4-27-2015

Probabilistic Inequalities and Bootstrap Percolation

Tomas Juskevicius

PROBABILISTIC INEQUALITIES AND BOOTSTRAP PERCOLATION

by

Tomas Juškevičius

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Mathematical Sciences

The University of Memphis

May 2015

ACKNOWLEDGEMENTS

I would like to thank my adviser, Dr. Béla Bollobás, for all of his guidance and assistance, his unfailing dedication to his students and for all of the wonderful opportunities he provided me with. Most of the results contained in this dissertation is joined work with mathematicians he introduced me to over the years.

I would also like to express my thanks to Dr. Paul Balister who always had his door open for me with any mathematical questions.

I am very much indebted to Gabriella Bollobás for her endless hospitality. She has always welcomed me in her home, made me feel very comfortable and cared for. She has helped me to develop as an individual as much as Béla has as a mathematician, and I am much in their debt for that.

I want to express my sincere gratitude to all students of our group (both former and current) and other graduate students in UoM for their companionship, good advice and collaboration. I would like to especially thank Richard Johnson, Jonathan Lee and Jee Zhou for their loyalty and friendship that extended beyond the University. Furthermore, I would like to mention the support and encouragement of my former colleagues - Jonas Jankauskas, Paulius Šarka and Matas Šileikis. A special thanks goes to my good non-mathematical friend Audrius Feigelovičius with whom I had a great pleasure to discuss mathematical puzzles and problems during the summers away from Memphis.

My thanks are also due to Ms. Tricia Simmons for her kindness and care. She has helped me with numerous issues during my stay in Memphis and I clearly would have been in trouble much more often if not for her.

And the last, but certainly not the least, I am indebted to my family whose unconditional support and constant care without which being far away from my country would have been much more difficult than it was.

ABSTRACT

Tomas Juškevičius. The University of Memphis. May 2015. Probabilistic Inequalities and Bootstrap Percolation. Major Professor: Béla Bollobás, Ph.D.

This dissertation focuses on two topics. Firstly, we address a number of extremal probabilistic questions:

- The Littlewood-Offord problem: we provide an alternative and very elementary proof of a classical result by Erdős that avoids using Sperner's Theorem. We also give a new simple proof of Sperner's Theorem itself.

- Upper bounds for the concentration function: answering a question of Leader and Radcliffe we obtain optimal upper bounds for the concentration function of a sum of real random variables when individual concentration information about the summands is given. The result can be viewed as the optimal form of a well-known Kolmogorov-Rogozin inequality.

- Small ball probabilities for sums of random vectors with bounded density: we provide optimal upper bounds the probability that a sum of random vectors lies inside a small ball and derive an upper bound for the maximum density of this sum. In particular, our work extends a result of Rogozin who proved the best possible result in one dimension and improves some recent results proved by Bobkov and Chystiakov [8]. This is joint work with Jonathan Lee.

- Two extremal questions of bounded symmetric random walks: we find distributions maximizing $\mathbb{P}(S_n \geq x)$ and $\mathbb{P}(S_n = x)$, where $S_n = X_1 + \cdots + X_n$ is a sum of independent bounded symmetric random variables. This is joint work with Matas Šileikis and Dainius Dzindzalieta [16].

The second part of the dissertation is concerned with a problem in Bootstrap Percolation. Let G be a graph and let $I \subset V(G)$ be a set of initially infected vertices. The

set of infected vertices is updated as follows: if a healthy vertex has the majority of its neighbours infected it itself becomes infected. Otherwise it stays healthy. In other words, we have a sequence of sets

$$I = I_0 \subset I_1 \subset \ldots \subset I_k \subset \ldots,$$

where $I_{k+1} = I_k \cup \{v \in V(G) : v \text{ has more infected than healthy neighbours}\}$. In the description of the bootstrap process above the superscripts of the sets correspond to the time steps when infections occur. If the process ends up infecting all of the vertices, i.e., $I_k = V(G)$ for some k, we say that percolation occurs.

In this dissertation we shall investigate this process on the Erdős-Renyi random graph $G(n,p)$. In this graph on n vertices each edge is included independently with probability p. We shall be interested in the smallest cardinality, say $m = m(n)$, of a uniformly chosen initially infected set of vertices I, such that the probability of percolation at least $1/2$. We call this quantity the critical size of the initially infected set. In the regime $p > c\log(n)/n$ (the connectivity threshhold) we prove sharp upper and lower bounds for m that match in the first two terms of the asymptotic expansion.

This is joint work with Nathan Kettle and Cecilia Holmgren. The problem was suggested to us by Béla Bollobás and Robert Morris.

TABLE OF CONTENTS

Part I

Probabilistic Inequalities

CHAPTER 1

THE LITTLEWOOD-OFFORD PROBLEM AND AN EXTENSION OF SPERNER'S THEOREM

In this section we will relate the classical Littlewood-Offord problem and Sperner's Theorem, giving new and elementary proofs for both.

The Littlewood-Offord problem is a combinatorial question in geometry that asks for the maximum number of subsums of vectors $v_1, \ldots, v_n \in \mathbb{R}^d$ of length at least 1 that fall into a ball radius 1. Denote this number by $f(n)$.

Littlewood and Offord [32] proved that in the cases $d = 1$ and 2 we have the upper bound

$$f(n) \leq 2^n \log(n)/\sqrt{n}.$$

Erdős [17] showed that the best upper bound in the case $d = 1$ and an interval of length 2 is

$$\binom{n}{\lceil \frac{n}{2} \rceil}$$

and for any interval of length $2k$ the optimal upper bound is provided by the sum of the k largest binomial coefficients in n. We shall henceforth denote this sum by $f(n,k)$.

The 1-dimensional problem has a very natural probabilistic formulation - that is how it actually appears in Erdős's work.

Theorem 1.0.1. *Consider n independent random variables* ε_i *such that* $\mathbb{P}(\varepsilon_i = \pm 1) = 1/2$ *and let* $|a_i| \geq 1$. *Then for all* $x \in \mathbb{R}$

$$\mathbb{P}(a_1\varepsilon_1 + \cdots + a_n\varepsilon_n \in (x-k, x+k]) \leq \mathbb{P}(\varepsilon_1 + \cdots + \varepsilon_n \in (-k, k]).$$

Denote the set $\{1, 2, \ldots, n\}$ by $[n]$. We call any collection of k sets $F_1, \ldots, F_k$ a chain of length k if these sets are ordered by inclusion. That is, if we have $F_1 \subset F_2 \subset \ldots \subset F_k$. Let us call a family of sets k-Sperner if it has no chains longer than k.

Sperner's Theorem is a classical result in finite set combinatorics that tells us that the largest 1-Sperner family of subsets of $[n]$ cannot have more elements than the middle layer. That is, it cannot have more sets than the family of all sets of cardinality $\lceil \frac{n}{2} \rceil$.

The main ingredient of Erdős's proof of Theorem 1.0.1 was the following extension of Sperner's Theorem.

Theorem 1.0.2. *Let $\mathcal{F}$ be a k-Sperner family of subsets of $[n]$. Then*

$$|\mathcal{F}| \leq f(n,k) = \sum_{j=\lfloor \frac{n-k+1}{2} \rfloor}^{\lfloor \frac{n+k-1}{2} \rfloor} \binom{n}{j}.$$

Although the two Theorems stated may appear unrelated, Erdős linked them by giving a short proof of Theorem 1.0.1 using Theorem 1.0.2. Let us give this proof to highlight the link between random sums and k-Sperner families as we shall use a similar idea in Section 2.4 of Chapter 2.

Erdős's Proof of Theorem 1.0.1. There is a natural correspondence between random sums $S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$ that fall into the interval $(x-k, x+k]$ and subsets of $[n]$. Namely, for each realization of S_n that falls into $(x-k, x+k]$ assign a set $A = \{i : \varepsilon_i = 1\}$. Denote the collection of all such sets by $\mathcal{F}_x$. The probability in question is just the proportion of $\mathcal{F}_x$ in the powerset of $[n]$.

Let us verify that $\mathcal{F}_x$ is k-Sperner. Notice that if $\mathcal{F}_x$ has a chain of length $k+1$ then there exists two sets A, B in $\mathcal{F}_x$ that differ in at least $k+1$ elements. But then the linear combinations corresponding to these sets differ by at least $2k+2$ and so both sets cannot lie in $\mathcal{F}_x$, which is a contradiction. Therefore $\mathcal{F}_x$ is k-Sperner. Using Theorem 1.0.2 we obtain

$$\begin{aligned} \mathbb{P}\left(a_1\varepsilon_1 + \cdots + a_n\varepsilon_n \in (x-k, x+k]\right) &= |\mathcal{F}_x|/2^n \leq f(n,k)/2^n \\ &= \mathbb{P}\left(\varepsilon_1 + \cdots + \varepsilon_n \in (-k, k]\right). \end{aligned}$$

■

We shall now present the new proof of Theorem 1.0.1 without using Theorem 1.0.2. Then we shall move to the new elementary proof of Theorem 1.0.2.

Proof of Theorem 1.0.1 Let us write $S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$ and $W_n = \varepsilon_1 + \cdots + \varepsilon_n$. We can assume that $a_1 \geq a_2 \geq \ldots \geq a_n \geq 1$. Without loss of generality we can also take $a_n = 1$. This is because

$$\begin{aligned}\mathbb{P}\left(S_n \in (x-k, x+k]\right) &\leq \mathbb{P}\left(S_n/a_n \in (x-k, x+k]/a_n\right) \\ &\leq \sup_{x\in\mathbb{R}} \mathbb{P}\left(S_n/a_n \in (x-k, x+k]\right).\end{aligned}$$

We shall argue by induction on n. The claim is trivial for $n = 0$. Let us assume that we have proved the statement for $1, 2, ..., n-1$. Then taking the expectation with respect to ε_n we obtain

$$\begin{aligned}&\mathbb{P}\left(S_n \in (x-k, x+k]\right) \\ =&\tfrac{1}{2}\mathbb{P}\left(S_{n-1} \in (x-k-1, x+k-1]\right) + \tfrac{1}{2}\mathbb{P}\left(S_{n-1} \in (x-k+1, x+k+1]\right) \\ =&\tfrac{1}{2}\mathbb{P}\left(S_{n-1} \in (x-k-1, x+k+1]\right) + \tfrac{1}{2}\mathbb{P}\left(S_{n-1} \in (x-k+1, x+k-1]\right) \\ \leq&\tfrac{1}{2}\mathbb{P}\left(W_{n-1} \in (-k-1, k+1]\right) + \tfrac{1}{2}\mathbb{P}\left(W_{n-1} \in (-k+1, k-1]\right) \\ =&\tfrac{1}{2}\mathbb{P}\left(W_{n-1} \in (-k-1, k-1]\right) + \tfrac{1}{2}\mathbb{P}\left(W_{n-1} \in (-k+1, k+1]\right) \\ =&\mathbb{P}\left(W_n \in (-k, k]\right),\end{aligned}$$

which completes the proof.■

Before we prove Theorem 1.0.2 let us establish a simple and well-known recurrence relation for $f(n,k)$. We shall adopt the convention that for $k \geq n$ we have $f(n,k) = 2^n$ and for $k = 0$ we set $f(n,k) = 0$ to deal with boundary cases.

Lemma 1.0.3. *For* $1 \leq k \leq n$ *we have*

$$f(n,k) = f(n-1,k-1) + f(n-1,k+1).$$

Proof of Lemma 1.0.3 The assertion is trivial for $n = 1,2$. For $n > 2$ using Pascal's identity and grouping terms we have

$$\begin{aligned} f(n-1,k-1) + f(n-1,k+1) &= \sum_{j=\lfloor \frac{n-k+1}{2} \rfloor}^{\lfloor \frac{n+k-3}{2} \rfloor} \binom{n-1}{j} + \sum_{j=\lfloor \frac{n-k-1}{2} \rfloor}^{\lfloor \frac{n+k-1}{2} \rfloor} \binom{n-1}{j} \\ &= \sum_{j=\lfloor \frac{n-k+1}{2} \rfloor}^{\lfloor \frac{n+k-1}{2} \rfloor} \left(\binom{n-1}{j-1} + \binom{n-1}{j} \right) \\ &= \sum_{i=\lfloor \frac{n-k+1}{2} \rfloor}^{\lfloor \frac{n+k-1}{2} \rfloor} \binom{n}{j} = f(n,k). \end{aligned}$$

■

The main idea of the new proof of Theorem 1.0.2 is in showing that a k-Sperner family on the ground set $[n]$ can be partitioned into two parts so that if we removed the element n from all sets we would arrive at two families on the ground set $[n-1]$ where one is $(k-1)$-Sperner and the other one is $(k+1)$-Sperner.

