^* \to \{0,1\}^*$ *be a pseudo-random number generator that passes all probabilistic polynomial time statistical tests. Then* G *with seeds generated by a quasi-random source also passes all probabilistic polynomial time statistical tests.* □

They also showed that quasi-random sequences can be used to replace the coin flips in any algorithm that generates a desired distribution from a sequence of good coin flips. Since quasi-random sequences seem to be very powerful, it would be nice to be able to extract quasi-random sequences from the bits output by a slightly-random

source. Unfortunately, [SV86] showed that the bits of a single slightly-random source are not sufficient to do this. They, however, provided an algorithm that takes the bits from $\Omega(\log n)$ independent slightly-random sources to generate a quasi-random sequence. Subsequently, Vazirani [Vaz87] showed how to do this with only two independent sources.

Vazirani and Vazirani [VV85] looked at some complexity theoretic issues related to slightly-random sources. Recall that RP is the class of problems solvable in polynomial time, using a truly random source. More formally,

Definition 2.4 *A language L is in RP if there exists a polynomial p, and a deterministic polynomial time algorithm M, which accepts a string r of $p(|x|)$ bits from a source of truly random bits, and such that:*

1) If $x \in L$, $\Pr(M \text{ accepts } x) \geq \frac{1}{2}$.

2) If $x \notin L$, $M(x, r)$ always rejects x.

Associated with a randomized algorithm is the notion of witness sets. The witness set $W(x) = \{r | M(x, r) \text{ accepts } x)\}$. Along similar lines to the definition of RP, Vazirani [Vaz87] defined the class SRP: the class of problems solvable in polynomial time, using a slightly-random source.

Definition 2.5 *A language L is in SRP if:*

For every $\epsilon < \frac{1}{2}$, there exists a deterministic polynomial time algorithm M that can accept bits from a slightly-random source having bias ϵ and such that:

1) If $x \in L$, $\Pr(M \text{ accepts } x) \geq \frac{1}{2}$.

2) If $x \notin L$, M always rejects x.

It is easy to see that $P \subseteq SRP \subseteq RP$. The central question here is the relation between P and RP. One hopes to achieve a better understanding of this problem by studying the class SRP. Vazirani and Vazirani [VV85] showed that indeed, $SRP = RP$.

Theorem 2.6 $SRP = RP$. □

To do this, they showed how to sample polynomially many strings using a slightly-random source, so that at least one of the sampled strings is a witness with probability at least $\frac{1}{2}$, if the input string $x \in L$. They also extended their proof to show that the complexity class BPP can be simulated using a slightly-random source. It is interesting to note that Chor and Goldreich [CG88] showed how to simulate BPP using a weaker source, called a "probability-bounded" source. Zuckerman [Zuc91] showed how to simulate BPP using an even weaker source, called a "δ-source". In all three cases, they transform the bits of the imperfect source to create a polynomial number of n-bit strings, most of which will be witnesses with high probability.

Alon and Rabin [AR89] study the properties of bits produced directly by a single slightly-random source. They note that it is not clear that we can assume the existence of two independent sources, for the bad behavior might be due to the environment in which case the two sources could influence each other. Also, no transformation is permitted; the random bits must be used directly. This restriction has the advantage that the resulting algorithm, when it works, will be more efficient.

We would need only n random bits to produce an n-bit number (unlike the large polynomial number of n-bit strings). No extra space is required either.

2.2.1 The Alon-Rabin result

In their paper, Alon and Rabin [AR89] study the properties of bits produced directly by a single slightly-random source. The framework is as follows. Recall that a slightly-random source (with *bias* $0 \leq \epsilon \leq \frac{1}{2}$) is a sequence $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ of random bits such that the conditional probability that $\mathbf{x}_i = 1$, given the outcomes of the first $i-1$ bits, is always between $\frac{1}{2} - \epsilon$ and $\frac{1}{2} + \epsilon$. Intuitively, we are flipping a bunch of coins, but our adversary, who knows the complete history of previous coin flips, gets to choose the bias of each coin.

In addition, there is a "witness set" S that we are trying to hit, where S is some subset of $\{0,1\}^n$ (the set of all binary sequences of length n). Define the *ϵ-biased probability* $\Pr_\epsilon(S)$ of S by

$$\Pr_\epsilon(S) = \min_x \Pr[x \in S],$$

where x ranges over all slightly-random sources with bias ϵ. For example, $\Pr_0(S) = |S|/2^n$, whereas $\Pr_{1/2}(S) = 0$ (unless $S = \{0,1\}^n$). Intuitively, the ϵ-biased probability measures the minimum odds of hitting S when our adversary, who knows S, gets to choose the source.

Is the ϵ-biased probability of a witness set S always within a constant factor of $|S|/2^n$? Unfortunately, the answer is no, as observed by Alon and Rabin [AR89].

As a counterexample, consider the majority set (for an odd integer n)

$$\mathrm{MAJ} = \{x \in \{0,1\}^n : \sum_{i=1}^{n} x_i > \frac{n}{2}\},$$

whose unbiased probability is $\frac{1}{2}$, but whose ϵ-biased probability is exponentially small for every fixed $\epsilon > 0$. Consequently it is impossible to obtain a strong bound for *every* witness set.

They then determine the "worst possible" set $S \subseteq \{0,1\}^n$ of cardinality k. Define a linear order on the set of all strings of length n as follows. For $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, set $x \leq y$ if and only if $\sum_{i=1}^n x_i < \sum_{i=1}^n y_i$ or $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and $\sum_{i=1}^n x_i 2^i < \sum_{i=1}^n y_i 2^i$. Call a set S compressed if $x \in S$ and $x \leq y$ implies that $y \in S$. It is easy to verify that a compressed set contains all strings with at most j 0's and possibly some strings with exactly $(j+1)$ 0's, where j satisfies the following inequality.

$$\sum_{i=0}^{j} \binom{n}{i} \leq |S| \leq \sum_{i=0}^{j+1} \binom{n}{i}.$$

For a set S, let CS denote the unique compressed set of size $|S|$. The following lemma shows that the best strategy for the adversary to avoid a compressed set is to bias each flip towards 0.

Lemma 2.7 *Let S and j satisfy the above inequality. Let $r = |S| - \sum_{i=1}^n \binom{n}{i}$. Then for every $0 \leq \epsilon \leq \frac{1}{2}$,*

$$\Pr_\epsilon(CS) = \sum_{i=0}^{j} \binom{n}{i} (\tfrac{1}{2}+\epsilon)^i (\tfrac{1}{2}-\epsilon)^{n-i} + r(\tfrac{1}{2}+\epsilon)^{j+1}(\tfrac{1}{2}-\epsilon)^{n-j-1}.$$

□

The following lemma shows that for a number k, the compressed set of k strings is the easiest set for the adversary to avoid.

Lemma 2.8 *For every* $0 \leq \epsilon \leq \frac{1}{2}$ *and for every set* S,

$$\Pr_\epsilon(S) \geq \Pr_\epsilon(CS).$$

Using the above two lemmas, one can show that if $\epsilon \leq \frac{1}{\sqrt{n}}$, then for sets S such that $|S| = c2^n$, the biased probability $\Pr_\epsilon(S)$ is bounded away from 0. However, if ϵ grows asymptotically faster than $\frac{1}{\sqrt{n}}$, then $\Pr_\epsilon(CS)$ goes to 0 as n approaches infinity.

What happens if the set S is a random set of binary strings? Is it true that $\Pr_\epsilon(S)$ is bounded away from 0 for almost every set S? (Here S ranges uniformly over all subsets of $\{0,1\}^n$, and "almost every" means a $1 - o(1)$ fraction as n tends to infinity.). Alon and Rabin [AR89] showed that for $\epsilon < \frac{1}{2}(\sqrt{2} - 1) \approx .207$, the ϵ-biased probability of almost every witness set is bounded away from 0. To do so, they use a recursive calculation of the ϵ-biased probability of a witness set S. Given a vector $x = (x_1, x_2)$ in $\Re^2$, define its *biased mean* by

$$B(x) = \min(p_1 x_1 + p_2 x_2, p_2 x_1 + p_1 x_2),$$

where $p_1 = \frac{1}{2} - \epsilon$ and $p_2 = \frac{1}{2} + \epsilon$. Let T be the complete binary tree of height n, whose leaves naturally correspond to the binary sequences of length n. A leaf is labelled 1 if the corresponding binary sequence is in S; otherwise it is labelled 0. Given an interior node whose two children are labelled x_1 and x_2, define its label

to be the biased mean $B(x_1, x_2)$. Then the label of the root is indeed the ϵ-biased probability of S. Taking expected values of the labelling above leads to

$$\mathrm{E}(B(x)) = \frac{\mathrm{E}(x_1) + \mathrm{E}(x_2)}{2} - \epsilon \mathrm{E}(|x_1 - x_2|).$$

Because x_1 and x_2 come from disjoint subtrees, they are independent, identically-distributed random variables. Alon and Rabin [AR89] used the second-moment method to estimate $\mathrm{E}(|x_1 - x_2|)$. They showed that, for $\epsilon < 0.207$, the variance of x_1 contracts as we go up the tree. This leads to their main result. Here we provide a simpler proof of their contraction lemma. For a random variable z, let $\mathrm{Var}(z)$ denote its variance.

Lemma 2.9 *If x_1 and x_2 are independent, identically distributed random variables, then*

$$\mathrm{Var}(B(x)) \leq 2(\tfrac{1}{2} + \epsilon)^2 \mathrm{Var}(x_1).$$

Proof: Without loss of generality, $E(x_1) = E(x_2) = 0$.

Let $m = \min(x_1, x_2)$ and $M = \max(x_1, x_2)$. By the definition of B, we have

$$B(x) = (\tfrac{1}{2} + \epsilon)m + (\tfrac{1}{2} - \epsilon)M.$$

Squaring both sides leads to

$$\begin{aligned} B(x)^2 &= (\tfrac{1}{2} + \epsilon)^2 m^2 + (\tfrac{1}{2} - \epsilon)^2 M^2 + 2(\tfrac{1}{2} + \epsilon)(\tfrac{1}{2} - \epsilon) mM \\ &\leq (\tfrac{1}{2} + \epsilon)^2 (x_1^2 + x_2^2) + 2(\tfrac{1}{2} + \epsilon)(\tfrac{1}{2} - \epsilon) x_1 x_2. \end{aligned}$$

Taking expected value of both sides gives

$$E(B(x)^2) \leq 2(\tfrac{1}{2} + \epsilon)^2 E(x_1^2).$$

This implies the lemma. □

2.3 Statement of Results

Our first result is an affirmative answer to the Alon-Rabin question. More precisely,

Theorem 2.10 *For every fixed $\epsilon < \frac{1}{2}$, there is a constant $c_\epsilon > 0$ such that for almost every witness set $S \subseteq \{0,1\}^n$, the ϵ-biased probability of hitting S, $\Pr_\epsilon(S) \geq c_\epsilon$.* □

The constant c_ϵ necessarily tends to 0 as ϵ increases to $\frac{1}{2}$. Our proof shows that c_ϵ is at least $p^{O(\log(1/p))}$, where $p = \frac{1}{2} - \epsilon$. We do not know if this lower bound can be improved.

Our work actually applies to a more general situation than the one described above. First, the witness set need not have a uniform distribution; all that matters is that the events "$x \in S$" (for x in $\{0,1\}^n$) be mutually independent. Second, the source need not output merely bits; any finite alphabet will do. (Example: dice.) For simplicity, we omit these generalizations here.

2.4 Significance of Work

The slightly-random source is quite general, and models many real-world sources. It allows many kinds of correlation among the random bits. It permits the adversary to know the witness set, to know the complete history of previous coins, and to be computationally powerful. Our result holds for any ϵ less than $\frac{1}{2}$, which means that the coins may become arbitrarily biased.

Our proof technique is interesting as well. Alon and Rabin [AR89] used a second-moment method to analyze the underlying random process. Unfortunately, this method provably fails for ϵ larger than $\frac{1}{2}(\sqrt{2}-1) \approx .207$. Instead, we analyze a higher moment of the random process. Because higher moments have fewer properties than does the second moment, many technical complications arise. Nevertheless, through the judicious use of classical inequalities, we are able to make this higher-moment method succeed.

2.5 Proof of Main Result

In this section, we prove that the ϵ-biased probability of almost every witness set is bounded away from 0, for every $\epsilon < 1/2$. To do so, we rely on a recursive calculation of the ϵ-biased probability of a witness set S. Given a vector $x = (x_1, x_2)$ in $\Re^2$, define its *biased mean* by

$$B(x) = \min(p_1 x_1 + p_2 x_2, p_2 x_1 + p_1 x_2),$$

where $p_1 = \frac{1}{2} - \epsilon$ and $p_2 = \frac{1}{2} + \epsilon$. Let T be the complete binary tree of height n, whose leaves naturally correspond to the binary sequences of length n. Recall the formulation given in section **2.1**. It is easy to see that the expectation E_n at the root is

$$E_n = E_0 - \epsilon \sum_{i=0}^{n-1} \mathrm{E}(|x_1^{(i)} - x_2^{(i)}|),$$

where $x_1^{(i)}$ and $x_2^{(i)}$ are two independent random variables on level i. (At the leaves, the expectation is $E_0 = \frac{1}{2}$.) Alon and Rabin [AR89] used the second-moment method

to estimate $\mathrm{E}(|x_1 - x_2|)$. They showed that $\mathrm{E}((x_1 - x_2)^2)$ contracts as we go up the tree. This implies an exponentially small bound for $\mathrm{E}(|x_1 - x_2|)$, thereby providing the result. Unfortunately, the contraction holds only when $\epsilon < \frac{1}{2}(\sqrt{2} - 1) \approx 0.207$.