Proof of Theorem 1.0.2 The result is trivial when $n = 1$ for all values of k. We thus assume that $n > 1$ and that the assertion is true for all integers smaller than n.

Let $\mathcal{F}$ be a k-Sperner family on the ground set $[n]$. Define $\mathcal{F}_n = \{A \in \mathcal{F} : n \in A\}$ and $\mathcal{F}_n^c = \mathcal{F} / \mathcal{F}_n$. Consider all elements of $\mathcal{F}_n$ that are on the bottom of some chain of k elements. Remove the element n from these sets and move the resulting sets to to $\mathcal{F}_n^c$ and denote the resulting collection by $\mathcal{G}$. Note that the sets we moved cannot coincide with any sets in $\mathcal{F}_n^c$ as if such a set existed we could create a chain of length $k+1$ in $\mathcal{F}$. Also, remove the element n from all remaining sets in $\mathcal{F}_n$ and denote the resulting family by $\mathcal{H}$.

Both families $\mathcal{G}$ and $\mathcal{H}$ are now defined on the groundset $[n-1]$ and the total number

of elements in both of them is exactly $|\mathcal{F}|$. Note that $\mathcal{H}$ cannot have any chains of length k as we removed one element from each such chain. Thus $\mathcal{H}$ is $(k-1)$-Sperner.

Furthermore, all sets that we moved are incomparable and so we could not have prolonged the chains in $\mathcal{F}_n^c$ by more than 1 when we added new elements to it to form $\mathcal{G}$. Thus $\mathcal{G}$ is $(k+1)$-Sperner. Using the induction hypothesis and Lemma 1.0.3 we have

$$|\mathcal{F}| = |\mathcal{H}| + |\mathcal{G}| \leq f(n-1,k-1) + f(n-1,k+1) = f(n,k).$$

■

CHAPTER 2

KOLMOGOROV'S INEQUALITY AND A QUESTION OF LEADER AND RADCFLIFFE

2.1 Introduction and the main result

The Levý concentration function of a real-valued random variable X is defined by

$$Q(X,\lambda) = \sup_{x\in\mathbb{R}} \mathbb{P}\left(X \in (x, x+\lambda]\right), \quad \lambda \geq 0.$$

Of special interest is the investigation of $Q(S_n,\lambda)$, where $S_n = X_1 + \cdots + X_n$ is a sum of independent random variables. The first inequality relating $Q(S_n,\lambda)$ to individual concentration functions $Q(X_i,\lambda)$ was proved by Kolmogorov [30]. Let us state a refined version of the latter inequality by Rogozin [35].

Theorem 2.1.1. *For $L \geq \lambda$ we have*

$$Q(S_n, L) \leq CL \left(\sum_{i=1}^{n} (1 - Q(X_i,\lambda)) \right)^{-1/2},$$

where C is an absolute constant.

Many generalizations and sharpenings of Theorem 2.1.1 were established by a number authors. These include the work of Esseen [18], Kesten [26] and Halász [19] among others.

In his celebrated paper Erdős [17], using Sperner's Theorem, provided the first exact result which is nowadays usually referred to as the Littlewood-Offord problem. Namely, for linear combinations $S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$ of independent random variables ε_i with $\mathbb{P}(\varepsilon_i = \pm 1)$ and $a_i \geq 1$ he showed that

$$\mathbb{P}\left(S_n \in (x-k, x+k]\right) \leq \mathbb{P}\left(\varepsilon_1 + \cdots + \varepsilon_n \in (-k,k]\right), \quad k \in \mathbb{N}.$$

Erdős conjectured that an analogous result holds with coefficients a_i replaced by vectors in any Banach space, which was confirmed by Kleitman [29].

The condition on the variables to be two point valued in Erdös's result was later removed by Leader and Radcliffe [31]. To be more precise, let us state their result.

Theorem 2.1.2. *Let $X_1, \dots, X_n$ be independent random variables satisfying $Q(X_i, 2) = 1/k$ for some $k \in \mathbb{N}$. Then we have*

$$\mathbb{P}\left(S_n \in (x-1, x+1]\right) \leq \mathbb{P}\left(U_1 + \cdots + U_n \in (-1, 1]\right),$$

where U_i are independent and uniformly distributed in the k point set

$$\{-k+1, -k+3, \dots, k-1\}.$$

Note that the case $k = 2$ corresponds exactly to the Littlewood-Offord problem. Leader and Radcliffe asked the question about what happens in the case when $Q(X_i, 2)$ is not of the form $1/k$. The main aim of this chapter is to prove an inequality in the spirit of Theorem 2.1.2 that deals with arbitrary values for the concentration functions $Q(X_i, \lambda) = \alpha_i$ and all lengths of intervals of concentration (not just $\lambda = 2$).

Before stating our result let us first adopt some notation. We shall denote by $\mathcal{L}(X)$ the law of X, that is, its probability distribution. Furthermore, let us denote by ν^k the uniform distribution on $\{-k+1, -k+3, \dots, k-1\}$. That is, we have $\nu^k = \mathcal{L}(U_1)$, where U_1 is as in 2.1.2. Furthermore, we shall write $Q(X)$ for $Q(X, 2)$.

Theorem 2.1.3. *Let $S_n = X_1 + \cdots + X_n$ be the sum of independent random variables X_i such that $Q(X_i) = \alpha_i$ and consider the integers k_i so that $\frac{1}{k_i+1} < \alpha_i \leq \frac{1}{k_i}$. Then for all $k \in \mathbb{N}$ we have*

$$\mathbb{P}\left(S_n \in (x-k, x+k]\right) \leq \mathbb{P}\left(T_1 + \cdots + T_n \in (-k, k]\right), \tag{1}$$

where T_i has the distribution $\mathcal{L}(T_i) = (1-\tau_i)\nu^{k_i+1} + \tau_i \nu^{k_i}$ and $\tau_i = k_i(k_i+1)\alpha_i - k_i$.

Note that the distribution of T_i is a convex combination of two uniform distributions. It is easy to see that in the case $k = 1$ and $\alpha_i = 1/k$ Theorem 2.1.3 reduces nicely to Theorem 2.1.2. Indeed, in this case $\tau_i = 1$ and so $\mathcal{L}(T_i) = \mathcal{L}(U_i)$.

The outline of the proof of Theorem 2.1.3 is as follows. Firstly, we narrow down the set of distributions under consideration. This is done by characterizing the extreme points of distributions with a condition on the concentration function. Secondly, by the use of a Sperner-type theorem for multisets we extend the result of Leader and Radcliffe to all intervals and with the condition $Q(X_i, 2) = 1/k$ replaced by $Q(X_i, 2) = 1/k_i$. Finally, we reduce Theorem 2.1.3 to this extension.

2.2 Reduction to discrete random variables

Before we proceed with the proof of Theorem 2.1.3 let us make a simple reduction. Firstly, we want to show that the all random variables X_i in Theorem 2.1.3 can be assumed take finitely many values. This is due to a well known fact in real analysis - we can approximate any bounded measurable function by step functions, giving us the required discretization. Let us be more precise. Consider a random variable X with distribution function $F(t) = \mathbb{P}(X \leq t)$. For all $n \in \mathbb{N}$ and $k = 0, 1, \ldots, n$ define the level sets $A_k = \left\{t : \ F(t) \in \left(\frac{k-1}{n}, \frac{k}{n}\right]\right\}$. The sets A_k are intervals (possibly infinite) as F is monotone. Furthermore, we define the sequence of functions F_m by setting

$$F_n(t) = \sum_{k=0}^{n} \frac{k}{n} \mathbb{1}_{A_k}.$$

Each function F_m is a distribution function since

$$\lim_{t \to \infty} F_m(t) = 1 \quad \text{and} \quad \lim_{t \to -\infty} F_m(t) = 0.$$

Consider the corresponding sequence of random variables $X^{(m)}$ with distribution function F_m. Since F_m is a step function with differences between consecutive steps $\frac{1}{m}$ it

follows that $X^{(m)}$ have a uniform distribution on a finite set. Furthermore, by the definition of the sequence F_m we have that for all $t \in \mathbb{R}$

$$|F(t) - F_n(t)| \leq \frac{1}{m}.$$

It follows that not only does $X^{(m)}$ converge to X weakly as $m \to \infty$, but also that the convergence is uniform. It immediately follows that

$$\begin{aligned} \left|C_h(X) - C_h(X^{(m)})\right| &\leq \sup_t |(F(t+h) - F(t)) - (F_n(t+h) - F_n(t))| \\ &\leq \sup_t |F(t+h) - F(t)| + \sup_t |F(t+h) - F(t)| \\ &\leq \frac{2}{m}. \end{aligned}$$

For a sum of independent random variables $S = X_1 + \cdots + X_n$ associate a corresponding sum $S_n^{(m)} = X_1^{(m)} + \cdots + X_n^{(m)}$, where $X_i^{(m)}$ are independent discretized versions of X_i as described above. It is a standard result in probability that for fixed n the sequence $S_n^{(m)}$ converges weakly to S_n as $m \to \infty$.

Tucker [38] showed that weak convergence of random variables implies the convergence of the concentration functions and therefore we can arbitrarily well approximate $Q(X_1 + \cdots + X_n)$ in Theorem 2.1.3 by the discretized sums.

We have to also discuss one more detail. After the discretization of a random variable X we may slightly alter $C_h(X)$. This effect turns out to be negligible in the context of Theorem 2.1.3. Indeed, notice that the upper bound in the theorem is continuous with respect to the values α_i. This can be easily seen by taking the expectation with respect to T_i - it then becomes a linear function of α_i.

In view of what we have just established, we shall henceforth assume that all variables under consideration take only finitely many values with probabilities that are themselves rational numbers.

2.3 Extremal distributions

For $k \geq 1$ define by μ^k a uniform distribution on some k points in $\mathbb{R}$ that are pairwise at distance at least 2. Note that the definition of μ^k depends on the choice of those points, which is not reflected in the notation. Usually we will supply μ^k with a subscript, which will mean that the distributions with distinct subscripts might be concentrated in different sets. When the set of k points will be $\{-k+1, -k+3, \ldots, k-3, k-1\}$, we are going to use the notation ν^k instead of μ^k.

Lemma 2.3.1. *Let X be a real valued random variable such that $Q(X) = m/n \in (1/(k+1), 1/k]$. Assume that X takes the values in the set $S = \{y_1, \ldots, y_m\}$ with with probabilities $\mathbb{P}(X = y_i) = m_i/n_i$. Let us define*

$$N = n\prod_i n_i, \quad K = (n-km)\prod_i n_i, \quad L = ((k+1)m-n)\prod_i n_i.$$

Then we can express the distribution of X as

$$\mathcal{L}(X) = \frac{1-\tau}{K}\sum_{l=1}^{K}\mu_l^{k+1} + \frac{\tau}{L}\sum_{l=K+1}^{K+L}\mu_l^k, \tag{2}$$

where $\tau = k(k+1)m/n - k$.

Proof. Assume that $y_1 \leq \ldots \leq y_M$. We can regard the distribution of X as the uniform distribution on a multiset S', where S' is obtained from S by taking the element y_i exactly $nm_i\prod_{j\neq i} n_i$ times. Let $x_1, \ldots, x_N$ be the elements of S' in increasing order.
The condition $Q(X) = m/n$ ensures than no more than $d = Nm/n$ points lie in the interval $(x, x+2]$ for all x. Thus the points x_l, x_{l+d} are at distance at least 2. For $l \leq L$ the points $x_l, x_{l+d}, \ldots, x_{l+kd}$ are pairwise at distance at least two. Each point has mass $1/N$, so in order make the measure on the latter set of points into a probability measure we must

divide it by it by $(k+1)/N$. We have

$$(k+1)/N = (k+1)(n-km)/(nK) = (1-(k(k+1)m/n-k))/K = (1-\tau)/K,$$

thus obtaining the first K distributions μ_l^{k+1} with the desired weights.

For $K+1 \leq l \leq K+L$ take the points $x_l, x_{l+d}, \ldots, x_{l+(k-1)d}$ and the measures concentrated on those points will give us the required L measures μ_l^k. It can be checked that the proportion is again correct, but that will follow from the fact that we used up all points from S' and took each of them only once. Indeed, we started constructing each measure in the representation from a different point in $x_1, \ldots, x_{K+L}$ and then added points with equally spaced indices. Thus we did not use any point twice. Furthermore, $K(k+1)+Lk = N$ and so we used them all.