Instead, we shall show that, for every $\epsilon < \frac{1}{2}$ and every sufficiently large number d, the dth moment $\mathrm{E}(|x_1 - x_2|^d)$ contracts.

Given a vector x in $\Re^2$, define its *biased norm* by

$$||x|| = \max(|p_1 x_1 + p_2 x_2|, |p_2 x_1 + p_1 x_2|).$$

Our first lemma says that the biased mean is a Lipschitz continuous function.

Lemma 2.11 *If x and y are two vectors in $\Re^2$, then*

$$|B(x) - B(y)| \leq ||x - y||.$$

Proof: Let $\delta = ||x-y||$. The definition of δ implies the following two inequalities:

$$\begin{aligned} p_1 x_1 + p_2 x_2 &\leq p_1 y_1 + p_2 y_2 + \delta \\ p_2 x_1 + p_1 x_2 &\leq p_2 y_1 + p_1 y_2 + \delta. \end{aligned}$$

By the definition of B, the minimum of the two left-hand expressions is $B(x)$; the minimum of the two right-hand expressions is $B(y)+\delta$. Hence $B(x) \leq B(y)+\delta$. By symmetry, it follows that $B(y) \leq B(x) + \delta$. Combining these last two inequalities leads to the desired result. □

The next lemma relates the biased norm to more classical norms.

Lemma 2.12 *If x is a vector in $\Re^2$ and $d > 1$ is a real number, then*

$$||(x_1, x_2)||^d + ||(-x_1, x_2)||^d \leq c\,(|x_1|^d + |x_2|^d),$$

where $c = (p_1^{d/d-1} + p_2^{d/d-1})^{d-1} + p_2^d$.

Proof: Assume that x_1 and x_2 are non-negative. (The other three cases are similar.) The first term on the left is dealt with using Hölder's inequality [HLP52]. That inequality tells us that

$$|p_1x_1 + p_2x_2|^d \leq (p_1^{d/d-1} + p_2^{d/d-1})^{d-1}(|x_1|^d + |x_2|^d).$$

The same bound holds for $|p_2x_1 + p_1x_2|^d$, and hence for $||(x_1, x_2)||$.

The second term on the left is even easier to deal with. We have

$$\begin{aligned} |-p_1x_1 + p_2x_2|^d &\leq \max(p_1|x_1|, p_2|x_2|)^d \\ &\leq p_1^d|x_1|^d + p_2^d|x_2|^d \\ &\leq p_2^d(|x_1|^d + |x_2|^d). \end{aligned}$$

The same bound holds for $|-p_2x_1 + p_1x_2|^d$, and hence for $||(-x_1, x_2)||$. Adding up the bounds for the two terms on the left completes the proof. □

We obtain the following corollary on the biased mean.

Corollary 2.13 *If x and y are two vectors in $\Re^2$ and $d > 1$ is a real number, then*

$$|B(x_1, x_2) - B(y_1, y_2)|^d + |B(y_1, x_2) - B(x_1, y_2)|^d \leq c\,(|x_1 - y_1|^d + |x_2 - y_2|^d),$$

where $c = (p_1^{d/d-1} + p_2^{d/d-1})^{d-1} + p_2^d$.

Proof: Apply Lemma **2.11** to both terms on the left. Now, apply Lemma **2.12** to prove the Corollary. □

We can now prove the contraction theorem that we have been looking for.

Theorem 2.14 *Let x and y be two random vectors in $\Re^2$ such that x_1, x_2, y_1, and y_2 are mutually independent, and such that x_1 and y_1 are identically distributed. Then for every real number $d > 1$,*

$$\mathrm{E}(|B(x) - B(y)|^d) \leq \frac{c}{2}\mathrm{E}(|x_1 - y_1|^d + |x_2 - y_2|^d),$$

where $c = \left(p_1^{d/d-1} + p_2^{d/d-1}\right)^{d-1} + p_2^d$.

Proof: Take expected values of both sides of the inequality in Corollary **2.13**. The two terms on the left side have the same expected value, because the tuple (x_1, x_2, y_1, y_2) has the same distribution as the tuple (y_1, x_2, x_1, y_2). Dividing by 2 gives the result. □

Theorem **2.14** is a contraction result. The constant c can be made less than 1 by choosing d sufficiently large. That is because as d tends to infinity the value of c tends to $p_1^{p_1} p_2^{p_2} < 1$. In fact, the choice $d = \frac{1}{p_1} \log \frac{1}{p_1}$ makes $c < 1$.

Next, we will attempt to show that the expectation E_n at the root is bounded away from 0. Define $M_i = \mathrm{E}(|x_1^{(i)} - x_2^{(i)}|^d)$.

Theorem 2.15 *For every $0 \leq \epsilon < 1/2$, the final expectation E_n is at least $E_0 - \epsilon M_0^{1/d}/(1 - c^{1/d})$, where* $c = \left(p_1^{d/d-1} + p_2^{d/d-1}\right)^{d-1} + p_2^d$.

Proof: Applying Theorem **2.14** iteratively shows that $M_i \leq c^i M_0$. Jensen's inequality [HLP52] implies that

$$\begin{aligned} \mathrm{E}(|x_1^{(i)} - x_2^{(i)}|) &\leq \mathrm{E}(|x_1^{(i)} - x_2^{(i)}|^d)^{1/d} \\ &= M_i^{1/d} \\ &\leq c^{i/d} M_0^{1/d}. \end{aligned}$$

Summing this bound for every $i \geq 0$ leads to a geometric series whose sum is $M_0^{1/d}/(1 - c^{1/d})$. Plugging this estimate into the displayed formula for E_n given near the beginning of this section finishes the proof. □

We still have not shown that E_n is bounded away from 0 for every $\epsilon < \frac{1}{2}$. That is because as ϵ approaches $\frac{1}{2}$, the value of d approaches infinity, causing $1 - c^{1/d}$ to approach 0, which renders the bound of Theorem **2.15** useless. Instead, we will start with another initial distribution for which it is clear that the final expectation is bounded away from 0, and then we will use a majorizing argument to show the same for our original initial distribution. The same idea was used by Alon and Rabin [AR89].

A random variable x is said to *stochastically dominate* a random variable y if $\Pr[x \geq t]$ is at least $\Pr[y \geq t]$ for every t.

Theorem 2.16 *For every $0 \leq \epsilon < \frac{1}{2}$, the final expectation E_n is bounded away from 0.*

Proof: Given a non-negative integer k to be chosen later, consider the new initial

distribution

$$y = \begin{cases} 0, & \text{with probability } 2^{-2^k}; \\ p_1^k, & \text{otherwise.} \end{cases}$$

The new initial expectation is at least $p_1^k(1 - 2^{-2^k}) \sim p_1^k$. The new initial moment is at most $2^{1-2^k} p_1^{kd}$. Therefore, we can choose k sufficiently large so that, by Theorem 5, the final expectation is bounded away from 0, if our initial distribution were y.

It is easy to verify that the variable $x^{(k)}$ at the kth level of our original process stochastically dominates y. Combined with the monotonicity of our biased mean operator B, we have the result. □

Not only is the expected value of $\Pr_\epsilon(S)$ bounded away from 0, but in fact almost every S will have $\Pr_\epsilon(S)$ bounded away from 0. That is because the moment M_n is exponentially small. Therefore all but an exponentially-small fraction of witness sets S will have ϵ-biased probability bounded away from 0.

Chapter 3

Collective Coin Flipping

3.1 Introduction

In parallel and distributed computing, a set of processors often have to produce the same random bit in order to perform some task. This is easy if we assume that no processor is faulty. To produce this bit, we could simply designate some processor as the leader, and have it generate the random bit. However, this processor could be faulty and then the bit generated need not be random. In fact, we have to contend with the worst case scenario. The faulty processor might behave like an adversary who is trying to prevent the protocol from achieving its goal. Also, there might be many faulty processors. The faulty processors can be considered to act like a coalition of dishonest players/cheaters. We would like to design a coin flipping protocol that would work even with many faulty processors. Closely related to the problem of collective coin flipping is the problem of leader election. Here,

a set of processors wants to elect a leader that is honest/not faulty. Note that a leader election protocol immediately implies a coin flipping protocol. Elect a leader, then have the leader flip a coin. The coin flipping problem has been extensively studied in the framework of Byzantine agreement [FM88,Rab83]. Feldman and Micali [FM88] present a randomized algorithm for leader election that works well even in the presence of a linear number of faulty processors and also runs in expected constant number of rounds. However, in their model, processors are allowed to send private messages to other processors in the network. We will permit only broadcast messages and therefore treat the problem as a perfect-information game. This model was first formalized by Ben-Or and Linial [BL85,BL89]. The following section discusses the major results in the area of perfect information collective coin flipping.

3.2 Previous Work

3.2.1 One round protocols

We will start with one-round coin flipping protocols [BL85,BL89]. Such a protocol is just a Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$. That is, all the processors supply one bit each simultaneously. Let the bit supplied by the ith processor be called x_i. The bit output is $f(x_1, \ldots, x_n)$. Let us assume that the function f is balanced: it is 1 for precisely 2^{n-1} strings. An example of such a protocol is the parity function. Note that as long as there is even one honest player, the bit output is truly random. How

much influence can each processor/variable have in general? (Note that processor i and variable x_i can be used interchangeably). For a set S of variables, assign values to variables outside S at random. Denote by p the probability that given the values assigned to variables not in S, it is possible to set the variables in S so as to make f equal to 0. Then, define the influence of S towards 0: $I_f^0(S) = p - \frac{1}{2}$. Similarly, define $I_f^1(S)$. The total influence of the coalition S is then $I_f(S) = I_f^0(S) + I_f^1(S)$. Ben-Or and Linial [BL89] present a one-round coin flipping protocol where the influence of any particular player is only $O(\frac{\log n}{n})$. It is easy to see that for any Boolean function f, there is always one variable with influence $\Omega(\frac{1}{n})$. Consider all the vertices in the n-cube associated with the function with function value 0. Since the function is balanced, there are at 2^{n-1} such vertices. Now, the edge isoperimetric inequality says that the minimum number of edges in the associated cut is at least 2^{n-1} [Bol86]. Therefore the sum of the influences is $\Omega(1)$. Ben-Or and Linial [BL89] conjectured that for every balanced boolean function f on n variables, there is always a variable with influence at least $\Omega(\frac{\log n}{n})$. Chor and Gereb-Graus [CG87] showed the existence of a variable with influence $\frac{3-o(1)}{n}$. Later, Kahn, Kalai and Linial [KKL88] showed, using techniques from harmonic analysis, that indeed, there is always a variable with influence $\Omega(\frac{\log n}{n})$.

3.2.2 Multi-round protocols

Ben-Or and Linial [BL85,BL89] formalized the notion of coin flipping protocols as perfect information games. A perfect-information coin flipping protocol for a set of

processors N is a rooted tree T. Every interior vertex v is labelled by the name of one processor. Also associated with v is a probability distribution D_v on its children. The leaves are labelled 0 or 1. The protocol starts at the root vertex r with the corresponding processor choosing one of r's children according to the distribution D_r; the protocol proceeds to the chosen vertex and repeats. Finally a leaf is reached; the value at the leaf is said to be the outcome of the protocol. Let $\Pr(T = 1)$ be the probability that the outcome of the protocol is 1. Let $S \subseteq N$ be a coalition of faulty/dishonest processors. For $i \in \{0, 1\}$, let $\Pr_S(T = i)$ be the minimum probability that the outcome of the protocol is i when the coalition plays the optimal strategy. Faulty processors need not use the probability distribution. A protocol T is immune to t cheaters if $\Pr_S(T = 0)$ and $\Pr_S(T = 1)$ are bounded away from 0, as n approaches infinity, for every coalition S of size t or less.

Saks [Sak89] noted that no coin-flipping protocol for n players could be immune to $\lceil n/2 \rceil$ cheaters. Using induction, one can easily show that either there are $\lceil n/2 \rceil$ processors that can completely influence the protocol towards 0 or there are $\lceil n/2 \rceil$ processors that can completely influence the protocol towards 1.