■

2.4 Proof of a Sperner-type Theorem for multisets

In this section we shall be dealing with multisets defined on the ground set $[n]$ such that each element has an upper bound, say k_i, on its multiplicity. The case $k_i = 1$ naturally reduces to the study of sets. In the latter case we can switch between talking about the powerset of $[n]$ to the study of indicator vectors in $\{0,1\}^n$ with set inclusion corresponding to the product order in $\{0,1\}^n$.

Analogously, we shall view multisets as vectors in the discrete rectangle $L(k_1, \ldots, k_n) = \{0, \ldots, k_1 - 1\} \times \cdots \times \{0, \ldots, k_n - 1\}$ by associating with a multiset the vector of multiplicities of each element in it.

For a vector $x \in \mathbb{R}^n$ we shall denote its i-th coordinate by x_i. We shall endow $L(k_1, \ldots, k_n)$ with the product order. That is, $v \leq w$ if and only if $v_i \leq w_i$. Multiset inclusion corresponds to this order as in the case with sets.

We shall call a collection of vectors $v_1, \ldots, v_k$ a **chain** if $v_1 \leq \cdots \leq v_k$ and refer to the

number k as its length. We say that a family of vectors $\mathcal{F}$ is k-**Sperner** if it has no chains of length $k+1$. In the case $k=1$ we shall say that $\mathcal{F}$ is an antichain rather than 1-Sperner.

Let us partition $L(k_1,\ldots,k_n)$ into classes L_i where

$$L_i = \{x \in L(k_1,\ldots,k_n) \,|\, x_1+\cdots+x_n = i\}.$$

Note that $|L_i|$ is a symmetric sequence in the sense that $|L_i| = |L_{N-i}|$ where $N = \sum(k_i-1)$. The sequence $|L_i|$ is non-decreasing for $i \le \lfloor \frac{N}{2} \rfloor$ and thus, by symmetry, it is non-increasing for $i \ge \lceil \frac{N}{2} \rceil$.

For $k \le k_1+\cdots+k_n+1$ write $f(k_1,k_2,\ldots,k_n,k)$ for the sum of the k largest sets L_i. These are just the k middle diagonals of the rectangle $L(k_1,\ldots,k_n)$.

In Chapter 1 we presented Erdős's proof of the Littlewood-Offord problem that used a Sperner type theorem. We shall need a similar result for multiset k-Sperner families.

Theorem 2.4.1. *Let $\mathcal{F}$ be a k-Sperner family of vectors in $L(k_1,\ldots,k_n)$. Then*

$$|\mathcal{F}| \le f(m_1,m_2,\ldots,m_n,k).$$

Before we proceed with the proof, let us state an inequality for antichains of multisets that will be instrumental in proving Theorem 2.4.1.

Lemma 2.4.2. *Let $\mathcal{F}$ be an antichain in $L(k_1,\ldots,k_n)$. For $0 \le i \le \sum_{j=1}^{n}(k_j-1)$ denote $\mathcal{F}_i = \mathcal{F} \cap L_i$. We have*

$$\sum_i \frac{|\mathcal{F}_i|}{|L_i|} \le 1.$$

The proof of Lemma 2.4.2 can be found in Chapter 10 of the book by I. Anderson [2]. In the case of sets Lemma 2.4.2 is known as the LYM inequality.

Proof of Theorem 2.4.1. Let $\mathcal{F}$ be a k-Sperner family. It is easy to see that $\mathcal{F}$ is a union of k antichains. Indeed, the maximal elements of $\mathcal{F}$ form an antichain and the remaining elements form a $(k-1)$-Sperner family and so the observation follows by induction on k.

Let $\mathcal{A}$ be one of the k antichains that decompose $\mathcal{F}$.

Using Lemma 2.4.2 we obtain

$$\sum_i \frac{|\mathcal{A}_i|}{|L_i|} \leq 1.$$

Summing this inequality over all k antichains we obtain

$$\sum_i \frac{|\mathcal{F}_i|}{|L_i|} \leq k. \tag{3}$$

For families of vectors of fixed cardinality the sum in (3) is minimized by families containing vectors with coordinate sums as close to $\sum_i (k_i - 1)/2$ as possible. This is because in view of (3) the vectors are assigned the smallest weight.

Suppose now that $|\mathcal{F}| > f(k_1, \ldots, k_n, k)$. Note for the family of vectors consisting of the middle k diagonals of $L(k_1, \ldots, k_n)$ the corresponding sum in (3) is exactly equal to 1 and is minimal among all families having $f(k_1, \ldots, k_n, k)$ vectors. Therefore for any family of vectors with more elements the corresponding sum in (3) is strictly greater than 1, which is a contradiction. Thus $|\mathcal{F}| \leq f(k_1, \ldots, k_n, k)$ and we are done. ■

2.5 Linearity of the problem and the case $Q(X_i) = 1/k_i$

Before we proceed to the proof of Theorem 2.1.3, we need establish two facts. Firstly, we show our problem is linear in each measure and so we will always be able to assume that the maximum is attained by an extreme point. Secondly, we extend the result of Leader and Radcliffe to the case where instead of the uniform condition $Q(X_i) = 1/k$ we have $Q(X_i) = 1/k_i$ and all possible interval lengths.

Lemma 2.5.1. *Let $X_1, \ldots, X_n$ be independent random variables with $\mathcal{L}(X_i) = \eta_i$. Furthermore, assume that each distribution η_i can be written as a convex combination of some collection of distributions, say $\{\gamma_{i,j} : j = 1, \ldots, K\}$ for some integer K. That is, for*

each i the exist non-negative numbers $\alpha_{i,1},\dots,\alpha_{i,K}$ such that

$$\eta_i = \sum_{j=1}^{K} \alpha_{i,j}\gamma_{i,j} \quad \textit{and} \quad \sum_{j=1}^{K} \alpha_{i,j} = 1.$$

Then for any measurable function $f:\mathbb{R}^n \to \mathbb{R}$ we have

$$\mathbb{E}f(X_1,\dots,X_n) \le \mathbb{E}f(Y_1,\dots,Y_n),$$

where the random variables Y_i are independent and for each i there is some j such that $\mathcal{L}(Y_i) = \gamma_{i,j}$.

Proof. First let us proof the assertion in the case $n = 1$. Denote by $Y_{1,j}$ a random variable with $\mathcal{L}(Y_{1,j}) = \gamma_{1,j}$. We have

$$\mathbb{E}f(X_1) = \sum_{j=1}^{L} \alpha_{1,j}\mathbb{E}f(Y_{1,j}) \le \max_{1\le k\le L} \mathbb{E}f(Y_{1,j}). \tag{4}$$

It is not difficult to see now that the general case reduces to the latter case. Indeed,

$$\mathbb{E}f(X_1,\dots,X_n) = \mathbb{E}(\mathbb{E}f(X_1,\dots,X_n)|X_i) = \mathbb{E}g_i(X_i),$$

where the function g_i is the conditional expectation of f given X_i. We can do the same for each coordinate step by step and are done. ■

Lemma 2.5.2. *Let $X_1,\dots,X_n$ be independent random variables such that $Q(X_i,2) = 1/k_i$. We have*

$$\mathbb{P}\left(S_n \in (x-k,x+k]\right) \le \mathbb{P}\left(U_1+\dots+U_n \in (-k,k]\right), \tag{5}$$

where U_i are independent and $\mathcal{L}(U_i) = \nu^{k_i}$.

Proof of Lemma 2.5.2. We can assume that X_i take finitely many values and that all probabilities are rational numbers. In view of Lemma 2.3.1, Lemma 2.5.1 and using the

notation of Section 2.3, we can assume that $\mathcal{L}(X_i) = \mu_i^{k_i}$. For each i let us denote the values X_i takes by $x_{i,1}, \ldots, x_{i,k_i}$. Let us define a family of vectors (or multisets)

$$\mathcal{F} = \left\{ v \in L(k_1, \ldots, k_n) | \sum_{j=1}^{n} x_{j,v_j} \in (-k, k] \right\}.$$

Note that by definition of measures $\mu_i^{k_i}$ the points $x_{i,1}, \ldots, x_{i,k_i}$ are all at distance at least 2 within each other. Therefore if we had a chain of vectors (or multisets) of length $k+1$ then the sums corresponding to the top and bottom vectors (or multisets) would differ by more than $2k$ as so we get a contradiction. Therefore the family $\mathcal{F}$ is k-Sperner.

Using Theorem 2.4.1, we therefore have

$$\begin{aligned} \mathbb{P}(S_n \in (x-k, x+k]) &= |\mathcal{F}| / \prod_{j=1}^{n} k_i \\ &\leq f(k_1, k_2, \ldots, k_n, k) / \prod_{j=1}^{n} k_i \\ &= \mathbb{P}(U_1 + \cdots + U_n \in (-k, k]). \end{aligned}$$

2.6 Proof of Theorem 2.1.3

Let $X_1, \ldots, X_n$ be independent random variables taking finitely many values, say $X_i \in \{y_{i,1}, \ldots, y_{i,A_i}\}$. Assume that $Q_2(X_i) = \alpha_i \in (1/(k_i+1), 1/k_i]$. Consider another sequence $Y_1, \ldots, Y_n$ of independent random variables with sum M_n such that $\mathcal{L}(Y_i) = \tau_i \nu_i^{k_i} + (1-\tau_i)\nu_i^{k_i+1}$, where $\tau_i = \alpha_i k_i (k_i+1) - k_i$.

Without loss of generality we can assume that α_i are rational. This is because the upper bound we want to establish is continuous with respect to α_i and so it is enough to deal with only rational values. Let us write $\alpha_i = \alpha_i = m_i/n_i \in (1/(k_i+1), 1/k_i]$.

Moreover, we can assume that the probabilities $\mathbb{P}\left(X_i = y_{i,k}\right)$ are also rational. Thus $\mathbb{P}\left(X_i = y_{i,k}\right) = m_{i,k}/n_{i,k}$. Writing $N_i = n_i \prod_{j=1}^{n} n_{i,j}$ we can look at the distribution of X_i as a uniform distribution on a multiset with N_i elements. By Lemma 2.3.1 we have

$$\mathcal{L}(X_i) = \frac{1-\tau_i}{K_i}\sum_{l_i=1}^{K_i}\mu_{i,l_i}^{k_i+1} + \frac{\tau_i}{L_i}\sum_{l_i=K_i+1}^{K_i+L_i}\mu_{i,l_i}^{k_i}, \tag{6}$$

where K_i, L_i and τ_i are defined as in Lemma 2.3.1.

We shall expand the product measure $\prod_{i=1}^n \mathcal{L}(X_i)$ into a sum of products of the measures $\mu_{i,l_i}^{\tilde{k}_i}$, where $\tilde{k}_i = k_i + 1$ for $l_i \leq K_i$ and $\tilde{k}_i = k_i$ otherwise. For the same ranges of l_i define $\tilde{\tau}_i$ in a natural way - the coefficient in front of $\mu_{i,l_i}^{\tilde{k}_i}$. Then using Lemma 2.5.2 term by term we obtain

$$\begin{aligned}
\mathbb{P}(S_n \in (-k,k]) &= \prod_{i=1}^n \mathcal{L}(X_i)((x-k,x+k]) \\
&= \prod_{i=1}^n \left(\frac{1-\tau_i}{K_i}\sum_{l_i=1}^{K_i}\mu_{i,l_i}^{k_i} + \frac{\tau_i}{L_i}\sum_{l=K_i+1}^{K_i+L_i}\mu_{i,l_i}^{k_i+1}\right)(B_k) \\
&= \prod_{i=1}^n \left(\tilde{\tau}_i \sum_{l=1}^{K_i+L_i}\mu_{i,l}^{\tilde{k}_i}\right)(B_k) \\
&= \sum_{l_1,\ldots,l_n}\prod_{i=1}^n \tilde{\tau}_i \mu_{i,l}^{\tilde{k}_i}(B_k) \\
&\leq \sum_{l_1,\ldots,l_n}\prod_{i=1}^n \tilde{\tau}_i \nu_{i,l}^{\tilde{k}_i}((-k,k]) \\
&= \prod_{i=1}^n \left(\tilde{\tau}_i \sum_{l=1}^{K_i+L_i}\nu_{i,l}^{\tilde{k}_i}\right)((-k,k]) \\
&= \prod_{i=1}^n \left(\frac{1-\tau_i}{K_i}\sum_{l_i=1}^{K_i}\nu_{i,l_i}^{k_i} + \frac{\tau_i}{L_i}\sum_{l=K_i+1}^{K_i+L_i}\nu_{i,l_i}^{k_i+1}\right)((-k,k]) \\
&= \prod_{i=1}^n (\tau_i \nu_i^{k_i} + (1-\tau_i)\nu_i^{k_i+1})((-k,k]) \\
&= \mathbb{P}(M_n \in (-k,k]).
\end{aligned}$$

■

Note that once we expand the product measure $\prod_{i=1}^n \mathcal{L}(X_i)$, use Lemma 2.5.2 term by term and group similar terms we obtain exactly the expansion of $\prod_{i=1}^n \mathcal{L}(Y_i)$.

CHAPTER 3

SMALL BALL PROBABILITIES FOR SUMS OF RANDOM VECTORS WITH BOUNDED DENSITY

Let μ be the Lebesgue measure and X a random vector in $\mathbb{R}^d$. If X has a density, say p, we define

$$M(X) = ||p||_\infty = \operatorname{ess\,sup} p = \sup\{\varepsilon : \mu(\{t : p(t) > \varepsilon\}) > 0\}.$$

For random variables with distributions that are not absolutely continuous with respect to μ we set $M(X) = \infty$. All of our density functions will be taken as equivalence classes up to alterations on sets of measure 0; that is, they are defined as elements of L_∞.

The aim of this chapter is to provide best possible upper bounds for the maximum density and small ball probabilities of sums of random vectors.

Our starting point is a result by Rogozin, who showed that in the case $d = 1$ the worst case is provided by uniform distributions over intervals. To be more precise, it was proved in [36] that for independent real-valued random variables $X_1, \ldots, X_n$ with $M(X_i) \leq M_i$ we have

$$M(X_1 + \cdots + X_n) \leq M(U_1 + \cdots + U_n),$$

where U_k are independent and uniformly distributed in $[-\frac{1}{2M_i}, \frac{1}{2M_i}]$.

We extend Rogozin's inequality to all dimensions. In fact, we prove a more general statement for small ball probabilities that almost instantly implies the latter .

Theorem 3.0.1. *Let $X_1, \ldots, X_n$ be independent random vectors in $\mathbb{R}^d$ with $M(X_i) \leq K_i$. Consider a collection of independent random vectors $U_1, \ldots, U_n$ with densities equal to K_i on a centered ball and* 0 *elsewhere. Then for every measurable set S we have*

$$\mathbb{P}(X_1 + \cdots + X_n \in S) \leq \mathbb{P}(U_1 + \cdots + U_n \in B), \tag{7}$$

where B is the centered ball such that $\mu(B) = \mu(S)$.

Corollary 3.0.2. *Under the same conditions as above we also have that*

$$M(X_1 + \cdots + X_n) \leq M(U_1 + \cdots + U_n).$$

Proof of Corollary 3.0.2 Note that for any variable X with density p

$$M(X) = \lim_{\varepsilon \to 0} \sup_{\mu(S)=\varepsilon} \varepsilon^{-1} \int_S p d\mu.$$

Let B^ε be the centered ball with volume ε. Using Theorem 3.0.1 we obtain

$$\begin{aligned} M(X_1 + \cdots + X_n) &= \lim_{\varepsilon \to 0} \sup_{\mu(S)=\varepsilon} \varepsilon^{-1} \mathbb{P}(X_1 + \cdots + X_n \in S) \\ &\leq \lim_{\varepsilon \to 0} \varepsilon^{-1} \mathbb{P}(U_1 + \cdots + U_n \in B^\varepsilon) \\ &= M(U_1 + \cdots + U_n). \end{aligned}$$

Hence the corollary holds. ■

Even for $d = 1$ our approach is quite different than that of Rogozin, who used discretization arguments together with an idea of Erdős to relate small ball probabilities to Sperner's theorem in finite set combinatorics. We avoid these subtleties by using a rearrangement inequality proved by Brascamp, Lieb and Luttinger.

3.1 Rearrangements of functions

For non-negative functions $f : \mathbb{R}^d \mapsto \mathbb{R}$ set $M_y^f = \mu\{t : f(t) \geq y\}$. Assume that $M_a^f < \infty$ for some $a \in \mathbb{R}$. Define $\tilde{f}$ to be a function such that:

$$1)\ \tilde{f}(x) = \tilde{f}(y), \text{ for } |x|_2 = |y|_2;$$

$$2)\ f(x) \leq f(y) \text{ for } x \leq y;$$

$$3)\ M_y^{\tilde{f}} = M_y^f.$$

The function $\tilde{f}$ is known as the spherically symmetric decreasing rearrangement of f. For existence, uniqueness and other properties of $\tilde{f}$ we refer the reader to [12] and [20] (Chapter X).

Having introduced the relevant symmetrization we can state the aforementioned rearrangement result.

Theorem 3.1.1. *Let f_j, $1 \leq j \leq k$ be non-negative measurable functions on $\mathbb{R}^d$ and let $a_{j,m}$, $1 \leq j \leq k, 1 \leq m \leq n$, be real numbers. Then*

$$\int_{\mathbb{R}^{nd}} \prod_{j=1}^{k} \left(f_j \left(\sum_{m=1}^{n} a_{j,m} x_m \right) \right) d^{nd}x \leq \int_{\mathbb{R}^{nd}} \prod_{j=1}^{k} \left(\tilde{f}_j \left(\sum_{m=1}^{n} a_{j,m} x_m \right) \right) d^{nd}x$$

A direct consequence of the latter result is the following.

Theorem 3.1.2. *Let $X_1, \ldots, X_n$ be independent random variables with given density functions p_i. Consider another collection of independent random variables $X_1', \ldots, X_n'$ with density functions $\tilde{p}_i$. Then for every measurable set S we have*

$$\mathbb{P}\left(X_1 + \cdots + X_n \in S\right) \leq \mathbb{P}\left(X_1' + \cdots + X_n' \in B\right), \tag{8}$$

where B is the centered ball such that $\mu(B) = \mu(S)$.

Proof. We have that

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \in S\right) = \int_{x_1,\ldots x_n} \prod_{i=1}^{n} p_i(x_i) \mathbb{1}_S \left(\sum_i x_i\right) d^n\mu.$$

Now apply Theorem 3.1.1 with the f_i taken to be $\{p_1, \ldots, p_n, \mathbb{1}_S\}$ and the $a_{j,m} = 1$ when $j = m$ or $j = n+1$ and $a_{j,m} = 0$ otherwise. We note that $\widetilde{\mathbb{1}_S} = \mathbb{1}_B$ and that $\tilde{p}_i$ are the densities of X_i', completing the proof. ■

3.2 Extremal measures

Let X be a normed space. Given a set $A \subset X$ say that a point $x \in A$ is an extremal point of A if x does not lie in the interior of any line segment within A. In other words, for all $y, z \in A$ we have that if $x \in \{(1-\lambda)y + \lambda z \,|\, \lambda \in [0,1]\}$, then either $x = y$ or $x = z$.

The aim of this section is to characterize the extreme points of the set of measures with bounded densities. The reason for doing this is that we want to narrow down the class of measures under consideration. The well-known Krein-Milman theorem tells us that in a normed space every convex compact set is equal to the closure of convex hull of its extreme points. This allows us to draw the conclusion that any linear function of a convex compact set is maximized by an extremal point. Unfortunately, in our case the set of measures under consideration will not be compact and so we cannot use this theorem. Fortunately, we shall be able to show that considering extreme points is sufficient. This will be done in the last section of this chapter.

It turns out that in the end we shall never actually use the characterization. We will only need some ideas from its proof. We shall nevertheless provide the characterization as we consider this to be of independent interest and possibly useful in future investigations.

Lemma 3.2.1. *Denote by $\mathcal{S}_K$ be the set of probability measures in $\mathbb{R}^d$ that have densities with essential suprema bounded by $K > 0$. The extreme points of $\mathcal{S}_K$ are measures having densities $p(t) = K\mathbb{1}_S(t)$ for some set S with $\mu(S) = 1/K$.*

Proof. Firstly, we note that all measures having densities $p = K\mathbb{1}_S$ are extremal. Suppose not. Then $p = \alpha p_1 + (1-\alpha)p_2$, where $\alpha \in (0,1)$ and p_1, p_2 are not equal to p. But then p_1 and p_2 differ from p on a set of positive measure, and so $\max(p_1, p_2) > K$ on some set of positive measure. Hence one of p_1, p_2 must exceed K on a set of positive measure, so is outside of $\mathcal{S}_K$.

Suppose that the density of a measure is not one of these extremal examples. Consider the sets

$$A_y = \{t : p(t) \geq y\}.$$

Now, there is some $y \in (0,K)$ such that $\mu(A_y) > 0$, as otherwise $p(t) = K$ almost everywhere on its support, and so p would be one of our extremal examples. We fix any such y, and define $X = \sup(p) \backslash A_y$. Furthermore, we partition X into two disjoint sets X_0, X_1 such that $\int_{X_0} p d\mu = \int_{X_1} p d\mu$.

We fix $\delta \in (0, K/y - 1) \cap (0,1)$, and construct two densities p_1, p_2 as follows:

$$p_i(t) = \begin{cases} p(t) & t \in A_y \\ (1-\delta)p(t) & t \in X_i \\ (1+\delta)p(t) & t \in X_{1-i} \end{cases}$$

First, we observe that $p = \frac{1}{2}(p_1 + p_2)$. Furthermore, each of p_1, p_2 are equal to p on A_y, and are bounded pointwise on X by:

$$(1+\delta)\sup_X p \leq (1+\delta)y \leq K.$$

Hence the essential suprema of p_1, p_2 are bounded by K, and so $p_1, p_2 \in \mathcal{S}_K$ as required. ■

3.3 Proof of the main Theorem

Combining the results section 3.1, the proof of Lemma 3.2.1 and Theorem 3.1.2 we will easily derive Theorem 3.0.1. In this section we shall view all densities as elements of $L_1(\mathbb{R})$ instead of $L_\infty(\mathbb{R})$.

Proof. We first observe that Equation 7 for each i can be written as

$$\mathbb{P}(X_1 + \cdots + X_n \in S) = \mathbb{E}\left[\mathbb{P}(X_1 + \cdots + X_n \in S) \,|\, X_i\right]. \tag{9}$$

Therefore the probability in question in linear with respect to each distribution and so also with respect to each density function. We shall now show that $\mathbb{P}(X_1 + \cdots + X_n \in S)$ is maximized when each X_i has a density function from the set $\mathcal{S}_{K_i}$.

We can without loss of generality assume that the densities are simple functions. That is, functions that take only finitely many values. This is because simple functions are dense in $L_1(\mathbb{R})$.

Assume that the density p of the i-th random variable in (9) is a simple function and that it takes values $K = x_1 \geq x_2 \geq \cdots x_t > 0$ and is zero elsewhere. Define

$$A_j = \left\{ t \,\middle|\, p(t) = x_j \right\}.$$

We have that $\mu(A_1) \geq 0$ and $\mu\left(A_j\right) > 0$ for $j > 1$. We shall now proceed quite similarly as in the proof of Lemma 3.2.1. Namely, we shall express this density as a convex combination of two different densities that both lie in $\mathcal{S}_{K_i}$. In view of (9) that we can replace p by one of those densities so that the corresponding expectation does not decrease.

For any $1 \leq j \leq t$ let B_0 and B_1 be a partition of A_j into two parts of equal measure.

We can write

$$p_i(t) = \begin{cases} p(t) & t \in A_j^c \\ (1-\delta)p(t) & t \in B_i \\ (1+\delta)p(t) & t \in B_{1-i}, \end{cases}$$

where $\delta \in [0,1]$ is picked in the following manner. If $x_j \leq \frac{K}{2}$ then $\delta = 1$ and otherwise δ is such that $(1+\delta)x_j = K$.

Note that p_1 and p_2 are indeed densities as by picking δ in the manner above we ensured that they are both positive. Furthermore, we have that $p = \frac{1}{2}p_1 + \frac{1}{2}p_2$.

It is also useful to note that we either by switching p by p_1 or p_2 in (9) we either in both cases decrease the measure of the set where the density is in $(0,K)$ by $\frac{\mu(A_j)}{2}$. After the procedure the other values x_j stay the same for the new density. As we can then perform the same procedure for each $j > 1$ we get that eventually for the final density the measure of the set where this density is in $(0,K)$ halved. We also increase the integral in (9) each time. Using the same procedure repeatedly for the obtained density we get a sequence of densities that converge to some density in $\mathcal{S}_{K_i}$ and we get the increase of the probability in (9) along this sequence. Thus the maximum is attained on an extreme point.

Thus by Lemma 3.2.1 and the reasoning above, the densities of every variable X_i can be assumed to be proportional to the indicator function of some set S_i of measure K_i^{-1} if we attain the maximum. From Theorem 3.1.2, we can infer that in order to maximize this expression in (9) we may replace each of the densities by their spherically decreasing rearrangements and the set S by B. The spherically decreasing rearrangement of an indicator function of a set S_i is the indicator function of the centered ball of the same volume. This means that we are replacing the density of X_i by the density of U_i when optimizing the expression in (9) and we are done.