Ben-Or and Linial [BL85] constructed a coin flipping protocol that is immune to $O(n^{\log_3 2})$ cheaters. Saks [Sak89] then came up with a baton passing scheme that is immune to $O(\frac{n}{\log n})$ cheaters. See also Ajtai and Linial [AL89]. The idea is as follows. Initially, the baton is held by some player. That player then passes the baton to one of the remaining players, making the choice randomly. The recipient then hands it over to one of the players that have not yet been selected, again

making the choice randomly. This game proceeds till all the players have been selected. The player holding the baton then flips a coin. If the last player is honest, the coin will be truly random. It is intuitively clear that the best strategy for the cheaters is to pass the baton to one of the honest players. Saks's proof formalizes this intuition to show that this protocol is immune to $O(\frac{n}{\log n})$ cheaters. Ben-Or and Linial [BL89] conjectured that this is the best possible bound. However, Alon and Naor [AN93] showed, using probabilistic methods, the existence of a protocol that is immune to ϵn cheaters, for every $\epsilon < \frac{1}{3}$. In [BN93a], we showed that their protocol works for every $\epsilon < 0.44$. Here we show that their protocol works for every $\epsilon < \frac{1}{2}$. We will elaborate on the protocol of Alon and Naor in the next section. In their paper, they also provide an explicit construction of a protocol that is immune to cn cheaters, for some small, positive constant c. Recently, Cooper and Linial [CL93] have shown the existence of fast protocols that are immune to cn cheaters, for some small constant c. Their protocol requires only polylogarithmically (in n) many rounds. Collective coin flipping and slightly-random sources [VV85,SV86] are closely related. For a delightful survey of the results on these topics, see Ben-Or, Linial, and Saks [BLS87].

3.2.3 The Alon-Naor Result

Alon and Naor [AN93] showed, using probabilistic methods, the existence of a protocol that is immune to ϵn cheaters, for every $\epsilon < \frac{1}{3}$. In [BN93a], we showed that their protocol works for every $\epsilon < 0.44$. Later, we show [BN93b] that it actually

works for every fixed $\epsilon < \frac{1}{2}$. In this section, we will discuss the Alon-Naor protocol. Alon and Naor show the existence of a perfect information coin flipping protocol by showing the existence of a perfect information leader election protocol. In the case of the leader election problem, we have a set of processors who wish to elect a leader. Some of the processors could be faulty and we wish to design a protocol that will often elect a leader that is good. A leader election protocol can be viewed as a rooted tree T. Every vertex v of the tree (including the leaves) is labelled by the name of a processor and has a distribution D_v on its children. The protocol starts at the root and proceeds down some path of the tree. It ends when a leaf is reached; the processor that owns that leaf is chosen the leader. For a coalition $S \subseteq N$, let $\Pr_S(T)$ be the minimum probability that the coalition fails in getting one of its members elected as the leader. A leader election protocol T is immune to t cheaters if $\Pr_S(T) > \delta > 0$ for every set S of size t or less.

Alon and Naor [AN93] showed the existence of a leader election protocol that is immune to ϵn cheaters, where $\epsilon < \frac{1}{3}$. Note that a good leader election protocol leads to a good coin flipping protocol: Once a leader is elected, let the leader flip the coin. This is how Alon and Naor obtain their coin flipping protocol. Their probabilistic construction is as follows: Let T be a complete binary tree of depth $k = \Theta(n)$. Label each vertex of T randomly and independently by a processor. The resulting tree is indeed a leader election protocol for the n processors. The protocol starts at the root node and proceeds down toward a leaf. At each step, the processor associated with the current node decides which of its two children the protocol should proceed

to. Let S be the set of cheaters. If the current processor is not in S, i.e., it is honest, then it chooses one of its two children randomly. On the other hand, if the current processor is a cheater, then it would choose one of its two children according to some optimal strategy. The leader is the processor associated with the leaf node the protocol ends at.

Let S be a fixed subset of N of size ϵn, for $\epsilon < \frac{1}{3}$. Alon and Naor show that there is a constant $\delta > 0$ such that $\Pr[\Pr_S(T) < \delta] < \frac{1}{\binom{n}{\epsilon n}}$. This would imply the existence of a leader election protocol meeting our requirements. They use estimates on the expectation and variance of $\Pr_S(T)$. Essentially, they use a recursive calculation of $\Pr_S(T)$. Then, they show that the variance of this random variable contracts as we go up the tree. Then, using Chebychev's inequality, they obtain their result. Alon and Naor ask the following question: Does there exist a coin flipping protocol that is immune to coalitions of size ϵn, for every $\epsilon < \frac{1}{2}$? A partial answer was provided in [BN93a], where we showed that the random protocol of Alon and Naor works for every $\epsilon < 0.44$.

3.3 Statement of Result

We answer the Alon-Naor question in the affirmative.

Theorem 3.1 *For every $\epsilon < \frac{1}{2}$, there exist coin flipping protocols that are immune to coalitions of size ϵn.* □

3.4 Significance of Work

Our work supplies an asymptotically-optimal bound on the immunity of coin-flipping and leader-election protocols. We show that ϵn faults (for fixed $\epsilon < \frac{1}{2}$) can be tolerated, whereas Saks showed that $\lceil n/2 \rceil$ faults cannot be tolerated. These two bounds leave only a tiny gap, namely the case of $n/2 - o(n)$ faults. We suspect that this number of faults is impossible to tolerate, but we have not been able to prove it.

The perfect-information model is quite strong. The adversary has unlimited computational power, so cryptographic techniques are useless. The faulty processors are allowed to coordinate strategy. No private communication is permitted.

The original motivation for collective coin flipping was its application to the Byzantine agreement problem. Byzantine agreement can be reached quickly if a global coin with not-too-large bias is available (see Rabin [R83]). So our protocol may be of use there.

3.5 Proof of Main Result

Recall the construction of Alon and Naor [AN93]. Let T be a complete binary tree of depth $k = \Theta(n)$. Label each vertex of T randomly and independently by a processor. The resulting tree is indeed a leader election protocol for the n processors. The protocol starts at the root node and proceeds down toward a leaf. At each step, the processor associated with the current node decides which of its two children the

protocol should proceed to. Let S be the set of cheaters. If the current processor is not in S, i.e. it is honest, then it chooses one of its two children randomly. On the other hand, if the current processor is a cheater, then it would choose one of its two children according to some optimal strategy. The leader is the processor associated with the leaf node the protocol ends at.

Alon and Naor obtain their result by analyzing the expectation and variance of the associated random process. We will, on the other hand, analyze higher moments. Interestingly enough, the random process we analyze here is similar to the one analyzed in [AR89,BN93a]. See also Chapter 2. We will use some of the ideas of the proof in [BN93a] although we have not been able to establish any formal relationship between the two random processes.

Consider a fixed set of cheaters S of size ϵn. Let y be an internal vertex of the tree T, let u and v be its children, and let T_y be the subtree rooted at node y. It is easy to see that

$$\Pr_S(T_y) = \begin{cases} \frac{\Pr_S(T_u)+\Pr_S(T_v)}{2} & \text{if } y \notin S; \\ \min(\Pr_S(T_u), \Pr_S(T_v)) & \text{if } y \in S. \end{cases}$$

Since the tree T is random, $y \in S$ itself is a Bernoulli random variable with expectation ϵ. Our aim is to understand the variable $\Pr_S(T)$ associated with the root. Note that all the variables associated with the internal nodes at the same level are identically distributed. Let w_l and w_l' be two independent copies of the random variable at height l. At the leaf level w_0 is a Bernoulli random variable with mean $1 - \epsilon$. Now, we would like to show that the expectation $E(|w_i - w_i'|^d)$, for

sufficiently large d, contracts as we go up the tree. This would enable us to say, as we will see later, that the expectation at the root $E(w_k)$ is bounded away from 0 and also that w_k is bounded away from 0 with very high probability. To do this, we need the following lemmas.

Let $x = (x_1, x_2, x_3, x_4)$ be a fixed vector in $\Re^4$. Define $\Delta(x)$ as follows:

$$\Delta(x) = \Delta_1(x) + \Delta_2(x) + \Delta_3(x) + \Delta_4(x),$$

where

$$\begin{aligned}
\Delta_1(x) &= (1-\epsilon)^2 \left| \frac{x_1 + x_2}{2} - \frac{x_3 + x_4}{2} \right|^d \\
\Delta_2(x) &= \epsilon(1-\epsilon) \left| \frac{x_1 + x_2}{2} - \min(x_3, x_4) \right|^d \\
\Delta_3(x) &= \epsilon(1-\epsilon) \left| \min(x_1, x_2) - \frac{x_3 + x_4}{2} \right|^d \\
\Delta_4(x) &= \epsilon^2 \left| \min(x_1, x_2) - \min(x_3, x_4) \right|^d .
\end{aligned}$$

The motivation for Δ will become clear in the proof of Lemma 2. The following lemma, an upper bound on Δ, will be the key tool in proving our contraction lemma.

Lemma 3.2 *If x is a vector in $\Re^4$, then*

$$\Delta(x_1, x_2, x_3, x_4) + \Delta(x_3, x_2, x_1, x_4) + \Delta(x_1, x_3, x_2, x_4) \leq \tfrac{1}{2} c \sum_{i<j} |x_i - x_j|^d,$$

where $c = \max(2\epsilon, \frac{1+\epsilon}{2}) + 6(3/4)^d$.

Proof: Since both sides of the inequality are symmetric functions of x, we may assume that $x_1 \leq x_2 \leq x_3 \leq x_4$. The proof now divides into two cases, depending on the value of x_2.

Case 1: $x_2 \geq \frac{x_1+x_3}{2}$. We will obtain an upper bound on each of the three Δ terms on the left hand side of the inequality, and then sum them up to obtain the desired result. First, let us bound $\Delta(x_1, x_2, x_3, x_4)$. We now bound each of its four terms.

$$\begin{aligned}
\Delta_1 &= (1-\epsilon)^2 \left|\frac{x_3+x_4}{2} - \frac{x_1+x_2}{2}\right|^d \leq (1-\epsilon)^2 \left|\frac{x_3+x_4}{2} - \frac{x_1+(x_1+x_3)/2}{2}\right|^d \\
&= (1-\epsilon)^2 \left|\frac{x_3+2x_4-3x_1}{4}\right|^d \leq (1-\epsilon)^2 (3/4)^d |x_4-x_1|^d \\
&\leq (3/4)^d |x_4-x_1|^d.
\end{aligned}$$

$$\begin{aligned}
\Delta_2 &= \epsilon(1-\epsilon) \left|x_3 - \frac{x_1+x_2}{2}\right|^d \leq \epsilon(1-\epsilon) \left|x_3 - \frac{x_1+(x_1+x_3)/2}{2}\right|^d \\
&= \epsilon(1-\epsilon)(3/4)^d |x_3-x_1|^d \leq (3/4)^d |x_3-x_1|^d.
\end{aligned}$$

$$\begin{aligned}
\Delta_3 &= \epsilon(1-\epsilon) \left|\frac{x_3+x_4}{2} - x_1\right|^d = \epsilon(1-\epsilon) \left|\frac{x_3-x_1}{2} + \frac{x_4-x_1}{2}\right|^d \\
&\leq \tfrac{1}{2}\epsilon(1-\epsilon)[|x_3-x_1|^d + |x_4-x_1|^d].
\end{aligned}$$

$$\Delta_4 = \epsilon^2 |x_3-x_1|^d.$$

Adding up these four inequalities, we have

$$\begin{aligned}
\Delta(x_1, x_2, x_3, x_4) \leq\ & \tfrac{1}{2}\epsilon(1+\epsilon)|x_3-x_1|^d + \tfrac{1}{2}\epsilon(1-\epsilon)|x_4-x_1|^d + \\
& (3/4)^d |x_3-x_1|^d + (3/4)^d |x_4-x_1|^d.
\end{aligned}$$

Now, let us bound the middle term $\Delta(x_3, x_2, x_1, x_4)$. Again, we bound each of

its four terms:

$$\begin{aligned}\Delta_1 &= (1-\epsilon)^2\left|\frac{x_1+x_4}{2}-\frac{x_3+x_2}{2}\right|^d \le (1-\epsilon)^2\left|\frac{x_4-x_1}{2}\right|^d \\ &= (1-\epsilon)^2(1/2)^d|x_4-x_1|^d \le (1-\epsilon)(3/4)^d|x_4-x_1|^d.\end{aligned}$$

$$\begin{aligned}\Delta_2 &= \epsilon(1-\epsilon)\left|\frac{x_3+x_2}{2}-x_1\right|^d \\ &\le \tfrac{1}{2}\epsilon(1-\epsilon)[|x_2-x_1|^d+|x_3-x_1|^d].\end{aligned}$$

$$\begin{aligned}\Delta_3 &= \epsilon(1-\epsilon)\left|x_2-\frac{x_1+x_4}{2}\right|^d \le \epsilon(1-\epsilon)\left|\frac{x_4-x_1}{2}\right|^d \\ &= \epsilon(1-\epsilon)(1/2)^d|x_4-x_1|^d \le \epsilon(3/4)^d|x_4-x_1|^d.\end{aligned}$$

$$\Delta_4 = \epsilon^2|x_2-x_1|^d.$$

Summing the above four terms, we get

$$\Delta(x_3,x_2,x_1,x_4) \le \tfrac{1}{2}\epsilon(1+\epsilon)|x_2-x_1|^d+\tfrac{1}{2}\epsilon(1-\epsilon)|x_3-x_1|^d+(3/4)^d|x_4-x_1|^d.$$

Finally, we obtain an upper bound for $\Delta(x_1,x_3,x_2,x_4)$.

$$\begin{aligned}\Delta_1 &= (1-\epsilon)^2\left|\frac{x_1+x_3}{2}-\frac{x_2+x_4}{2}\right|^d \le (1-\epsilon)^2\left|\frac{x_4-x_1}{2}\right|^d \\ &= (1-\epsilon)^2(1/2)^d|x_4-x_1|^d \le (3/4)^d|x_4-x_1|^d.\end{aligned}$$

$$\Delta_2 = \epsilon(1-\epsilon)\left|\frac{x_1+x_3}{2}-x_2\right|^d \le \epsilon(1-\epsilon)\left|\frac{x_3-x_1}{2}\right|^d$$

$$= \epsilon(1-\epsilon)(1/2)^d|x_3-x_1|^d \leq (3/4)^d|x_3-x_1|^d.$$

$$\Delta_3 = \epsilon(1-\epsilon)\left|\frac{x_4+x_2}{2}-x_1\right|^d \leq \tfrac{1}{2}\epsilon(1-\epsilon)[|x_2-x_1|^d+|x_4-x_1|^d]$$

$$\Delta_4 = \epsilon^2|x_2-x_1|^d \leq \epsilon^2|x_4-x_1|^d.$$

Summing up the four terms, we get

$$\Delta(x_1,x_3,x_2,x_4) \leq \tfrac{1}{2}\epsilon(1-\epsilon)|x_2-x_1|^d + \tfrac{1}{2}\epsilon(1+\epsilon)|x_4-x_1|^d + (3/4)^d|x_3-x_1|^d + (3/4)^d|x_4-x_1|^d.$$

Now that we have bounded all three Δ terms, we can add up the bounds. We obtain

$$\Delta(x_1,x_2,x_3,x_4)+\Delta(x_3,x_2,x_1,x_4)+\Delta(x_1,x_3,x_2,x_4)$$

$$\leq \epsilon|x_2-x_1|^d+\epsilon|x_3-x_1|^d+\epsilon|x_4-x_1|^d+2(3/4)^d|x_3-x_1|^d+3(3/4)^d|x_4-x_1|^d$$

$$\leq [\epsilon+3(3/4)^d]\sum_{i<j}|x_i-x_j|^d \leq \tfrac{1}{2}c\sum_{i<j}|x_i-x_j|^d,$$

by the definition of c. This settles Case 1.

Case 2: $x_2 \leq \frac{x_1+x_3}{2}$. Again, we bound each of the three Δ terms, and then add up the three bounds. First, let us bound the four terms of $\Delta(x_1,x_2,x_3,x_4)$.

$$\Delta_1 = (1-\epsilon)^2\left|\frac{x_1+x_2}{2}-\frac{x_3+x_4}{2}\right|^d$$

$$= (1-\epsilon)^2 \left| \frac{x_4 - x_1}{4} + \frac{x_4 - x_2}{4} + \frac{x_3 - x_1}{4} + \frac{x_3 - x_2}{2} \right|^d$$

$$\leq \tfrac{1}{4}(1-\epsilon)^2 [|x_4 - x_1|^d + |x_4 - x_2|^d + |x_3 - x_1|^d + |x_3 - x_2|^d].$$

$$\Delta_2 = \epsilon(1-\epsilon) \left| \frac{x_1 + x_2}{2} - x_3 \right|^d \leq \tfrac{1}{2}\epsilon(1-\epsilon)[|x_3 - x_1|^d + |x_3 - x_2|^d].$$

$$\Delta_3 = \epsilon(1-\epsilon) \left| x_1 - \frac{x_3 + x_4}{2} \right|^d \leq \tfrac{1}{2}\epsilon(1-\epsilon)[|x_3 - x_1|^d + |x_4 - x_1|^d].$$

$$\Delta_4 = \epsilon^2 |x_3 - x_1|^d .$$

Adding up the four terms, we get

$$\begin{aligned} \Delta(x_1, x_2, x_3, x_4) &\leq \tfrac{1}{4}(1+\epsilon)^2 |x_3 - x_1|^d + \tfrac{1}{4}(1-\epsilon^2) |x_4 - x_1|^d \\ &\quad + \tfrac{1}{4}(1-\epsilon^2) |x_3 - x_2|^d + \tfrac{1}{4}(1-\epsilon)^2 |x_4 - x_2|^d \\ &\leq \frac{1+\epsilon}{4} |x_3 - x_1|^d + \frac{1+\epsilon}{4} |x_4 - x_1|^d \\ &\quad + \tfrac{1}{4}(1-\epsilon^2) |x_3 - x_2|^d + \tfrac{1}{4}(1-\epsilon)^2 |x_4 - x_2|^d \\ &\leq \frac{1+\epsilon}{4} \sum_{i<j} |x_i - x_j|^d . \end{aligned}$$

Secondly, we bound the four terms of the term $\Delta(x_3, x_2, x_1, x_4)$.

$$\Delta_1 = (1-\epsilon)^2 \left| \frac{x_3 + x_2}{2} - \frac{x_1 + x_4}{2} \right|^d \leq (1-\epsilon)^2 (1/2)^d |x_4 - x_1|^d$$

$$\leq (1-\epsilon)(3/4)^d |x_4 - x_1|^d .$$

$$\Delta_2 = \epsilon(1-\epsilon) \left| \frac{x_3 + x_2}{2} - x_1 \right|^d = \epsilon(1-\epsilon) \left| \frac{x_3 + (x_1 + x_3)/2}{2} - x_1 \right|^d$$

$$= \epsilon(1-\epsilon)(3/4)^d |x_3 - x_1|^d \leq (1-\epsilon)(3/4)^d |x_3 - x_1|^d .$$

$$\begin{aligned}
\Delta_3 &= \epsilon(1-\epsilon)\left|x_2 - \frac{x_1+x_4}{2}\right|^d \leq \epsilon(1-\epsilon)(1/2)^d |x_4 - x_1|^d \\
&\leq \epsilon(3/4)^d |x_4 - x_1|^d .
\end{aligned}$$

$$\begin{aligned}
\Delta_4 &= \epsilon^2 |x_2 - x_1|^d \leq \epsilon^2 \left|\frac{x_1+x_3}{2} - x_1\right|^d \\
&\leq \epsilon^2 (1/2)^d |x_3 - x_1|^d \leq \epsilon(3/4)^d |x_3 - x_1|^d .
\end{aligned}$$

Adding these four inequalities yields

$$\Delta(x_3, x_2, x_1, x_4) \leq (3/4)^d |x_3 - x_1|^d + (3/4)^d |x_4 - x_1|^d .$$

Finally, we bound the four terms of $\Delta(x_1, x_3, x_2, x_4)$.

$$\begin{aligned}
\Delta_1 &= (1-\epsilon)^2 \left|\frac{x_1+x_3}{2} - \frac{x_2-x_4}{2}\right|^d \leq (1-\epsilon)^2 (1/2)^d |x_4 - x_1|^d \\
&\leq (1-\epsilon)(3/4)^d |x_4 - x_1|^d .
\end{aligned}$$

$$\begin{aligned}
\Delta_2 &= \epsilon(1-\epsilon)\left|\frac{x_1+x_3}{2} - x_2\right|^d \leq \epsilon(1-\epsilon)(1/2)^d |x_3 - x_1|^d \\
&\leq (1-\epsilon)(3/4)^d |x_3 - x_1|^d .
\end{aligned}$$

$$\begin{aligned}
\Delta_3 &= \epsilon(1-\epsilon)\left|x_1 - \frac{x_2+x_4}{2}\right|^d \leq \epsilon(1-\epsilon)\left|x_1 - \frac{(x_1+x_3)/2 + x_4}{2}\right|^d \\
&= \epsilon(1-\epsilon)\left|\frac{2x_4 - x_3 - 3x_1}{4}\right|^d \leq \epsilon(1-\epsilon)(3/4)^d |x_4 - x_1|^d \\
&\leq \epsilon(3/4)^d |x_4 - x_1|^d .
\end{aligned}$$

$$\begin{aligned}\Delta_4 &= \epsilon^2 |x_2 - x_1|^d \leq \epsilon^2 \left| \frac{x_1 + x_3}{2} - x_1 \right|^d \\ &= \epsilon^2 (1/2)^d |x_3 - x_1|^d \leq \epsilon (3/4)^d |x_3 - x_1|^d .\end{aligned}$$

Adding the four terms, we get

$$\Delta(x_1, x_3, x_2, x_4) \leq (3/4)^d |x_3 - x_1|^d + (3/4)^d |x_4 - x_1|^d$$

Again, now that we have bounded all the three Δ terms, we add up all three bounds to get

$$\Delta(x_1, x_2, x_3, x_4) + \Delta(x_3, x_2, x_1, x_4) + \Delta(x_1, x_3, x_2, x_4)$$

$$\begin{aligned}&\leq \tfrac{1}{4}(1+\epsilon) \sum_{i<j} |x_i - x_j|^d + 2(3/4)^d |x_3 - x_1|^d + 2(3/4)^d |x_4 - x_1|^d \\ &\leq [\frac{1+\epsilon}{4} + 2(3/4)^d] \sum_{i<j} |x_i - x_j|^d \leq \tfrac{1}{2} c \sum_{i<j} |x_i - x_j|^d .\end{aligned}$$

This settles Case 2 as well as the Lemma. □

Using Lemma 1 above, we can prove the contraction result that we desire. Let $x = (x_1, x_2, x_3, x_4)$ be a random vector in $\Re^4$ such that x_1, x_2, x_3, and x_4 are all mutually independent, and x_1, x_2, and x_3 are identically distributed. Consider independent random variables z_1 and z_2 with the following distributions:

$$z_1 = \begin{cases} \frac{x_1 + x_2}{2} & \text{with probability } 1 - \epsilon; \\ \min(x_1, x_2) & \text{otherwise.} \end{cases}$$

$$z_2 = \begin{cases} \frac{x_3+x_4}{2} & \text{with probability } 1-\epsilon; \\ \min(x_3, x_4) & \text{otherwise.} \end{cases}$$

Lemma 3.3 (Contraction Lemma) *If x_1, x_2, x_3, x_4, z_1, and z_2 are random variables defined as above, then*

$$E(|z_1 - z_2|^d) \leq \frac{c}{6} \sum_{i<j} E(|x_i - x_j|^d),$$

where $c = \max(2\epsilon, \frac{1+\epsilon}{2}) + 6(3/4)^d$.

Proof: For fixed $x = (x_1, x_2, x_3, x_4)$, we have $E(|z_1 - z_2|^d) = \Delta(x_1, x_2, x_3, x_4)$. Therefore, $E(|z_1 - z_2|^d) = E[\Delta(x_1, x_2, x_3, x_4)]$. By the hypothesis, the three tuples (x_1, x_2, x_3, x_4), (x_3, x_2, x_1, x_4), and (x_1, x_3, x_2, x_4) all have the same distribution. Take expected values of the inequality in Lemma 3.2. All three expected values on the left are equal. The lemma is now immediate. □

To make this a contraction result, we need the constant c to be less than 1. Given $\epsilon < \frac{1}{2}$, clearly, we can choose d sufficiently large so that $c < 1$. In fact, $d = 8 \log \frac{6}{1-2\epsilon}$ suffices.

Let us go back to our random tree. From our contraction lemma, it is immediate that $E(|w_l - w'_l|^d) \leq cE(|w_{l-1} - w'_{l-1}|^d)$, where c is the same constant as in the contraction lemma. Therefore, $E(|w_k - w'_k|^d) \leq c^k E(|w_0 - w'_0|^d) \leq c^k 2\epsilon(1-\epsilon)$.

Next, we would like to show that the expectation at the root $E(w_k)$ is bounded away from 0.

$$E(w_k) = (1-\epsilon)E(\frac{w_{k-1} + w'_{k-1}}{2}) + \epsilon E(\min(w_{k-1}, w'_{k-1}))$$

$$
\begin{aligned}
&= (1-\epsilon)E(w_{k-1}) + \epsilon E(\frac{w_{k-1}+w'_{k-1}}{2} - \frac{|w_{k-1}-w'_{k-1}|}{2}) \\
&= E(w_{k-1}) - \tfrac{1}{2}\epsilon E(|w_{k-1}-w'_{k-1}|).
\end{aligned}
$$

Iterating, we have

$$
\begin{aligned}
E(w_k) &= E(w_0) - \tfrac{1}{2}\epsilon \sum_{i=0}^{k-1} E(|w_i - w'_i|) \\
&\geq 1 - \epsilon - \tfrac{1}{2}\epsilon \sum_{i=0}^{k-1} \sqrt[d]{E(|w_i - w'_i|^d)} \\
&\geq 1 - \epsilon - \tfrac{1}{2}\epsilon \frac{(2\epsilon(1-\epsilon))^{1/d}}{1-c^{1/d}}.
\end{aligned}
$$

Here, we applied Jensen's inequality [HLP52] to bound $E(|w_i - w'_i|)$ by $\sqrt[d]{E|w_i - w'_i|^d}$, and then summed up the geometric series with ratio $c^{1/d} < 1$. This itself does not prove the theorem for all $\epsilon < \frac{1}{2}$ because $c^{1/d}$ might be very close to 1 causing the last term to become much bigger than $1-\epsilon$. However, we can the majorizing argument in Chapter **2**, Lemma **11** (see also [BN93a]) to complete the proof for all fixed $\epsilon < \frac{1}{2}$.