■

CHAPTER 4

TWO QUESTIONS ON SYMMETRIC RANDOM WALKS

Let $S_n = X_1 + \cdots + X_n$ be a sum of independent random variables X_i such that

$$|X_i| \leq 1 \quad \text{and} \quad \mathbb{E}X_i = 0. \tag{10}$$

Let $W_n = \varepsilon_1 + \cdots + \varepsilon_n$ be the sum of independent Rademacher random variables, i.e., such that $\mathbb{P}(\varepsilon_i = \pm 1) = 1/2$. We will refer to W_n as a simple random walk with n steps.

By a classical result of Hoeffding [22], we have the bound

$$\mathbb{P}(S_n \geq x) \leq \exp\left\{-x^2/2n\right\}, \quad x \in \mathbb{R}. \tag{11}$$

If we take $S_n = W_n$ on the left-hand side of (11), then in view of the Central Limit Theorem we can infer that the exponential function on the right-hand side is the minimal one. Yet a certain factor of order x^{-1} is missing, since $\Phi(x) \approx (\sqrt{2\pi}x)^{-1} \exp\left\{-x^2/2\right\}$ for large x.

Furthermore, it is possible to show that the random variable S_n is sub-Gaussian in the sense that

$$\mathbb{P}(S_n \geq x) \leq c\,\mathbb{P}\left(\sqrt{n}Z \geq x\right), \quad x \in \mathbb{R},$$

where Z is the standard normal random variable and c is some explicit positive constant (see, for instance, [6]).

Perhaps the best upper bound for $\mathbb{P}(S_n \geq x)$ was given by Bentkus [5]. He proved, in particular, that for integer x we have

$$\mathbb{P}(S_n \geq x) \leq 2\mathbb{P}(W_n \geq x - 1). \tag{12}$$

Although there are numerous improvements of the Hoeffding inequality, to our knowledge there are no examples where the exact bound for the tail probability is found.

In this chapter we give an optimal bound for the tail probability $\mathbb{P}(S_n \geq x)$ under the additional assumption of symmetry.

We henceforth reserve the notation S_n and W_n for random walks with symmetric steps satisfying (10) and a simple random walk with n steps respectively.

Theorem 4.0.1. *For $x > 0$ we have*

$$\mathbb{P}(S_n \geq x) \leq \begin{cases} \mathbb{P}(W_n \geq x) & \text{if } \lceil x \rceil + n \in 2\mathbb{Z}, \\ \mathbb{P}(W_{n-1} \geq x) & \text{if } \lceil x \rceil + n \in 2\mathbb{Z}+1. \end{cases} \tag{13}$$

Kwapień proved (see [37]) that for arbitrary i.i.d. symmetric random variables X_i and real numbers a_i with absolute value less than 1 we have

$$\mathbb{P}(a_1 X_1 + \ldots + a_n X_n \geq x) \leq 2\,\mathbb{P}(X_1 + \ldots + X_n \geq x), \quad x > 0.$$

The case $n = 2$ with $X_i = \varepsilon_i$ and $x = 2$ shows that the constant 2 cannot be improved. Theorem 4.0.1 improves Kwapień's inequality for Rademacher sequences.

In this chapter we also consider the problem of finding the quantity

$$\sup_{S_n} \mathbb{P}(S_n = x),$$

which can be viewed as a non-uniform bound for the concentration of the random walk S_n at a point.

Theorem 4.0.2. *For $x > 0$ and $k = \lceil x \rceil$ we have*

$$\mathbb{P}(S_n = x) \leq \mathbb{P}(W_m = k), \tag{14}$$

where

$$m = \begin{cases} \min\left\{n, k^2\right\}, & \text{if } n + k \in 2\mathbb{Z}, \\ \min\left\{n-1, k^2\right\}, & \text{if } n + k \in 2\mathbb{Z}+1. \end{cases}$$

Equality in (14) is attained for $S_n = \frac{x}{k} W_m$.

We provide two different proofs for both inequalities. The first approach is based on induction on the number of random variables (§4.1). To prove Theorem 4.0.2, we also need Theorem 1.0.1.

Interestingly, Theorems 4.0.1 and 4.0.2 can also be proved by applying results from extremal combinatorics (§4.2). Namely, we use the bounds for the size of intersecting families of sets (hypergraphs) by Katona [24] and Milner [33].

Using a strengthening of Katona's result by Kleitman [28], we extend Theorem 4.0.1 to odd 1-Lipschitz functions rather than just sums of the random variables X_i (§4.3). It is important to note that the bound of Theorem 4.0.1 cannot be true for all Lipschitz functions since the extremal case is not provided by odd functions.

4.1 Proofs by induction on dimension

We will first show that it is sufficient to prove Theorems 4.0.1 and 4.0.2 in the case where S_n is a linear combination of independent Rademacher random variables ε_i with coefficients $|a_i| \leq 1$.

Lemma 4.1.1. *Let $g : \mathbb{R}^n \to \mathbb{R}$ be a bounded measurable function. Then we have*

$$\sup_{X_1,\ldots,X_n} \mathbb{E}g(X_1,\ldots,X_n) = \sup_{a_1,\ldots,a_n} \mathbb{E}g(a_1\varepsilon_1,\ldots a_n\varepsilon_n),$$

where the supremum on the left-hand side is taken over symmetric independent random variables $X_1,\ldots,X_n$ such that $|X_i| \leq 1$ and the supremum on the right-hand side is taken over numbers $-1 \leq a_1,\ldots,a_n \leq 1$.

Proof. Define $S = \sup_{a_1,\ldots,a_n} \mathbb{E}g(a_1\varepsilon_1,\ldots a_n\varepsilon_n)$. Clearly

$$S \leq \sup_{X_1,\ldots,X_n} \mathbb{E}g(X_1,\ldots,X_n).$$

By symmetry of $X_1, \ldots, X_n$, we have

$$\mathbb{E}g(X_1, \ldots, X_n) = \mathbb{E}g(X_1\varepsilon_1, \ldots, X_n\varepsilon_n).$$

Therefore

$$\mathbb{E}g(X_1, \ldots, X_n) = \mathbb{E}\mathbb{E}[g(X_1\varepsilon_1, \ldots, X_n\varepsilon_n) \,|\, X_1, \ldots, X_n] \leq \mathbb{E}S = S.$$

■

Thus, in view of Lemma 4.1.1 we will henceforth write S_n for the sum $a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$ instead of a sum of arbitrary symmetric random variables X_i.

Proof of Theorem 4.0.1. First note that the inequality is true for $x \in (0,1]$ and all n. This is due to the fact that $\mathbb{P}(S_n \geq x) \leq 1/2$ by symmetry of S_n and for all n the right-hand side of the inequality is given by the tail of an odd number of random signs, which is exactly $1/2$. We can also assume that the largest coefficient $a_i = 1$ as otherwise if we scale the sum by a_i then the tail of the this new sum would be at least as large as the former. We will thus assume, without loss of generality, that $0 \leq a_1 \leq a_2 \leq \ldots \leq a_n = 1$. Define a function $\mathbb{I}(x,n)$ to be 1 if $\lceil x \rceil + n$ is even, and zero otherwise. Then we can rewrite the right-hand side of (13) as $\mathbb{P}(W_{n-1} + \varepsilon_n \mathbb{I}(x,n) \geq x)$, making a convention that $\varepsilon_0 \equiv 0$.

For $x > 1$ we argue by induction on n. Case $n = 0$ is trivial. Observing that $\mathbb{I}(x-1,n) = \mathbb{I}(x+1,n) = \mathbb{I}(x,n+1)$ we have

$$\begin{aligned}
\mathbb{P}(S_{n+1} \geq x) &= \tfrac{1}{2}\mathbb{P}(S_n \geq x-1) + \tfrac{1}{2}\mathbb{P}(S_n \geq x+1) \\
&\leq \tfrac{1}{2}\mathbb{P}(W_{n-1} + \varepsilon_n \mathbb{I}(x-1,n) \geq x-1) \\
&+ \tfrac{1}{2}\mathbb{P}(W_{n-1} + \varepsilon_n \mathbb{I}(x+1,n) \geq x+1) \\
&= \mathbb{P}(W_n + \varepsilon_{n+1}\mathbb{I}(x,n+1) \geq x).
\end{aligned}$$

■

Before proving Theorem 4.0.2, we will obtain an upper bound for $\mathbb{P}(S_n = x)$ under an additional condition that all a_i are nonzero.

Lemma 4.1.2. *Let $x > 0$, $k = \lceil x \rceil$. Suppose that $0 < a_1 \leq \cdots \leq a_n \leq 1$. Then*

$$\mathbb{P}(S_n = x) \leq \begin{cases} \mathbb{P}(W_n = k), & \text{if} \quad n+k \in 2\mathbb{Z}, \\ \mathbb{P}(W_{n-1} = k), & \text{if} \quad n+k \in 2\mathbb{Z}+1. \end{cases} \tag{15}$$

Proof. We first prove the Lemma for $x \in (0,1]$ and any n. By Theorem 1.0.1 we have

$$\mathbb{P}(S_n = x) \leq 2^{-n} \binom{n}{\lceil n/2 \rceil}. \tag{16}$$

On the other hand, if $x \in (0,1]$, then $k = 1$ and

$$2^{-n} \binom{n}{\lceil n/2 \rceil} = \begin{cases} 2^{-n} \binom{n}{(n+1)/2} = \mathbb{P}(W_n = 1), & \text{if} \quad n+1 \in 2\mathbb{Z}, \\ 2^{-n} \binom{n}{n/2} = \mathbb{P}(W_{n-1} = 1), & \text{if} \quad n+1 \in 2\mathbb{Z}+1, \end{cases}$$

where the second equality follows by Pascal's identity:

$$2^{-n} \binom{n}{n/2} = 2^{-n} \left[\binom{n-1}{n/2} + \binom{n-1}{n/2-1} \right] = 2^{1-n} \binom{n-1}{n/2} = \mathbb{P}(W_{n-1} = 1).$$

Let $\mathbb{N} = \{1, 2, \dots\}$ stand for the set of positive integers. Let us write $B_n(x)$ for the right-hand side of (15). Note that it has the following properties:

$$x \mapsto B_n(x) \text{ is non-increasing;} \tag{17}$$

$$x \mapsto B_n(x) \text{ is constant on each of the intervals } (k-1, k], \quad k \in \mathbb{N}; \tag{18}$$

$$B_n(k) = \tfrac{1}{2} B_{n-1}(k-1) + \tfrac{1}{2} B_{n-1}(k+1), \quad \text{if } k = 2, 3, \dots. \tag{19}$$

We proceed by induction on n. The case $n = 1$ is trivial. To prove the induction step

for $n \geq 2$, we consider two cases: (i) $x = k \in \mathbb{N}$; (ii) $k-1 < x < k \in \mathbb{N}$.

Case (i). For $k = 1$ the Lemma has been proved, so we assume that $k \geq 2$. By the inductional hypothesis we have

$$\begin{aligned}\mathbb{P}(S_n = k) &= \tfrac{1}{2}\mathbb{P}(S_{n-1} = k - a_n) + \tfrac{1}{2}\mathbb{P}(S_{n-1} = k + a_n) \\ &\leq \tfrac{1}{2}B_{n-1}(k - a_n) + \tfrac{1}{2}B_{n-1}(k + a_n). \end{aligned} \tag{20}$$

By (17) we have

$$B_{n-1}(k - a_n) \leq B_{n-1}(k-1), \tag{21}$$

and by (18) we have

$$B_{n-1}(k + a_n) = B_{n-1}(k+1). \tag{22}$$

Combining (20), (21), (22), and (19), we obtain

$$\mathbb{P}(S_n = k) \leq B_n(k). \tag{23}$$

Case (ii). For $x \in (0,1]$ the Lemma has been proved, so we assume $k \geq 2$. Consider two cases: (iii) $x/a_n \geq k$; (iv) $x/a_n < k$.