Chapter 4

A Ramsey-theoretic result

4.1 Introduction

The power of the probabilistic method in combinatorics was demonstrated first by Erdős [Erd47] and later by Erdös and Rényi [ER60] when they laid the foundation for the theory of random graphs. Since then, numerous new combinatorial results have been proved using this beautiful technique and elegant proofs provided for classical theorems [ASE92]. Also, the theory of random graphs has developed into a rich field with many exciting problems. Ramsey theory [GRS90] is one fascinating area where the theory of random graphs has seen a lot of application. In fact, Ramsey theory had a lot to do with the early development of the theory of random graphs. See [Erd47,Erd61] for more on this.

4.2 Brief History

Prove that in any collection of six people, either three of them mutually know each other or three of them mutually do not know each other. This popular puzzle is probably the most famous result that can be considered Ramsey-theoretic. Burkill and Mirsky [BM73,GRS90] state: *There are numerous theorems in mathematics which assert, crudely speaking, that every system of a certain class possesses a large subsystem with a higher degree of organization than the original system.* The aforementioned puzzle is a simple example of this phenomenon, of which Ramsey theory is an important part. Results in Ramsey theory usually state that if we partition a large system into many parts, then at least one of the parts contains a subsystem of a given size [Bol85]. For example, Van der Waerden's theorem says that if the positive integers are partitioned into finitely many classes, then at least one class contains arithmetic progressions of arbitrary length. Similarly, Schur's theorem states that if the positive integers are partitioned into finitely many classes, at least one class contains x, y, z such that $x + y = z$. A generalization of the earlier puzzle problem would say that if a graph contains sufficiently many vertices, then it has either a complete subgraph of k vertices or an independent set of k vertices. This is a special case of Ramsey's theorem. The partitioning of a set can be thought of as a coloring of the elements of the set. The Ramsey number $R(k, l)$ is defined to be the smallest integer n such that if the edges of the complete graph on n vertices are colored red or blue, then there will be either a complete subgraph of k vertices

whose edges are all red, or a complete subgraph of l vertices whose edges are all blue. Ramsey [Ram30] showed that $R(k,l)$ is finite for any two integers k and l. One of the first applications of the probabilistic method in combinatorics came when Erdős [Erd47] showed that $R(k,k) > 2^{k/2}$ for $k \geq 3$. Simply color each edge red or blue, randomly and independently. If $n \leq 2^{k/2}$, the expected number of monochromatic complete subgraphs of size k is seen to be less than 1. This shows the existence of a coloring that avoids monochromatic cliques of size k. A small improvement can be obtained by using the Lovász Local Lemma [ASE92]. Ramsey numbers are notoriously difficult to determine. Even $R(5,5)$ is not known precisely. For more on this, see [Spe87].

4.3 The Rödl-Ruciński Result

Let us now consider the problem of obtaining Ramsey type results for random graphs. In the late 1960s, Erdős asked if there was a graph G whose maximum clique size was 3 and such that any r coloring of the edges of the graph would result in a monochromatic triangle. This was answered in the affirmative first by Folkman [Fol70] for two colors and later by Nesetril and Rödl [NR84]. Spencer [Spe88] used the probabilistic method to show the existence of a 'small' graph satisfying this property for $r = 2$, the two coloring problem. (Typically, the numbers involved here are VERY big and the 'small' graph has roughly three hundred million vertices.) Basically, he showed that a random graph on n vertices, for n sufficiently large, with

edge probability $p = \frac{c}{\sqrt{n}}$ would be good enough. See also, Frankl and Rödl [FR86].

Formally, let $G(n,p)$ denote the Bernoulli random graph obtained by deleting each edge of the complete graph on n vertices K_n, independently with probability $1-p$. $F \rightarrow (G)^2_r$ means that any r-coloring of the edges of F results in at least 1 monochromatic copy of G in F. Rödl and Ruciński [RR94] proved the following:

Theorem 4.1 *For all $r \geq 2$, there exists $C > 0$ such that if $p > Cn^{-1/2}$ then almost surely every r-coloring of the edges of $G(n,p)$ results in a monochromatic triangle.*

Note that the property $G \rightarrow (K_3)^2_r$ is an increasing property of G— adding an edge can only make the graph satisfy the property, not destroy it. Therefore, it follows from [BT86] that there is a threshold function $p^*(n)$, for each r, so that if the edge probability p is such that $p = o(p^*)$, then almost surely, $G(n,p) \not\rightarrow (K_3)^2_r$, while if $p^* = o(p)$, then almost surely $G(n,p) \rightarrow (K_3)^2_r$. Luczak, Ruciński and Voigt [LRV92] showed that if $p < cn^{-1/2}$, for c sufficiently small, $\Pr(G(n,p) \rightarrow (K_3)^2_r) \rightarrow 0$, as $n \rightarrow \infty$. Combining this with the above theorem, Rödl and Rucinski noted that $p = n^{-1/2}$ is a threshold for the property $G(n,p) \rightarrow (K_3)^2_r$, independent of r.

4.4 Main Result

4.4.1 Statement of Main Result

In this thesis, we extend the result of Rödl and Ruciński [RR94] in the following sense. Let $MT_r(G)$ be the number of monochromatic triangles in any r-coloring of G. If $p > C/\sqrt{n}$, then from [RR94], almost surely there is a monochromatic triangle

in all r-colorings of the random graph. Is it true that, almost surely, a fraction of the triangles will be monochromatic? Our main theorem shows that this is indeed the case.

Theorem 4.2 *For all $r \geq 1$, there exists $c > 0$ and $\alpha > 0$ such that the following holds: Let $p > \frac{c}{\sqrt{n}}$. Then, almost surely,*

$$MT_r(G(n,p)) \geq \alpha \frac{c^3}{6} n^{3/2}.$$

□

We prove our theorem by modifying the proof of [RR94]. First, we need some definitions. The next subsection contains the important definitions and a brief description of the main tools used.

4.4.2 Definitions and Tools

Let G be a bipartite graph on vertex sets A, B. For $X \subseteq A$ and $Y \subseteq B$, let $e(X,Y)$ denote the number of edges in the induced subgraph on $X \times Y$. Then the density of G on $X \times Y$ is $\rho(X,Y) = \frac{e(X,Y)}{|X| \cdot |Y|}$.

Definition 4.3 *G is said to be ϵ-regular if for all X, Y such that $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$,*

$$|\rho(X,Y) - \rho(A,B)| < \epsilon.$$

G is said to be $(\geq d, \epsilon)$-regular if it is ϵ-regular with density $\rho \geq d$.

Let H be a graph with vertex set $A_1 \cup \ldots \cup A_f$, the A_i being disjoint.

Definition 4.4 *H is said to be $(\geq d, \epsilon, f)$-regular (with respect to this fixed partition of the vertices) if for <u>all</u> i, j, $1 \leq i < j \leq f$, H is $(\geq d, \epsilon)$-regular on $A_i \times A_j$.*

For properties of $(\geq d, \epsilon)$-regular graphs, see [RR94].

Now, we state a version of the celebrated Szemerédi Regularity Lemma that is suitable for our purposes. This version is implied by the original proof of Szemerédi [Sze78] and can be explicitly found in [EHS+93].

Lemma 4.5 (The Szemerédi Regularity Lemma) *Let $A_1, \ldots, A_m$ be disjoint sets, each of n vertices. Let $H_1, \ldots, H_w$ be graphs on the union of $A_1, \ldots, A_m$. Then, for all $\epsilon > 0$, there exists a $C < C_0(m, \epsilon)$ such that there exists a partitioning of the sets A_i into $E_i, A_i^1, \ldots, A_i^C$ such that $|E_i| \leq \epsilon n$, $|A_i^s| = |A_i^t|$, for all $s, t \geq 1$ and for all x, H_x is ϵ-regular on $A_i^s \times A_j^t$ on all but $\epsilon m^2 C^2$ pairs.*

Definition 4.6 *$n \rightarrow (l_1, \ldots, l_r)$ if for every r-coloring of the complete graph on n vertices, there exists $i, 1 \leq i \leq r$, and a complete subgraph T of l_i vertices with all edges colored i.*

We now state a version of Ramsey's theorem [GRS90] suitable for our purposes.

Lemma 4.7 (Ramsey's Theorem) *For all $f_1, \ldots, f_r$, there exists n_0 so that for $n \geq n_0$,*

$$n \rightarrow (f_1, \ldots, f_r).$$

Corollary 4.8 *For all f, there exists $n_0(f, r)$ so that for $n \geq n_0$,*

$$n \rightarrow (f_1, \ldots, f_{r+1}, r+2)$$

where $f_1 = \ldots = f_{r+1} = f$.

Definition 4.9 *Let* $R_r(f) = n_0(f, r)$.

We will use this definition later to decide the size of the graph that we start with.

Now, we mention two large deviation theorems that will be repeatedly used through out the proof. The first one, commonly known as the Chernoff bound, deals with finite sums of mutually independent random variables. The second is Janson's inequality for the sums of random variables with limited dependence.

Let $Y_1, \ldots, Y_n$ be mutually independent zero-one random variables. Let $Y = \sum_{i=1}^{n} Y_i$. Let $\mu = E(Y)$ be the expected value of the sum.

Lemma 4.10 (Chernoff's bound, see [ASE92]) *For all* $\epsilon > 0$, *there exists a positive constant* c_ϵ *such that*

$$\Pr(|Y - \mu| > \epsilon\mu) < 2\exp[-c_\epsilon \mu].$$

Let Ω be an arbitrary finite set and let R be a random subset of Ω given by $\Pr(r \in R) = p_r$, these events mutually independent over $r \in \Omega$. Let I be a finite index set and for $A_i, i \in I$, given subsets of Ω, let $B_i, i \in I$, be the event that $A_i \subseteq R$ and $Y_i, i \in I$, be the associated indicator random variables. Let $Y = \sum_{i \in I} Y_i$. Furthermore, let $\mu = E(Y)$ be the expected value of Y. Can we get a large deviation theorem for Y? Janson [Jan90] provides an inequality to bound the lower tail. The nice part about this inequality is that it is only in terms of the expectation of Y and some version of a second moment of Y. This makes it easy to use. Define $\Delta = \sum_{i \sim j} \Pr(B_i \wedge B_i)$. This sum is over all ordered pairs (i, j) such that $A_i \cap A_j \neq \phi$.

Lemma 4.11 (Janson's Inequality) *With notations as above, if* $0 \leq \epsilon \leq 1$, *then*

$$\Pr(Y \leq (1-\epsilon)\mu) \leq \exp[-\tfrac{1}{2}\frac{(\epsilon\mu)^2}{\mu+\Delta}].$$

For a graph G, let G_p denote the random subgraph of G obtained by independently selecting each edge with probability p.

4.4.3 Proof of Main Result

We prove our theorem by modifying the proof of Rödl and Ruciński [RR94]. The proof is by induction on r, the number of colors. In order to keep the induction going, we (and [RR94]) are forced to prove a stronger theorem. For intuition why, [RR94] provides the following argument. Consider a straightforward inductive approach common in Ramsey theory. Partition the vertex set into two parts, U and V. Expose the edges in two phases. First, generate the edges of $G(n,p)$ that go between U and V. Let an adversary color them. For each vertex u in U consider the most common color with respect to edges incident on u. Now, there is at least one color, say red, such that red is the most common color for at least $\frac{|U|}{r}$ of the vertices. Call this set S and consider the subgraph H of the complete graph on S defined by edges (x,y) such that both x and y are connected to a common vertex in U by edges colored red. If H contains a complete subgraph of size $\Omega(n)$, then we could apply induction on $r-1$ colors to show that there will, almost surely, be at least one monochromatic triangle in $G(n,p)$. This is because, if the adversary does not want to create even one monochromatic triangle, then he cannot color any edge of H (that is generated) by the color red. In our case, the adversary does not mind generating a

few monochromatic triangles, so our task in finding the subgraph H is even harder. Unfortunately, H does not have such a nice structure. However, using the Szemerédi Regularity Lemma, one can find a well structured, large subgraph of H, on which induction can be applied. (This will become clear later.) Therefore, the induction hypothesis has to be strengthened. We will not deal with the random subgraph of the complete graph but with random subgraphs of $(\geq d, \epsilon, f)$-regular graphs. We will however, retain the two-round exposure of edges ($p = c/\sqrt{n} = (c_1 + c_2)/\sqrt{n}$), first putting in edges with probability $c_1/\sqrt{n}$ and later with probability $c_2/\sqrt{n}$. Almost surely, the graph obtained after the first round will have $\Theta(c_1 n^{3/2})$ edges, so we have to contend with the $r^{c_1 n^{3/2}}$ colorings the adversary can choose from. So, we have to prove our theorem with failure probability $\exp[-\Omega(cn^{3/2})]$. The actual theorem is as follows:

Theorem 4.12 *For all $r \geq 1$ there exists f so that for all $d \in (0, 1]$ there exists $\epsilon > 0$ so that the following holds: If G is any $(\geq d, \epsilon, f)$-regular graph on nf vertices with respect to f sets of size n each, then, for sufficiently large c, with $p = c/\sqrt{n}$, there exists $\alpha' > 0$ so that for all sufficiently small α,*

$$\Pr(MT_r(G_p) < \alpha c^3 n^{3/2}) < \exp[-\alpha' cn^{3/2}].$$

□

It is easy to see that the main theorem, Theorem 4.2 follows from the above theorem. Note that K_{fn}, the complete graph on fn vertices, contains a $(\geq d, \epsilon, f)$-regular graph (with $d = 1$). Apply Theorem 4.12 to this subgraph after renormalizing the

number of vertices and, by the monotonicity of MT_r, we are done.

We start with the base case: $r = 1$. Even this, very unusual in Ramsey theory, is quite nontrivial. We are now simply looking at the number of triangles in the random subgraph G_p and want to show, that with probability $1 - \exp[-cn^{3/2}]$, there will be $\Omega(cn^{3/2})$ triangles. For this, we need Janson's inequality for the lower tail of sums of random variables with limited dependence.

Base Case: One Color

In this case, $r = 1$. Fix $f = 3$. Let A_1, A_2, A_3 be sets such that $|A_1| = |A_2| = |A_3| = n$. Define G_1 to be a 3-partite graph on (A_1, A_2, A_3). Recall that $\rho(A_i, A_j)$ is the density of the induced bipartite subgraph $G(A_i, A_j)$. Let $\rho(A_i, A_j) = d_{ij}$. Let $d = \min\{d_{12}, d_{13}, d_{23}\}$. Let ϵ satisfy the following condition $\epsilon < d/2$. Also let $G(A_i, A_j)$ be $(\geq d, \epsilon)$-regular. Let G_p denote the Bernoulli random graph where each edge of G_1 is placed with probability p, independently of each other. For each triangle $t = \{x, y, z\}$ in G_1, let A_t be the event that the three edges xy, yz and xz all lie in G_p. Let T be the random variable denoting the number of triangles in G_p. Note that $T = MT_1$. Let μ be the expected number of triangles in G_p.

Lemma 4.13 *Let $d \in (0, 1]$ and $\epsilon < \frac{d}{2}$. Let G_1 be a 3-partite graph, as above, that is $(\geq d, \epsilon, 3)$-regular. Then, for $p = c/\sqrt{n}$, with $c > 1$, there exists $\alpha' > 0$ so that for all sufficiently small α the following holds:*

$$\Pr(T \leq \alpha c^3 n^{3/2}) \leq \exp[-\alpha' c n^{3/2}].$$

Proof: For any vertex x, let $N_i(x)$ denote the set of vertices adjacent to x in A_i and let $\deg(x, A_i) = |N_i(x)|$. For $i \in \{2, 3\}$, let X_i be the set of $x \in A_1$ such that $\deg(x, A_i) < (d_{1i} - \epsilon)n$. Then, by the definition of ϵ-regularity, we have $|X_i| < \epsilon n$.

Consider $S = A_1 - (X_2 \cup X_3)$. Then $|S| > (1 - 2\epsilon)n$. Note that for $x \in S$, we have $N_i(x) \geq (d_{1i} - \epsilon)n > \epsilon n$ [by the condition on ϵ]. Fix a vertex x in S. Consider $G(N_2(x), N_3(x))$. Again, by the definition of ϵ-regularity, this graph has density d' such that $|d' - d_{23}| \leq \epsilon$.

Let $T(x)$ denote the number of triangles in G_1 containing x. Then $T(x) \geq (d_{12} - \epsilon)(d_{13} - \epsilon)(d_{23} - \epsilon)n^2$. Summing over all $x \in S$, we get the total number of triangles in G_1 is $T_1 \geq (d_{12} - \epsilon)(d_{13} - \epsilon)(d_{23} - \epsilon)(1 - 2\epsilon)n^3 \geq n^3(d^3 - 5\epsilon)$. Note that this is a strictly positive fraction of n^3 because of the condition on ϵ.

Now we use Janson's estimate for the lower tail to show that, with very high probability, there will be $\Theta(c^3 n^{3/2})$ triangles in G_p. By Janson's inequality, for $\beta \in [0, 1]$,

$$Pr(T \leq (1 - \beta)\mu) \leq \exp[-\frac{1}{2}\frac{(\beta\mu)^2}{\mu + \Delta}],$$

where $\Delta = \sum_{s \sim t} E[A_s \wedge A_t]$, s and t being triangles in G_1; and μ is the expected number of triangles in G_p. Then $\mu = \sum \Pr[A_t] \geq n^{3/2} c^3(d^3 - 5\epsilon)$, and $\Delta \leq 3n^4 p^5 \leq 3c^5 n^{3/2}$. For $c > 1$, this gives $\mu + \Delta \leq 4c^5 n^{3/2}$.

Therefore, we have

$$\Pr(T \leq (1 - \beta)(d^3 - 5\epsilon)c^3 n^{3/2}) \leq \exp[-\frac{\beta^2}{8}(d^3 - 5\epsilon)^2 c n^{3/2}].$$

It is easy to see that we can renormalize the constants and then the base case is proven. □

We go from r to $r+1$ in two phases. Start with a $(\geq d, \epsilon, f^+)$-regular graph G. We will choose $f^+ = R_r(f)$ as in Definition 4.9; f being the requisite number of vertex sets for the r-color case. Put in edges with probability $c_1/\sqrt{n}$. (Later, we will put in edges with probability $c_2/\sqrt{n}, c = c_1 + c_2$, with c_2 much bigger than c_1). Let the adversary color these edges. Then, with failure probability $\exp[-\Omega(c_1 n^{3/2})]$, for any coloring, we show that we can pick a color, say red, and a subgraph H of G that is $(\geq d', \epsilon', f)$-regular. For the Rödl-Ruciński theorem, we need that every edge of H is NotRed, i.e., every edge (x, y) of H has both vertices x and y connected to a common vertex z by edges colored red (by the adversary). Therefore, no edge of H can be colored red if it is picked in the second phase because it would result in a monochromatic triangle. However, for our theorem, this is not sufficient since the adversary does not mind creating a few monochromatic triangles. So we need the notion of VeryNot$w$ edges. A pair (x, y) is called VeryNot$w$ if there are δc_1^2 vertices, $0 < \delta < 1$, all joined to both x and y by edges colored w (by the adversary after Phase 1). The subgraph H consists of only VeryNot$w$ edges, for some color w. A VeryNot$w$ edge if colored w creates δc_1^2 monochromatic triangles. For simplicity, let us say that H contains VeryNotRed edges. In the second phase we put in edges with probability $c_2/\sqrt{n}$. But we will focus our attention on H. Every edge in the random subgraph H_p of H that is colored red creates δc_1^2 monochromatic triangles. The total number of monochromatic triangles created is then at least the number

of monochromatic triangles of H_p plus δc_1^2 times the number of edges of H_p colored red. We will show that this term is at least $\Omega(c_1^3 n^{3/2})$, with failure probability $\Omega(-c_2 n^{3/2})$. We will do this by considering an $r+\frac{1}{2}$ coloring of H. An $r+\frac{1}{2}$ coloring is really an $r+1$ coloring where the use of one particular color, say red, is costly. For any $r+1$ coloring of a graph, let λ be the number of red edges and MT_{r+1} the number of monochromatic triangles. Then let $\Lambda = \delta c_1^2 \lambda + MT_{r+1}$. The theorem for $r+\frac{1}{2}$ colors would say that with high probability, for any $r+1$ coloring of the random subgraph H_p, we have $\Lambda = \Omega(c_1^3 n^{3/2})$. The precise statement is given in the next subsection. The theorem for $r+1$ colors will follow directly from the theorem for $r+\frac{1}{2}$ colors. We will see this after the next subsection.

4.4.4 r and a half colors

For convenience, we will state the theorem for $r+\frac{1}{2}$ colors as an implication from the theorem for r colors to the statement about the $r+\frac{1}{2}$ coloring.

Theorem 4.14 *Suppose that if G is $(\geq d, \epsilon, f)$-regular on fn vertices with respect to f groups of n each then for $r \geq 1, c \geq c_0$, c_0 independent of n, any r-coloring of the edges of G'_p, $p' = c/\sqrt{n}$, satisfies*

$$\Pr(MT_r(G'_p) < \beta c^3 n^{3/2}) < \exp[-\beta' c n^{3/2}]$$

where β' is dependent upon ϵ, d, f, and β is sufficiently small.

Let $0 < c_1 < c_2$ and $\gamma \in (0,1)$. Let $p = c_2/\sqrt{n}$. Consider an $(r+1)$-coloring of G_p. Let R be the subgraph of G_p consisting of all the red edges of G_p. Let

MT_r be the number of monochromatic triangles in the r-coloring of $G_p \backslash R$. Let $\Lambda = \gamma c_1^2 |R| + MT_r(G_p \backslash R)$; and let F_α be the event: $\Lambda < \alpha c_1^3 n^{3/2}$. Then, for sufficiently large c_2, there exists α' dependent upon ϵ, r, d, γ, such that, for sufficiently small α, the following holds:

$$\Pr(F_\alpha) \leq \exp[-\alpha' c_2 n^{3/2}].$$

Proof: The idea is as follows: First, we will condition on G_p having roughly $c_2 n^{3/2}$ edges. Given that, we will bound the probability of the event F (which we will call a failure) by the failure probability of an adversary using a random coloring times some appropriately small factor.

Let the failures be:

1) *failure event A*: G_p has too few or too many edges.

$$|G_p| < \tfrac{1}{4} c_2 n^{3/2} d f^2 \text{ or } |G_p| > \tfrac{1}{2} c_2 n^{3/2} f^2.$$

By the Chernoff bound, there exists $\eta > 0$ such that

$$\Pr(A) < \exp[-\eta c_2 n^{3/2}].$$

2) *failure event B*: $B = \vee B_{e,e_1}$, where $\frac{1}{4} c_2 n^{3/2} d f^2 \leq e < \frac{1}{2} c_2 n^{3/2} f^2$, $e_1 < \frac{\alpha}{\gamma} c_1 n^{3/2}$, and B_{e,e_1} is the event that

$|G_p| = e$ and $\exists R, |R| = e_1, R \subseteq G_p$, and event C: There exists an r-coloring of G_p such that $MT_r(G_p \backslash R) < \alpha c_1^3 n^{3/2}$.

Choose α sufficiently small so that:

1. $\frac{1}{2} f^2 H(\frac{5\alpha}{df^2\gamma}) < \frac{\beta'}{2f^2}(\frac{1}{4} f^2 d - \frac{\alpha}{\gamma})$, H being the binary entropy function. ($H(x) =$

$-x\log_2 x-(1-x)\log_2(1-x).$) 2. $\alpha < \frac{\beta}{2f^2}(\frac{1}{2}f^2d - \frac{\alpha}{\gamma})$.

3. $\alpha < \frac{1}{4}\gamma f^2 d$.

First, note that if there indeed is an $(r+1)$-coloring of the e edges of G_p with e_1 red edges that causes failure, then with a random choice of e_1 red edges, the probability of failure is more than $\frac{1}{\binom{e}{e_1}}$. Also, the probability of an event occurring in a random graph with e edges of which e_1 are removed randomly is the same as the probability of the event occurring in a random graph with $e-e_1$ edges. Therefore,

$$\Pr(B_{e,e_1}) \leq \binom{e}{e_1} \Pr[\text{random graph with } e-e_1 \text{ edges satisfies event C}].$$

We now switch back to the Bernoulli model. We state, without proof, the following easy lemma.

Lemma 4.15 *Let G be a fixed graph on n vertices with m edges. Let $g(n,e)$ be the probability of some event A occurring on the random subgraph of G with e edges. Also, let g be a decreasing function of e. Let $f(n,p)$ be the probability that the event occurs on the Bernoulli random subgraph of G.*

Then, if $p = \frac{e}{m}$, $g(n,e) \leq 2f(n,\frac{p}{2})$. □

We will apply the above lemma for

$$p' = \frac{e-e_1}{2\binom{n}{2}(\sum d_{ij})} \geq \frac{\frac{1}{4}f^2dc_2n^{3/2} - \frac{\alpha}{\gamma}c_1n^{3/2}}{2f^2n^2} \geq \frac{\frac{1}{4}f^2dc_2n^{3/2} - \frac{\alpha}{\gamma}c_2n^{3/2}}{2f^2n^2} = \delta c_2/\sqrt{n}.$$

From the conditions on the choice of α, the value of δ is positive. Choose c_2 sufficiently large so that $\delta c_2 \geq c_0$.

$$\Pr(B_{e,e_1}) \leq 2\binom{e}{e_1} \Pr[G_{p'} \text{ satisfies event C}],$$

Summing over all possible choices of e and e_1,

$$\Pr(B) \leq n^4 \max_{e,e_1} \Pr(B_{e,e_1}) \leq 2n^4 \binom{e}{e_1} P(G_{p'} \text{ satisfies event C}).$$

Applying Lemma 1,

$$\Pr(B) \leq 2n^4 \exp[\tfrac{1}{2}c_2 n^{3/2} f^2 H(\frac{4\alpha c_1}{df^2\gamma c_2} + o(1))] \exp[-\beta'\delta c_2 n^{3/2}],$$

where H is the binary entropy function. By our choice of α, there exists a $\zeta > 0$ such that for sufficiently large n,

$$\Pr(B) \leq \exp[-\zeta c_2 n^{3/2}].$$

Combining the two cases, we have

$\Pr(F_\alpha) \leq \exp[-\eta c_2 n^{3/2}] + \exp[-\zeta c_2 n^{3/2}]$. Choose $\alpha' = \frac{1}{2}\min\{\eta, \zeta\}$. Then, for sufficiently large n,

$$\Pr(F_\alpha) \leq \exp[-\alpha' c_2 n^{3/2}].$$

This completes the proof of the theorem for $r + \frac{1}{2}$ colors. □

4.4.5 r + 1 colors

Now, we can complete the induction from r to $r+1$ colors. The basic idea is that we will start with a $(\geq d, \epsilon, f^+)$-regular graph, and put in edges of Phase 1 with probability $p = c_1/\sqrt{n}$. Then we will show that with high probability (i.e. with failure probability $\exp[-\Omega(c_1 n^{3/2})]$), we can extract a $(\geq d', \epsilon', f)$-regular graph H of VeryNot$w$ edges, for some color w. Now, we can apply the result we proved above

for $r+\frac{1}{2}$ colors, because each VeryNot$w$ edge creates $\Omega(c_1^2)$ monochromatic triangles if colored w. In order to construct the graph H, we need a few lemmas which we obtain by modifying the lemmas in [RR94].