Case (iii). Define $S'_n = a'_1\varepsilon_1 + \cdots + a'_n\varepsilon_n$, where $a'_i = ka_i/x$, so that $S'_n = \frac{k}{x}S_n$. Recall that $a_n = \max_i a_i$, by the hypothesis of Lemma. Then $a'_i \leq ka_n/x$ and the assumption $x/a_n \geq k$ implies that $0 < a'_1, \ldots, a'_n \leq 1$. Therefore, by (23) and (18) we have

$$\mathbb{P}(S_n = x) = \mathbb{P}\left(S'_n = k\right) \leq B_n(k) = B_n(x).$$

Case (iv). Without loss of generality, we can assume that $a_n = 1$, since

$$\mathbb{P}(S_n = x) = \mathbb{P}\left(\frac{a_1}{a_n}\varepsilon_1 + \cdots + \frac{a_n}{a_n}\varepsilon_n = \frac{x}{a_n}\right)$$

and $k-1 < x/a_n < k$, by the assumption of the present case. Sequentially applying the induction hypothesis, (18), (19), and again (18), we get

$$\begin{aligned}
\mathbb{P}(S_n = x) &= \tfrac{1}{2}\mathbb{P}(S_{n-1} = x-1) + \tfrac{1}{2}\mathbb{P}(S_{n-1} = x+1) \\
&\leq \tfrac{1}{2}B_{n-1}(x-1) + \tfrac{1}{2}B_{n-1}(x+1) \\
&= \tfrac{1}{2}B_{n-1}(k-1) + \tfrac{1}{2}B_{n-1}(k+1) \\
&= B_n(k) = B_n(x).
\end{aligned}$$

■

Proof of Theorem 4.0.2. Writing $B_n(k)$ for the right-hand side of (15), we have, by Lemma 4.1.2, that

$$\mathbb{P}(S_n = x) \leq \max_{j=k}^{n} B_j(k).$$

If $j+k \in 2\mathbb{Z}$, then $B_j(k) = \mathbb{P}\left(W_j = k\right) = B_{j+1}(k)$ and therefore

$$\max_{j=k}^{n} B_j(k) = \max_{\substack{k \leq j \leq n \\ k+j \in 2\mathbb{Z}}} \mathbb{P}\left(W_j = k\right). \tag{24}$$

To finish the proof, note that the sequence $\mathbb{P}\left(W_j = k\right) = 2^{-j}\binom{j}{(k+j)/2}$, $j = k, k+2, k+4, \ldots$ is unimodal with a peak at $j = k^2$, i.e.,

$$\mathbb{P}\left(W_{j-2} = k\right) \leq \mathbb{P}\left(W_j = k\right), \quad \text{if} \quad j \leq k^2,$$

and

$$\mathbb{P}\left(W_{j-2} = k\right) > \mathbb{P}\left(W_j = k\right), \quad \text{if} \quad j > k^2.$$

Indeed, elementary calculations yield that the inequality

$$2^{-j+2}\binom{j-2}{(k+j)/2-1} \leq 2^{-j}\binom{j}{(k+j)/2}, \qquad j \geq k+2,$$

is equivalent to the inequality $j \leq k^2$. ■

4.2 Proofs based on results in extremal combinatorics

Let $[n]$ stand for the finite set $\{1,2,\ldots,n\}$. Consider a family $\mathcal{F}$ of subsets of $[n]$. We denote by $|\mathcal{F}|$ the cardinality of $\mathcal{F}$. The family $\mathcal{F}$ is called k-**intersecting** if for all $A,B \in \mathcal{F}$ we have $|A \cap B| \geq k$ and an **antichain** if for all $A,B \in \mathcal{F}$ we have $A \nsubseteq B$.

A well known result by Katona [24] (see also [9], p. 98, Theorem 4) gives the optimal upper bound for the size a k-intersecting family.

Theorem 4.2.1. *If $k \geq 1$ and $\mathcal{F}$ is a k-intersecting family of subsets of $[n]$ then*

$$|\mathcal{F}| \leq \begin{cases} \sum_{j=t}^{n} \binom{n}{j}, & \text{if } k+n=2t, \\ \sum_{j=t}^{n} \binom{n}{j} + \binom{n-1}{t-1}, & \text{if } k+n=2t-1. \end{cases} \tag{25}$$

Notice that if $k+n=2t$, then

$$\sum_{j=t}^{n} \binom{n}{j} = 2^n \mathbb{P}(W_n \geq k). \tag{26}$$

If $k+n=2t-1$, then using the Pascal's identity $\binom{n}{j} = \binom{n-1}{j} + \binom{n-1}{j-1}$ we get

$$\sum_{j=t}^{n} \binom{n}{j} + \binom{n-1}{t-1} = 2 \sum_{j=t-1}^{n-1} \binom{n-1}{j} = 2^n \mathbb{P}(W_{n-1} \geq k). \tag{27}$$

The exact upper bound for the size of a k-intersecting antichain is given by the following result of Milner [33].

Theorem 4.2.2. *If a family $\mathcal{F}$ of subsets of $[n]$ is a k-intersecting antichain, then*

$$|\mathcal{F}| \leq \binom{n}{t}, \qquad t = \left\lceil \frac{n+k}{2} \right\rceil. \tag{28}$$

Note that we have

$$\binom{n}{t} = 2^n \mathbb{P}\left(W_n = k\right), \quad \text{if} \quad n+k = 2t, \tag{29}$$

and

$$\binom{n}{t} = 2^n \mathbb{P}\left(W_n = k+1\right), \quad \text{if} \quad n+k = 2t-1. \tag{30}$$

By Lemma 4.1.1 it is enough to prove Theorems 4.0.1 and 4.0.2 for the sums

$$S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n,$$

where $0 \le a_1, \ldots, a_n \le 1$. Denote as A^c the complement of the set A. For each $A \subset [n]$, write $s_A = \sum_{i\in A} a_i - \sum_{i\in A^c} a_i$. We define two families of sets:

$$\mathcal{F}_{\ge x} = \{A \subset [n] : s_A \ge x\}, \quad \text{and} \quad \mathcal{F}_x = \{A \subset [n] : s_A = x\}.$$

Proof of Theorem 4.0.1. We have

$$\mathbb{P}\left(S_n \ge x\right) = 2^{-n}|\mathcal{F}_{\ge x}|.$$

Let $k = \lceil x \rceil$. Since W_n takes only integer values, we have

$$\mathbb{P}\left(W_n \ge k\right) = \mathbb{P}\left(W_n \ge x\right) \qquad \text{and} \qquad \mathbb{P}\left(W_{n-1} \ge k\right) = \mathbb{P}\left(W_{n-1} \ge x\right).$$

Therefore, in the view of (25), (26), and (27), it is enough to prove that $\mathcal{F}_{\ge x}$ is k-intersecting. Suppose that there are $A, B \in \mathcal{F}_{\ge x}$ such that $|A \cap B| \le k-1$. Writing $\sigma_A = \sum_{i\in A} a_i$, we have

$$s_A = \sigma_A - \sigma_{A^c} = \left(\sigma_{A\cap B} - \sigma_{A^c\cap B^c}\right) + \left(\sigma_{A\cap B^c} - \sigma_{A^c\cap B}\right) \tag{31}$$

and

$$s_B = \sigma_B - \sigma_{B^c} = (\sigma_{A\cap B} - \sigma_{A^c\cap B^c}) - (\sigma_{A\cap B^c} - \sigma_{A^c\cap B}). \tag{32}$$

Since

$$\sigma_{A\cap B} - \sigma_{A^c\cap B^c} \le \sigma_{A\cap B} \le |A\cap B| \le k-1 < x,$$

from (31) and (32) we get

$$\min\{s_A, s_B\} < x,$$

which contradicts the fact $s_A, s_B \ge x$. ■

The following Lemma implies Theorem 4.0.2. It also gives the optimal bound for $\mathbb{P}(S_n = x)$ and thus improves Lemma 4.1.2.

Lemma 4.2.3. *Let $0 < a_1, \ldots, a_n \le 1$ be strictly positive numbers, $x > 0$, $k = \lceil x \rceil$. Then*

$$\mathbb{P}(S_n = x) \le \begin{cases} \mathbb{P}(W_n = k), & \text{if} \quad n+k \in 2\mathbb{Z}, \\ \mathbb{P}(W_n = k+1), & \text{if} \quad n+k \in 2\mathbb{Z}+1. \end{cases}$$

Proof. We have

$$\mathbb{P}(S_n = x) = 2^{-n}|\mathcal{F}_x|.$$

In the view of (28), (29), and (30), it is enough to prove that $\mathcal{F}_x$ is a k-intersecting antichain. To see that $\mathcal{F}_x$ is k-intersecting it is enough to note that $\mathcal{F}_x \subset \mathcal{F}_{\ge x}$. To show that $\mathcal{F}_x$ is an antichain is even easier. If $A, B \in \mathcal{F}_x$ and $A \subsetneq B$, then $s_B - s_A = 2\sum_{i\in B\setminus A} a_i > 0$, which contradicts the assumption that $s_B = s_A = x$. ■

Proof of Theorem 4.0.2. Lemma 4.2.3 gives

$$\mathbb{P}(S_n = x) \le \max_{j=k}^{n} \mathbb{P}\left(W_j = k+1-\mathbb{I}(k,j)\right),$$

where again $\mathbb{I}(k,j) = \mathbb{I}\{k+j \in 2\mathbb{Z}\}$. Note that if $k+j \in 2\mathbb{Z}$ we have

$$\begin{aligned}\mathbb{P}\left(W_j = k\right) &\geq \frac{1}{2}\mathbb{P}\left(W_j = k\right) + \frac{1}{2}\mathbb{P}\left(W_j = k+2\right) \\ &= \mathbb{P}\left(W_{j+1} = k+1\right), \qquad k > 0.\end{aligned}$$

Hence

$$\max_{j=k}^{n} \mathbb{P}\left(W_j = k+1-\mathbb{I}(k,j)\right) = \max_{\substack{k \leq j \leq n \\ k+j \in 2\mathbb{Z}}} \mathbb{P}\left(W_j = k\right),$$

the right-hand side being the same as the one of (24). Therefore, repeating the argument following (24) we are done. ■

4.3 Extension to Lipschitz functions

One can extend Theorem 4.0.1 to odd Lipschitz functions taken of n independent random variables. Consider the cube $C_n = [-1,1]^n$ with the ℓ^1 metric d. We say that a function $f : C_n \to \mathbb{R}$ is K-Lipschitz with $K > 0$ if

$$|f(x) - f(y)| \leq Kd(x,y), \qquad x,y \in C_n. \tag{33}$$

We say that a function $f : C_n \to \mathbb{R}$ is odd if $f(-x) = -f(x)$ for all $x \in C_n$. An example of an odd 1-Lipschitz function is the function mapping a vector to the sum of its coordinates:

$$f(x_1,\ldots,x_n) = x_1 + \cdots + x_n.$$

As in Theorems 4.0.1 and 4.0.2, the crux of the proof is dealing with two-valued random variables. The optimal bound for a k-intersecting family is not sufficient for this case, therefore we use the following generalization of Theorem 4.2.1 due to Kleitman [28] (see also [9][p. 102]) which we state slightly reformulated for our convenience. Let us define the diameter of a set family $\mathcal{F}$ by $\operatorname{diam}\mathcal{F} = \max_{A,B\in\mathcal{F}} |A \triangle B|$.

Theorem 4.3.1. *If $k \geq 1$ and $\mathcal{F}$ is a family of subsets of $[n]$ with* $\operatorname{diam} \mathcal{F} \leq n-k$*, then*

$$|\mathcal{F}| \leq \begin{cases} \sum_{j=t}^{n} \binom{n}{j}, & \text{if } k+n=2t, \\ \sum_{j=t}^{n} \binom{n}{j} + \binom{n-1}{t-1}, & \text{if } k+n=2t-1. \end{cases} \tag{34}$$

To see that Theorem 4.3.1 implies Theorem 4.2.1, observe that $|A \cap B| \geq k$ implies $|A \triangle B| \leq n-k$.

Theorem 4.3.2. *Suppose that a function $f : C_n \to \mathbb{R}$ is 1-Lipschitz and odd. Let $X_1, \ldots, X_n$ be symmetric independent random variables such that $|X_i| \leq 1$. Then, for $x > 0$, we have that*

$$\mathbb{P}\left(f(X_1, \ldots, X_n) \geq x\right) \leq \begin{cases} \mathbb{P}\left(W_n \geq x\right), & \text{if} \quad n + \lceil x \rceil \in 2\mathbb{Z}, \\ \mathbb{P}\left(W_{n-1} \geq x\right), & \text{if} \quad n + \lceil x \rceil \in 2\mathbb{Z}+1. \end{cases} \tag{35}$$

Proof. Applying Lemma 4.1.1 with the function

$$g(y_1, \ldots, y_n) = \mathbb{I}\{f(y_1, \ldots, y_n) \geq x\},$$

we can see that it is enough to prove (35) with

$$X_1 = a_1 \varepsilon_1, \ldots, X_n = a_n \varepsilon_n$$

for any 1-Lipschitz odd function f. In fact, we can assume that $a_1 = \cdots = a_n = 1$, since the function

$$(x_1, \ldots, x_n) \mapsto f(a_1 x_1, \ldots, a_n x_n)$$

is clearly 1-Lipschitz and odd.