Lemma 4.16 *Let A be a set such that $|A| \leq kn^2$. Let $A_1, \ldots, A_l$ be subsets of A, $l = \epsilon c^2, |A_i| \geq \epsilon n^2$. Then, there exists ϵ', dependent upon ϵ and k, and $S \subseteq A, |S| \geq \epsilon' n^2$ such that every $x \in S$ lies in at least $\epsilon' c^2$ of the A_i.*

Proof: Suppose that $|S| < \epsilon' n^2$. Then, we would have $\epsilon' n^2(\epsilon c^2) + kn^2(\epsilon' c^2) \geq \epsilon^2 c^2 n^2$. This implies $\epsilon'\epsilon + k\epsilon' \geq \epsilon^2$; choosing ϵ' small enough we get a contradiction. □

Lemma 4.17 (Subset Lemma) [RR94]) *For all $d, \epsilon' > 0$, there exists $\epsilon > 0$ so that the following holds: Let G be $(\geq d, \epsilon)$-regular on A, B, both of size n. Let A', B' be random subsets of A, B respectively, each with probability $p = cn^{-1/2}$. Then the probability that G is <u>not</u> $(\geq d/2, \epsilon')$-regular on $A' \times B'$ is less than $\exp[-\Omega(c\sqrt{n})]$.*

Lemma 4.18 *For all $d, \gamma > 0$, there exist $\epsilon, \delta > 0$ such that the following holds: Let G be $(\geq d, \epsilon)$-regular on A, B both of size n. Let $A_i, B_i, 1 \leq i \leq n/c^2$ be independent random subsets of A, B respectively, all with probability $p = cn^{-1/2}$. Let failure F be: $\exists I \subseteq \{1, \ldots, n/c^2\}, |I| = \gamma n/c^2$ and there exist for $i \in I$ subsets $A_i^* \subseteq A_i, B_i^* \subseteq B_i$ with $|A_i^*| \geq \gamma |A_i|, |B_i^*| \geq \gamma |B_i|$ such that*

$$|\bigcup_{i \in I} (A_i^* \times B_i^*) \cap G| < \delta n^2.$$

Then, for c sufficiently large, the probability of failure F is $\Pr(F) < \exp[-\Omega(\frac{n^{3/2}}{c})]$.

Proof: First, we ensure that for most $i \in I$, both A_i and B_i have the right size. Let $H_i = 1$ if $|A_i - c\sqrt{n}| > \frac{1}{2}c\sqrt{n}$ or $|B_i - c\sqrt{n}| > \frac{1}{2}c\sqrt{n}$, and 0 otherwise. Call index i *bad* if $H_i = 1$, *good* otherwise. By the Subset Lemma from [RR94], $E(H_i) < \exp[-\Omega(c\sqrt{n})]$. Let H be the failure event that $\sum H_i > \frac{1}{2}\gamma\frac{n}{c^2}$. By independence of the different H_i,

$$\Pr(H) < 2^n \exp[-c\sqrt{n}\tfrac{1}{2}\gamma\frac{n}{c^2}] = \exp[-\Omega(\frac{n^{3/2}}{c})].$$

Let I_0 be the set of good indices. With high probability, $|I_0| \geq \frac{n}{c^2}(1 - \frac{1}{2}\gamma)$. Failure F' occurs if, given H does not occur, there exists $I_1 \subseteq I_0, |I_1| > \frac{1}{2}\gamma\frac{n}{c^2}$, and

$$|\bigcup_{i\in I_1}(A_i^* \times B_i^*) \cap G| < \delta n^2$$

To count the total number of edges in the union, we use a sequencing argument. Order the elements of $I_1 : 1, \ldots, |I_1|$. Consider a sequence of graphs H_i on $A \times B$. Initially, H_0 has no edges. After the ith round, $H_i = \bigcup_{i\in\{1,\ldots,i\}}(A_i^* \times B_i^*) \cap G$. Let N_i be the new edges added in the ith round. $N_i = |H_i| - |H_{i-1}|$. Clearly, we can reorder the elements of I_1 so that for all i: $N_i \geq N_{i+1}$. Also, let J be the index so that for all $i \leq J$, we have $N_i \geq (\frac{d}{16}\gamma^2)c^2n$, and for all $i > J$, the reverse holds. There are at most $2^n n! n$ ways to choose I_1, the ordering $<$ on I_1, and the index J. For failure to occur, definitely $J < \lambda\gamma\frac{n}{c^2}$, where λ is proportional to δ. For each choice of $I_0, <, J$, with high probability, there are at most $2^{2J(2c\sqrt{n})}$ choices of A_i^*, B_i^* for $i \leq J$. Pick δ small enough so that the removal of any δn^2 edges from G leaves a G^- that is $(\geq d/2, \epsilon/2)$-regular (by Lemma 2.3 of [RR94]). Let $G^- = G \setminus \bigcup_{i\in\{1,\ldots,J\}}(A_i^* \times B_i^*) \cap G$. For any $\epsilon' < \gamma, d/4$, if G^- is $(\geq d/2, \epsilon')$-regular

on any $A_i \times B_i$, then $|A_i^* \times B_i^*| \cap G^-| \geq (\frac{d}{16}\gamma^2)c^2n$. But this cannot be true for $i > J$, by definition of J. This implies that for all $i > J$, G^- is not $(\geq d/2, \epsilon')$-regular on $A_i \times B_i$. For each i, let this event be F_i. Then, for $i > J$, we have $\Pr(F_i) = \exp[-kc\sqrt{n}]$. By independence of the F_i,

$$\Pr(\bigwedge_{i>J} F_i) < \exp[-k(1-\lambda)\frac{n^{3/2}}{c}].$$

Therefore, the total failure probability of the event F is

$$\Pr(F) \leq \Pr(H) + (2^n n! n)\exp[2\lambda\frac{n}{c^2}(2c\sqrt{n})]\exp[-k(1-\lambda)\frac{n^{3/2}}{c}].$$

Choose δ small enough so that λ is small enough so that $\Pr(F) \leq \exp[-\Omega(\frac{n^{3/2}}{c})]$. □

Lemma 4.19 *For all $d, \gamma > 0$, there exist $\epsilon, \delta > 0$ such that the following holds: Let G be $(\geq d, \epsilon)$regular on A, B both of size n. Let $A_{ji}, B_{ji}, 1 \leq i \leq n/c^2, 1 \leq j \leq c^2$ be independent random subsets of A, B respectively, all with probability $p = cn^{-1/2}$. For each j, let F_j be the failure F defined above. Let X_j be 1 if F_j occurs, 0 otherwise. Let $X = \sum_{j=1}^{c^2} X_j$. Then for $\zeta > 0$, $\Pr(X \geq \zeta c^2) < \exp[-\Omega(cn^{3/2})]$.*

Proof: By Lemma 4.18, $\Pr(X_j = 1) < \exp[-\Omega(\frac{n^{3/2}}{c})]$. Then $\Pr(X \geq \zeta c^2) \leq \binom{c^2}{\zeta c^2}\exp[-\Omega(\zeta cn^{3/2})] \leq \exp[-\Omega(cn^{3/2})]$ □

Lemma 4.20 *Let A be a function from $\{1,\ldots,n\} \times \{1,\ldots,m\} \times \{1,\ldots,n\} \times \{1,\ldots,m\} \rightarrow \{0,1\}$. Let $|A^{-1}(1)| = N$, and let $N < n^2$. Then there exists a choice of indices $S = \{a_1,\ldots,a_m\}$ where $1 \leq a_i \leq n$ so that for any distinct $\alpha, \beta \in S$, and $i, j \in \{1,\ldots,m\}$, we have $A(a_i, i, a_j, j) = 0$.*

Proof: Choose $b_1, \ldots, b_m$ uniformly and independently from $\{1, \ldots, m\}$. Then, the number of 1's picked is the sum $\sigma = \sum_{i,j} A(b_i, i, b_j, j)$. The expectation of σ is $E(\sigma) = N/n^2 < 1$. Therefore, there is a choice of the b_i, which we call a_i, so that no 1 is picked. □

We are now ready to complete the proof for $r+1$ colors. We will do this in two phases. Let $p = \frac{c}{\sqrt{n}} = \frac{c_1+c_2}{\sqrt{n}}$. We will set $c_1 = \zeta c$ and $c_2 = (1-\zeta)c$, where ζ, to be fixed later, will be a small positive constant less than 0.5. In phase 1, we put in edges with probability $\frac{c_1}{\sqrt{n}}$. Then, we let the adversary color the edges. We label a pair $(x, y) \in G$ VeryNot-w if after Phase 1, there exist δc_1^2 vertices all connected to both x and y by edges colored w. The choice of δ will follow from the next lemma. Then we show that for any adversarial coloring, one can find a color w such that the graph of VeryNot-w edges of G contains a $(\geq d', \epsilon', f)$-regular graph H. Then apply the result for $(r + \frac{1}{2})$ colors for the induction to work.

Phase 1: Consider a $(\geq d, \epsilon, f^+)$-regular graph G on f^+ sets of n vertices each. Here, f is requisite number of sets for the r-color case, and $f^+ = R_r(f)$, as in Definition 4.9. The choice of the constants will become clear over the course of the proof. Put in edges with probability $p = \frac{c_1}{\sqrt{n}} = \frac{c_1' C_0}{\sqrt{2n}}$. Here, C_0 will correspond to the number of classes in the Szemerédi Regular Partitioning, that each vertex set is split into. We need the following useful lemma.

Lemma 4.21 *Let $A_0, \ldots, A_{r+2}$ be any distinct $r+3$ sets of the vertex sets of G. Let $p = c_1' \sqrt{\frac{C_0}{n}}$. Let $m = \frac{n}{C_0}$. Consider the random subgraph of G with edge probability p. Let S be the following event: For any $r+1$ coloring of the random subgraph of G,*

and for all $B_0, \ldots, B_{r+2}, B_i \subseteq A_i, |B_i| = m$, *there exist* w, i, j, δ *such that* $|E| > \delta m^2$, *where* $x \in E$ *if* $x \in (A_i \times A_j) \cap G$ *and* x *is VeryNot-*w.

Then, $P(\neg S) < \exp[-\Omega(c_1' n^{3/2})]$.

Proof: There are less than $2^{O(n)}$ choices of the B_i. Fix a choice $B_0, \ldots, B_{r+2}$. Let $\gamma = \frac{1}{(r+1)\binom{r}{2}}$. Let for $x \in B_0$, let B_i^x denote the neighbors of x in B_i in the random graph. For any vertex $x \in B_0$, for any coloring, there will be a color w and sets B_j, B_k such that $\frac{1}{r+1}$ of the edges from i to both B_j^x and B_k^x are colored w. Then, there will be γm vertices $x \in B_0$ all having the same pair B_j, B_k and color w. Call a pair of vertices (x, y) Not-w if there exists a vertex z connected to both x and y by edges colored w. Split these γm vertices into c^2 groups of equal size. By Lemma 4.18, each group produces $\delta' m^2$ Not-w pairs, with failure probability $\exp[-\Omega(\frac{n^{3/2}}{c_1'})]$. Now, by Lemma 4.19, with failure probability $\exp[-\Omega(c_1' n^{3/2})]$, we have δm^2 pairs each of which becomes Not-w $\delta(c_1')^2$ times. That is, we have $|E| > \delta m^2$. Therefore, $P(\neg S) < 2^{O(n)} \exp[-\Omega(c_1' n^{3/2})] \leq \exp[-\Omega(c_1' n^{3/2})]$. □

Let $H_w \subseteq G$ be the graph of VeryNot-w edges. We have, therefore, $r + 1$ graphs $H_1, \ldots, H_w$. Apply Lemma 4.5 to obtain a common ϵ'-regular partitioning. Choosing $\epsilon' < \frac{1}{(f^+)^2}$, we can apply Lemma 4.20 to obtain sets $B_1, \ldots, B_{f+}$ each of size $\frac{n}{C_0}$, such that B_i, B_j are all ϵ'-regular with respect to all H_k simultaneously. From these $\mathcal{B} = \{B_1, \ldots, B_{f+}\}$, we will extract the new graph. (We want it to be $(\geq d', \epsilon', f)$-regular.) Create a new graph M on f^+ vertices $N = \{1, \ldots, f^+\}$. Add an edge labelled Not-w between i and j if $B_i \times B_j$ contains $\geq \delta m^2$ VeryNot-w edges. If there is more than one candidate color/label for any pair (i, j), choose one arbitrarily. If

there is no candidate label for (i,j), add an edge labelled Not-Not. It is immediate from Lemma 4.21 that with probability $\exp[-\Theta(c_1 n^{3/2})]$, M does not have a clique of size $r+2$ of Not-Not edges. Therefore, by Ramsey's theorem, Lemma 4.7, there is a clique of size f of Not-w edges for some color w. Let the winning color be denoted by W; $W = w$. Renumber the sets in $\mathcal{B}$ so that $B_1, \ldots, B_f$ form this clique. Let $H^* = H_w$ restricted to the union of the sets $B_1, \ldots, B_f$. This, by Lemma 4.21, is a $(\geq d', \epsilon', f)$-regular graph on f sets of n/C_0 vertices each. Remove the edges of H^* already generated in Phase 1 to obtain the graph H. There are only $O(n^{3/2}) = o(n^2)$ of them (otherwise we count it as a failure), so the graph is still ϵ''-regular, by basic facts about *epsilon*-regularity [RR94].

First, we define Phase 1 failures.

Failure Q_1 occurs if the total number of edges generated in Phase 1 is more than $2\frac{c_1}{\sqrt{n}}(f^+)^2 n^2$. By Chernoff bounds, $\Pr(Q_1) < \exp[-\Omega(c_1 n^{3/2})]$.

Failure Q_2 occurs if the $(\geq d', \epsilon'', f)$-regular graph H cannot be produced. By Lemma 4.19, $\Pr(Q_2) < \exp[-\Omega(c_1 n^{3/2})]$. If Phase 1 is a success, there are only $(r+1)^{O(c_1 n^{3/2})}$ possible choices of H.

Phase 2: Now we apply induction. Consider the random subgraph of H with probability $p = \frac{c_2}{\sqrt{n}}$. Consider any $r+1$ coloring of the graph. Any time an edge is colored W (the winning color of Phase 1), it creates δc_1^2 monochromatic triangles. Let R be the subgraph obtained by selecting all the edges of H colored W. Therefore, the total number of monochromatic triangles is at least

$$\delta c_1^2 |R| + M T_r(H \backslash R)$$

Let a Phase 2 failure be the event that this sum is less than $\alpha c_1^3 n^{3/2}$. Choosing c_2 large enough so that we can apply the Theorem 4.14, this probability is less than $\exp[-\gamma' c_2 n^{3/2}]$, where γ' depends on δ but not on c_1. Therefore, the total probability of failure is

$$\Pr(F) < \exp[-K_1 c_1 n^{3/2}] + \exp[K_2 c_1 n^{3/2}] \exp[-K_3 c_2 n^{3/2}].$$

The first term accounts for Phase 1 failures. The $\exp[K_2 c_1 n^{3/2}]$ term accounts for the different ways the adversary can color the graph H. The $\exp[-K_3 c_2 n^{3/2}]$ term accounts for Phase 2 failures. Pick small positive ζ so that

$$K_2\zeta - K_3(1-\zeta) = -K_4$$

with K_4 positive. Set $K_1\zeta = K_5, c_1 = \zeta c$, and $c_2 = (1-\zeta)c$, so that $c = c_1 + c_2$. Therefore, the failure probability is $\Pr(F) = \exp[-\Omega(cn^{3/2})]$. This proves the theorem. □

Chapter 5

Conclusions, and Future Work

In this thesis, we have discussed three problems arising in Theoretical Computer Science and Combinatorics, all involving randomness. We now mention some relevant open problems.

In Chapter **2**, we showed that, for most sets, slightly-random sources are nearly as good as truly random sources. In particular, we proved that, for most sets, the ϵ-biased probability of hitting the set is bounded away from 0. What are the exceptional sets? The set of binary strings with more 1s than 0s, the Majority set, is an example. Can we get a precise characterization of all the exceptional sets? Shamir [Sha87] used a Martingale technique to prove a strong concentration result for the ϵ-biased probability of hitting a random set. Unfortunately, his technique fails for $\epsilon > 0.207$. It would be interesting to extend his result to all fixed $\epsilon < \frac{1}{2}$.

In Chapter **3**, we proved the existence of perfect information coin flipping protocols and leader election protocols that are immune to ϵn cheaters, for every $\epsilon < \frac{1}{2}$.

What about explicit constructions of such protocols? Alon and Naor [AN93] and Cooper and Linial [CL93] provide explicit protocols for small, but fixed, ϵ. It would be very interesting to extend these to every $\epsilon < \frac{1}{2}$.

Finally, in Chapter 4, we looked at a problem in Ramsey theory. In particular, we showed that if an adversary r-colors the edges of the Binomial random graph, with edge probability $\frac{c}{\sqrt{n}}$, c being a large enough constant, then almost surely, a fraction of the triangles in the graph will be monochromatic. Extending this result to hypergraphs is a natural problem.

Bibliography

[AL89] M. Ajtai and N. Linial. The influence of large coalitions. *IBM Research Report 7133*, 67380, November 1989. Journal version to appear in Combinatorica.

[AN93] N. Alon and M. Naor. Coin-flipping games immune against linear-sized coalitions. *SIAM Journal on Computing*, 22:403–417, 1993.

[AR89] N. Alon and M.O. Rabin. Biased coins and randomized algorithms. In S. Micali, editor, *Advances in Computing Research*, volume 5, pages 499–507. JAI Press, Greenwich, Connecticut, 1989.

[ASE92] N. Alon, J.H. Spencer, and P. Erdős. *The Probabilistic Method.* John Wiley and Sons, Inc., New York, 1992.

[BL85] M. Ben-Or and N. Linial. Collective coin flipping, robust voting schemes and minima of Banzhaf values. *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science, Portland, Oregon*, pages 408–416, 1985.

[BL89] M. Ben-Or and N. Linial. Collective coin flipping. In S. Micali, editor, *Advances in Computing Research*, volume 5, pages 91–116. JAI Press, Greenwich, Connecticut, 1989.

[BLS87] M. Ben-Or, N. Linial, and M. Saks. Collective coin flipping and other models of imperfect randomness. *Colloq. Math. Soc. Janós Bolyai No. 52 Combinatorics Eger*, pages 75–112, 1987. North-Holland Publishing Company.

[Blu82] M. Blum. Coin flipping by telephone. *IEEE COMPCOM*, 1982.

[Blu84] M. Blum. Independent unbiased coin flips from a correlated biased source: a finite state Markov chain. *Proceedings of the 25th Annual Symposium on the Foundations of Computer Science*, pages 425–433, 1984.

[BM73] H. Burkill and L. Mirsky. Monotonicity. *Journal of Mathematical Analysis and Applications*, 41:391–410, 1973.

[BN93a] R. Boppana and B. Narayanan. The biased coin problem. *Proceedings of the 25th Annual ACM Symposium on Theory of Computing, San Diego, California*, pages 252–257, 1993.

[BN93b] R. Boppana and B. Narayanan. Collective coin flipping and leader election with optimal immunity. 1993. Submitted to the 34th Annual IEEE Symposium on Foundations of Computer Science.

[Bol85] B. Bollobás. *Random Graphs.* Academic Press, Orlando, Florida, 1985.

[Bol86] B. Bollobas. *Combinatorics.* Cambridge University Press, New York, New York, 1986.

[BT86] B. Bollobás and A. G. Thomasson. Threshold functions. *Combinatorica*, 7:35–38, 1986.

[CG87] B. Chor and M. Gereb-Graus. On the influence of a single participant in coin flipping schemes. Technical Report TR-06-87, Harvard University, 1987.

[CG88] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17:230–261, 1988.

[CL93] J. Cooper and N. Linial. Fast perfect-information leader-election protocol with linear immunity. *Proceedings of the 25th Annual ACM Symposium on Theory of Computing, San Diego, California*, 1993.

[EHS+93] P. Erdős, A. Hajnal, M. Simonovits, V. T. Sos, and E. Szemerédi. Turan-Ramsey theorems and simple asymptotically extremal structures. *Combinatorica*, 13:31–56, 1993.

[ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Kut. Int. Közl*, 5:17–61, 1960.

[Erd47] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematics Society*, 53:292–294, 1947.

[Erd61] P. Erdős. Graph theory and probability II. *Canadian Journal of Mathematics*, 13:346–352, 1961.

[FM88] P. Feldman and S. Micali. Optimal algorithms for Byzantine agreement. *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, Chicago, Illinois*, pages 148–161, 1988.

[Fol70] J. Folkman. Graphs with monochromatic complete subgraphs in every edge coloring. *SIAM Journal of Applied Mathematics*, 18:19–29, 1970.

[FR86] P. Frankl and V. Rődl. Large triangle-free subgraphs in graphs without K_4. *Graphs and Combinatorics*, 2:135–144, 1986.

[GM82] S. Goldwasser and S. Micali. How to generate cryptographically strong sequences of pseudo-random bits. *Proceedings of the 23th Annual Symposium on the Foundations of Computer Science, Chicago, Illinois*, 1982.

[GRS90] R.L. Graham, B.L. Rothschild, and J.H. Spencer. *Ramsey Theory.* Wiley, New York, 2nd edition, 1990.

[HLP52] G.H. Hardy, J.E. Littlewood, and G. Pòlya. *Inequalities.* Cambridge University Press, New York, New York, 2nd edition, 1952.

[Jan90] S. Janson. Poisson approximation for large deviations. *Random Structures and Algorithms*, 1:221–230, 1990.

[KG] W. Kennedy and J. Gentle. *Statistical Computing.* Marcel Dekker, Inc.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science, White Plains, New York*, pages 68–80, 1988.

[LRV92] T. Luczak, A. Ruciński, and B. Voigt. Ramsey properties of random graphs. *Journal of Combinatorial Theory, Series B*, 56, 1992.

[Mur70] H.F. Murry. A general approach for generating natural random variables. *IEEE Transactions on Computers*, C-19:1210–1213, 1970.

[NR84] J. Nesetril and V. Rödl. Sparse Ramsey graphs. *Combinatorica*, 4:71–78, 1984.

[Rab76] M. Rabin. Probabilistic algorithms. In J. Traub, editor, *Algorithms and Complexity*. Academic Press, New York, 1976.

[Rab83] M. Rabin. Randomized Byzantine generals. *Proceedings of the 24th Annual IEEE Symposium on Foundations of Computer Science, Tucson, Arizona*, pages 403–409, 1983.

[Ram30] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 30:264–286, 1930.

[RR94] V. Rödl and A. Ruciński. Random graphs with monochromatic triangles in every edge coloring. *Random Structures and Algorithms*, 5(2), 1994. (To appear).

[Sak89] M. Saks. A robust noncryptographic protocol for collective coin flipping. *SIAM Journal on Discrete Mathematics*, 2:240–244, 1989.

[Sch80] J. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27:701–717, 1980.

[Sha87] E. Shamir. Concentration of a random set function on the n-cube, related to slightly random sources. manuscript, 1987.

[Spe87] J.H. Spencer. *Ten Lectures on the Probabilistic Method.* Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1987.

[Spe88] J. Spencer. Three hundred million points suffice. *Journal of Combinatorial Theory*, 49(2):210–217, 1988.

[Spe92] J.H. Spencer. The probabilistic method. *Proceedings of the 3rd Symposium on Discrete Algorithms, Orlando, Florida*, 1992.

[SV86] M. Santha and U.V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33(1):75–87, 1986.

[Sze78] E. Szemerédi. Regular partitions of graphs. *Problems Combinatories et Theorie des graphs ,Editions du C.N.R.S*, 260:399–402, 1978.

[Vaz86] U. V. Vazirani. *Randomness, Adversaries and Computation.* PhD thesis, University of California, Berkeley, 1986.

[Vaz87] U. V. Vazirani. Strong communication complexity or generating random sequences from two communicating slightly random sources. *Combinatorica*, 7(4):375–392, 1987.

[von63] J. von Neumann. Various techniques used in connection with random digits. In *Von Neumann's Collected Works*, pages 768–770. Pergamon Press, New York, 1963.

[VV85] U.V. Vazirani and V.V. Vazirani. Random polynomial time is equal to slightly random polynomial time. *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science, Portland, Oregon*, pages 417–428, 1985.

[Zuc91] D. Zuckerman. Simulating BPP using a general weak random source. *Proceedings of the 32nd Annual Symposium on the Foundations of Computer Science, San Juan, Puerto Rico*, pages 79–89, 1991.