Given $A \subseteq [n]$, write f_A for $f(2\,\mathbb{I}_A(1) - 1, \ldots, 2\,\mathbb{I}_A(n) - 1)$, where $\mathbb{I}_A$ is the indicator

function of the set A. Note that

$$|f_A - f_B| \leq 2|A \triangle B| \tag{36}$$

by the Lipschitz property. Consider the family of finite sets

$$\mathcal{F} = \{A \subseteq [n] : f_A \geq x\},$$

so that

$$\mathbb{P}\left(f(\varepsilon_1, \ldots, \varepsilon_n) \geq x\right) = 2^{-n}|\mathcal{F}|.$$

Write $k = \lceil x \rceil$. Note that W_{n-1} and W_n take only integer values. Therefore by (26) and (27) we see that the right-hand side of (34) is equal, up to the power of two, to the right-hand side of (35). Consequently, if $\operatorname{diam} \mathcal{F} \leq n-k$, then Theorem 4.3.1 implies (35). Therefore, it remains to check that for any $A, B \in \mathcal{F}$ we have $|A \triangle B| \leq n-k$.

Suppose that for some A, B we have $f_A, f_B \geq x$ but $|A \triangle B| \geq n-k+1$. Then

$$|A \triangle B^c| = |(A \triangle B)^c| = n - |A \triangle B| \leq k-1,$$

and hence by (36) we have

$$|f_A - f_{B^c}| \leq 2k-2. \tag{37}$$

On the other hand, we have that $f_{B^c} \leq -x$, as f is odd. Therefore

$$f_A - f_{B^c} \geq 2x > 2k-2,$$

which contradicts (37). ■

Part II

Bootstrap Percolation

CHAPTER 5

MAJORITY BOOTSTRAP PERCOLATION ON THE ERDŐS-RENYI RANDOM GRAPH

The mathematical study of percolation took off after Broadbent and Hammersley introduced the following problem in [13]. Given an infinite graph G, with finite maximum degree, select each edge to be open or closed independently and with probabilities p or $1-p$ respectively. We ask the question whether there is a non-zero probability of a vertex v having an infinite connected component in the open edge subgraph? For G connected the answer to this question is the same irrespective of the vertex v considered. This is because the probability that v_1 is in an infinite component is at least p^d times the probability v_2 is in an infinite component, where d is the distance between v_1 and v_2 in G. The probability of a fixed vertex v being in an infinite open component is also clearly increasing with p, and so a lot of work has gone into determining $p_c = \inf\{p : \mathbb{P}_p(v \text{ in } \infty \text{ component}) > 0\}$, the critical edge percolation probability, for many graphs G.

The most natural class of graphs to study this problem on are lattices. Perhaps the most celebrated result in this area is Harris [21] and Kesten's [27] proof that $p_c(\mathbb{Z}^2) = \frac{1}{2}$. Much further study has gone into this problem and the critical probability has been found for many lattices. For example, Wierman in [39] and [40] found the critical probability of certain self-dual planar lattices, a result which was vastly extended by Bollobás and Riordan in [11]. Despite all this progress there are still many open cases, for example it is still not known, or even commonly conjectured, what the value of $p_c(\mathbb{Z}^3)$ is.

The classical bootstrap percolation, called r-neighbour bootstrap percolation, concerns a deterministic process on a graph. Firstly, a subset of the vertices of a graph G is initially infected. Then at each time step the infection spreads to any vertex with at least r infected neighbours. This process is a cellular automaton, of the type first introduced by von Neumann in [34]. This particular model was introduced by Chalupa, Leith and Reich in [14], where G was taken to be the Bethe lattice.

A standard way of choosing the initially infected vertices is to independently infect each vertex with probability p. The probability that the entire graph eventually becomes infected is increasing with p. It is therefore sensible to study the quantity $p_c = \inf\{p : \mathbb{P}_p(G \text{ infected}) \geq c\}$, in particular the critical probability $p_{1/2}$ and the size of the critical window $p_{1-\varepsilon} - p_\varepsilon$.

A natural setting for this problem is the finite grid $[n]^d$. Many of the results on bootstrap percolation concern this problem. The first to study this graph were Aizenman and Lebowitz in [1], who showed that in 2-neighbour bootstrap percolation when d is fixed we have $p_{1/2} = \Theta((\log n)^{1-d})$.

The r-neighbour bootstrap percolation process has also been studied on the random regular graph by Balogh in [4] and on the Erdős-Rényi random graph $G(n,p)$ by Janson Łuczak, Turova and Vallier in [23].

In majority bootstrap percolation a vertex becomes infected if a majority of its neighbours are. In [3] Balogh, Bollobás and Morris studied this process on the hypercube and showed that if the vertices of the n-dimensional hypercube are independently infected with probability

$$q = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\log n}{n}} + \frac{\lambda \log\log n}{\sqrt{n \log n}},$$

then with high probability percolation occurs if $\lambda > \frac{1}{2}$ and does not occur if $\lambda \leq -2$.

We shall study majority bootstrap percolation on the Erdős-Rényi random graph above the connectivity threshold.

5.1 Main Results

In this section we shall state our main results and discuss two different ways of selecting the initially infected set. Throughout this section we shall make use of some technical lemmas, which we include at the very of this chapter in Section 5.4 so as not to

disrupt the flow of our arguments.

For G a graph with some subset $I_0 \subset V(G)$ of initially infected vertices, the majority bootstrap process on G is defined by setting $I_{t+1} = I_t \cup \{v \in V(G) : |I_t \cap \Gamma(v)| \geq \frac{|\Gamma(v)|}{2}\}$. For a finite graph G, this process will terminate with $I_{T+1} = I_T$. Denote by $I = I_T$ the set of eventually infected vertices.

We shall look at the case of $G = G(n,p)$, the graph on n vertices where each edge is included independently with probability p. Our initial setup is also slightly different, instead of infecting each vertex independently with some probability q, we shall infect a random set of vertices of size m.

In the normal setup for the majority bootstrap process on $G(n,p)$, we would first choose the edges of $G(n,p)$ and then choose an initially infected set I_0 uniformly from $[n]^{(m)}$. As these two choices are independent we shall equivalently set $I_0 = [m]$ and then choose the edges of $G(n,p)$. This is the $MB(n,p\,;m)$ process.

We now introduce some notation and conventions that shall be used in this chapter. We set $d = \frac{np}{1-p}$, thus d is roughly the average degree in $G(n,p)$ for $p = o(1)$. We denote the binomial distribution with mean with parameters n and p by $B(n,p)$. We shall sometimes abuse the notation and denote by $B(n,p)$ a random variable that has a binomial distribution. We reserve m for the size of I_0 and shall always assume that $m = \frac{n}{2} - \frac{n}{2}\sqrt{\frac{\log d}{d}} + \lambda n \frac{\log\log\log d}{\sqrt{d\log d}}$, for some constant λ. We use the standard asymptotic little-o notation and this is always taken as n or N tend to infinity. An increasing unbounded function shall be denoted by $\omega(n)$. Unless otherwise stated any random variables mentioned will be independent. Throughout this chapter the inequalities are only claimed to be true for n large enough. For the $MB(n,p\,;m)$ process, define

$$\mathcal{P}_m(G(n,p)) = \mathbb{P}(I = [n]).$$

The main result we obtained jointly with Nathan Kettle and Cecilia Holmgren is the

following theorem.

Theorem 5.1.1. *Fix some number* $\varepsilon > 0$*. Assume that*

$$(1+\varepsilon)\log n \leq p(1-p)n \quad \textit{and} \quad p \leq 0.99.$$

If the initially infected set I_0 *has size*

$$m = \frac{n}{2} - \frac{n}{2}\sqrt{\frac{\log d}{d}} + \lambda n \frac{\log\log\log d}{\sqrt{d\log d}}$$

then

$$\mathcal{P}_m(G(n,p)) \to \begin{cases} 1, & \textit{if } \lambda > \frac{1}{2}, \\ 0, & \textit{if } \lambda < 0. \end{cases} \tag{38}$$

Our second result concerns a more natural setup, where each vertex is initially independently infected with probability q, we have that with high probability $|I_0| = qn \pm \omega(n)\sqrt{q(1-q)n}$. When $\sqrt{n} \ll \frac{n\log\log\log d}{\sqrt{d\log d}}$, i.e, when $p \ll \frac{(\log\log\log n)^2}{\log n}$, our result above shall still hold in this setting for $q = m/n$.

More formally define the $MB'(n,p\,;q)$ to be the process in which the graph $G(n,p)$ is chosen and each vertex is initially infected independently with probability q. Then the infection spreads by the majority bootstrap percolation process. For the process $MB'(n,p\,;q)$ define

$$\mathcal{P}'_q(G(n,p)) = \mathbb{P}(I = [n]).$$

Corollary 5.1.2. *Fix a number* $\varepsilon > 0$*. Assume that*

$$(1+\varepsilon)\log n \leq p(1-p)n \quad \textit{and} \quad p \ll \frac{(\log\log\log n)^2}{\log n},$$

then with $q = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\log d}{d}} + \lambda \frac{\log\log\log d}{\sqrt{d\log d}}$, *we have*

$$\mathcal{P}'_q(G(n,p)) \to \begin{cases} 1, & \text{if } \lambda > \frac{1}{2}, \\ 0, & \text{if } \lambda < 0. \end{cases} \tag{39}$$

If $p \leq 0.99$ *and* $p \gg \frac{(\log\log\log n)^2}{\log n}$, *then with* $q = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\log d}{d}} + \theta\frac{1}{\sqrt{n}}$, *we have*

$$\mathcal{P}'_q(G(n,p)) \to \Phi(2\theta),$$

where $\Phi(x)$ *denotes the distribution function of the standard Normal random variable.*

Proof. As each vertex is infected independently, $|I_0|$ has distribution $B(n,q)$. Thus with high probability it holds that $||I_0| - qn| \leq \omega(n)\sqrt{q(1-q)n}$. If $p \ll \frac{(\log\log\log n)^2}{\log n}$, then $n\frac{\log\log\log d}{\sqrt{d\log d}} \gg \sqrt{n}$ and the result follows from Theorem 5.1.1.

If $p \gg \frac{(\log\log\log n)^2}{n}$ then for each fixed $\delta > 0$ by the Central Limit Theorem we obtain

$$\begin{aligned} & \mathcal{P}'_q(G(n,p)) \\ = & \sum_{m=0}^{n} \mathbb{P}(B(n,q) = m)\, \mathcal{P}_m(G(n,p)) \\ \geq & \mathbb{P}\left(B(n,q) \geq qn + (\delta-\theta)\sqrt{n}\right) \mathcal{P}_{qn+(\delta-\theta)\sqrt{n}}(G(n,p)) \\ = & \mathbb{P}\left(B(n,q)/\sqrt{q(1-q)n} \geq (qn + (\delta-\theta)\sqrt{n})/\sqrt{q(1-q)n}\right)(1+o(1)) \\ \to & \Phi(2(\theta-\delta)), \end{aligned} \tag{40}$$

where the fourth line follows as $\mathcal{P}_{qn+(\delta-\theta)\sqrt{n}}(G(n,p)) \to 1$ for $p \gg \frac{(\log\log\log n)^2}{\log n}$ by Theorem 5.1.1. A similar argument shows that $1 - \mathcal{P}'_q(G(n,p)) \geq \Phi(-2(\theta+\varepsilon))(1+o(1))$ and so $\mathcal{P}'_q(G(n,p)) \to \Phi(2\theta)$. ■

When p is smaller than the connectivity threshold, $G(n,p)$ contains isolated vertices. Due to the way we define the $MB(n,p\,;m)$ process, any uninfected isolated vertex becomes infected in the first time step, so this is not an obstruction to complete

percolation. However once p drops to below $\frac{\log n}{2n}$, then with high probability $G(n,p)$ contains isolated edges and neither endpoint of an isolated edge becomes infected if both endpoints are initially uninfected. This means that $\mathcal{P}_m(G(n,p)) \to 0$ unless $m = n - o(n)$.

5.2 Upper Bound

As G is finite the $MB(n,p\,;m)$ process will eventually terminate with some set $I_0 \subset [n]$ of infected vertices. If we do not infect the whole graph, or, equivalently, we have that $I_0 \neq [n]$, then we can say something about the structure of I_0. We shall call a proper subset S of $[n]$ *closed* if for all $v \in [n] \setminus S$ we have $|\Gamma(v) \cap S| < \frac{|\Gamma(v)|}{2}$. As $I_0 \neq [n]$ we must have that the initially infected vertices I_0 are contained in a closed set. We shall show that with high probability I_0 is contained in no closed sets in three stages. Using Lemma 5.2.2 will allow us that with high probability the graph $G(n,p)$ has no "large" closed sets. After that we shall bound the expected number of medium sized closed sets that I_0 is in. But before we proceed with proving these two facts let us show that with high probability the number of infected vertices after one time step, $|I_1|$, is large, and so I_0 can rarely be contained in a small closed set. Recall that

$$m = \frac{n}{2} - \frac{n}{2}\sqrt{\frac{\log d}{d}} + \lambda n \frac{\log\log\log d}{\sqrt{d\log d}}.$$

Lemma 5.2.1. *In the $MB(n,p\,;m)$ process,*

$$|I_1 \setminus I_0| \geq \frac{n(\log\log d)^{2\lambda}}{e^8\sqrt{d\log d}}$$

with high probability.

Proof. For $i \in [n] \setminus I_0$, denote by A_i the event that vertex i is infected at time one, that is the event that i has fewer neighbours in $[n] \setminus I_0$ than it does in I_0. The events A_i are identical and very weakly correlated but not independent. Let X be the number of vertices

infected at the first step of the process. Then $X = |I_1 \setminus I_0| = \sum \mathbb{1}(A_i)$. We shall use Chebyshev's inequality to bound the probability that X is small.

As the events A_i are identical we shall set $r = \mathbb{P}(A_i)$, so $\mathbb{E}(X) = (n-m)r$. Let μ_1 and μ_2 are the means of $B(m,p)$ and $B(n-m-1,(1-p))$ respectively. We have that

$$\begin{aligned} r &= \mathbb{P}\left(|\Gamma(i) \cap I_0| \geq \Gamma(i) \cap ([n] \setminus I_0)\right) \\ &= \mathbb{P}\left(B(m,p) \geq B(n-m-1,p)\right) \\ &= \mathbb{P}\left(B(m,p) + B(n-m-1,(1-p)) \geq \mu_1 + \mu_2 + p(n-2m-1)\right). \end{aligned}$$

We have $p(n-2m-1) = \omega(n)\sqrt{p(1-p)n}$ and $p(n-2m-1)^2 = o(n\sqrt{p(1-p)n})$. Applying the bound from Proposition 5.4.11 to the last equality with $N = \frac{n-1}{2}$, $S = \frac{n-1-2m}{2}$ and $h = p(n-2m-1)$, we obtain

$$\begin{aligned} r &> \frac{\sqrt{p(1-p)(n-1)}}{2\pi p(n-2m-1)} \exp\left(-\frac{p(n-2m-1)^2}{2(1-p)(n-1)} - 4 - o(1)\right) \\ &> \frac{1}{2\pi\sqrt{\log d}} \exp\left(-\frac{\log d}{2} + 2\lambda \log\log\log d - 4 + o(1)\right) \\ &> \left(\frac{(\log\log d)^{2\lambda}}{2\pi e^4 \sqrt{d \log d}}\right)(1+o(1)). \end{aligned} \tag{41}$$

where in the second line we have used the asymptotic relation

$$d(n-2m-1)^2 = n^2 \log d - 4\lambda n^2 \log\log\log d + o(n^2).$$

Let us calculate the variance of X. We have

$$\begin{aligned} \mathrm{Var}(X) &= \sum_{i,j \in [n]\setminus[m]} (\mathbb{P}(A_j|A_i) - \mathbb{P}(A_j))\mathbb{P}(A_i) \\ &= (1-r)r(n-m) + r'r(n-m)(n-m-1), \end{aligned} \tag{42}$$

where $r' = \mathbb{P}(A_j|A_i) - r$, this being the same for any $i \neq j$. In (42) the first term is the sum over $i = j$ and the second term is the sum over $i \neq j$. Let B_{ij} and $\overline{B}_{ij}$ be the events that ij is, or is not, an edge in G respectively. We bound r' by

$$\begin{aligned}
\mathbb{P}(A_j|A_i) - \mathbb{P}(A_j) &= \mathbb{P}(A_j|B_{ij})\mathbb{P}(B_{ij}|A_i) + \mathbb{P}\left(A_j|\overline{B}_{ij}\right)\mathbb{P}\left(\overline{B}_{ij}|A_i\right) - \mathbb{P}\left(A_j\right) \\
&\leq \mathbb{P}\left(A_j|\overline{B}_{ij}\right) - \mathbb{P}\left(A_j\right) \\
&= \mathbb{P}\left(A_j|\overline{B}_{ij}\right)(1-(1-p)) \\
&= p\left(\mathbb{P}\left(A_j|\overline{B}_{ij}\right)\right) \\
&= p\mathbb{P}\left(B(m,p) = B(n-m-2,p)\right), \qquad (43)
\end{aligned}$$

where the last equality follows because the probabilities of the events $A_j|B_{ij}$ and $A_j|\overline{B}_{ij}$ are equal to $\mathbb{P}(B(m,p) \geq B(n-m-2,p)+1)$ and $\mathbb{P}(B(m,p) \geq B(n-m-2,p))$ respectively.

As $p(\frac{n}{2} - m - 1) = \omega(n)\sqrt{p(1-p)n}$ we get from Proposition 5.4.13 applied with $(N,S,T) = (\frac{n}{2}-1, \frac{n}{2}-m-1, 0)$, that r' is at most

$$\begin{aligned}
&\frac{p(\frac{n}{2}-m-1)}{2\pi(1-p)(\frac{n}{2}-1)} \exp\left(-\frac{p(\frac{n}{2}-m-1)^2}{(1-p)(\frac{n}{2}-1)} + o(1)\right) \\
&+ \frac{3}{\pi(\frac{n}{2}-m-1)} \exp\left(-\frac{9p(\frac{n}{2}-m-1)^2}{8(1-p)(\frac{n}{2}-1)}\right) \\
&< \frac{p\sqrt{\log d}}{2\pi(1-p)\sqrt{d}} \exp\left(-\frac{\log d}{2} + 2\lambda \log\log\log d + o(1)\right) \\
&+ \frac{6\sqrt{d}}{\pi n\sqrt{\log d}} \exp\left(-\frac{9\log d}{16} + \frac{9\lambda \log\log\log d}{4} + o(1)\right). \qquad (44)
\end{aligned}$$

The second term is much smaller than the first term and so,

$$r' < \left(\frac{\sqrt{\log d}(\log\log d)^{2\lambda}}{\pi n}\right). \qquad (45)$$

We are now able to bound the probability that X is small. From (42) and Chebyshev's

inequality we get

$$\begin{aligned}
\mathbb{P}\left(X \leq \frac{(n-m)r}{2}\right) &\leq \mathbb{P}\left(|X-(n-m)r| \geq \frac{(n-m)r}{2}\right) \\
&\leq \frac{4\mathrm{Var}(X)}{((n-m)r)^2} \\
&= \frac{4((1-r)+(n-m-1)r')}{(n-m)r} \\
&< \frac{4r'}{r}+o(1). \qquad (46)
\end{aligned}$$

From (45) and (41) this is at most

$$\left(\frac{2e^4 \log d\sqrt{d}}{n}\right)(1+o(1))+o(1) = o(1), \qquad (47)$$

and so we have with high probability that $|I_1 \setminus I_0|$ is at least $\frac{(n-m)r}{2}$, which for large n is greater than

$$\frac{n(\log\log d)^{2\lambda}}{e^8\sqrt{d\log d}},$$

which completes the proof. ■

We now show that $G(n,p)$ contains no large closed sets by a simple edge set comparison.

Lemma 5.2.2. *Suppose that for some fixed* $\varepsilon > 0$ *we have* $p(1-p)n \geq (1+\varepsilon)\log n$. *Then with high probability* $G(n,p)$ *contains no closed set of size greater than* $\frac{n}{2}+\frac{7n}{2\sqrt{d}}$.

Proof. Let us write s for the size of the set S. In order for the set S to be closed each vertex v has to have the majority of its neighbours outside S. In other words, we must have $|\Gamma(v)\cap([n]\setminus S)| > |\Gamma(v)\cap S|$. Summing over the vertices in $[n]\setminus S$ we have that the number of edges from S to $[n]\setminus S$ must be fewer than twice the number of edges in $[n]\setminus S$. If $\frac{n}{2}+\frac{7n}{2\sqrt{d}} < s < \frac{4n}{5}$ then $p(2s-n) \geq 7\sqrt{p(1-p)n}$ and so

$$ps(n-s)-3(n-s)\sqrt{p(1-p)s} > 2p\binom{n-s}{2}+4(n-s)\sqrt{p(1-p)(n-s)}.$$

By Proposition 5.4.14 every set of size s has at most $p\binom{n-s}{2}+2(n-s)\sqrt{p(1-p)(n-s)}$ edges in its complement with probability at least $1-\frac{1}{4^s}$ and by Proposition 5.4.15 every set S of size s has at least $ps(n-s)-3(n-s)\sqrt{p(1-p)s}$ edges between it and its complement with probability at least $1-\frac{1}{4^s}$. Therefore with high probability every set S of size $\frac{n}{2}+\frac{7n}{2\sqrt{d}} < s < \frac{4n}{5}$ is not closed.

If $s \geq \frac{4n}{5}$ and $p(1-p)n \geq 4\log n$, then we know from Proposition 5.4.16 that with probability at least $1-n^{-\frac{n-s}{120}}$ there does not exist a closed set of size s in $G(n,p)$. The result follows as $\sum_{i=1} n^{-\frac{i}{120}} = o(1)$.

If $n-n^{\frac{27}{28}} \geq s \geq \frac{4n}{5}$ and $5\log n \geq p \geq (1+\varepsilon)\log n$, then we know from Corollary 5.4.17 that with probability at least $1-n^{-\frac{n-s}{120}}$ there does not exist a closed set of size s in $G(n,p)$.

If $s \geq n-n^{\frac{27}{28}}$ and $5\log n \geq p \geq (1+\varepsilon)\log n$, then we know from Proposition 5.4.19 that with probability at least $1-n^{-\frac{n-s}{120}}$ every set $[n]\setminus S$ of size $n-s$ has at most $2(n-s)$ edges and so has a vertex v_S of degree at most 4. By Proposition 5.4.18 we have that with high probability the minimum degree of $G(n,p)$ is at least 9 and so v_S will become infected if all of S is and so S is not closed.

■

Lastly, we turn to bounding the expected number of medium sized closed sets I_0 is contained in. We shall therefore want a bound on the probability that a set S of size at least s in a particular range of s is closed. To do this we shall pick a test set T of a suitable size and bound the probability that none of the vertices in T are infected by S.

Lemma 5.2.3. *Fix* $\varepsilon > 0$ *and define*

$$s = \frac{n}{2} - \frac{n\sqrt{\log d}}{2\sqrt{d}} + \frac{n(\log\log d)^{1+\varepsilon}}{\sqrt{d\log d}}.$$

Take any set of vertices S *in* $G(n,p)$ *of size* $s \leq |S| < \frac{2n}{3}$. *Then*

$$\mathbb{P}(S \textit{ is closed}) \leq \exp\left(-\frac{n(\log d)^{(\log\log d)^{\varepsilon}-2}}{e^7\sqrt{d}}\right).$$

Proof. Without loss of generality we shall set $S = [|S|]$ and $T = [t+|S|] \setminus S$, where $t = \left\lfloor \frac{n}{(\log d)^2} \right\rfloor$. We shall condition on the edge set of T as once we have done so the events F_v, that v is infected by S for each vertex $v \in T$, are independent.

Denote by $\mathcal{E}(T)$ the family of all possible edge sets on the vertex set T and set $d_E(v)$ to be the degree of vertex $v \in T$ when T has edge set E. We have that

$$\mathbb{P}(F_v|E) = \mathbb{P}(|\Gamma(v) \cap S| < d_E(v) + |\Gamma(v) \cap ([n] \setminus (S \cup T))|).$$

Therefore,

$$\begin{aligned} &\mathbb{P}(S \text{ is closed}) && (48)\\ &\leq \sum_E \mathbb{P}(E) \prod_{v \in T} \mathbb{P}(F_v|E) \\ &= \sum_E \mathbb{P}(E) \prod_{v \in T} \mathbb{P}(B(|S|,p) < B(n-|S|-t,p) + d_E(v)), && (49) \end{aligned}$$

where $\mathbb{P}(E)$ is the probability of a particular edge set $E \subset \{0,1\}^{\binom{t}{2}}$ and is equal to

$p^{|E|}(1-p)^{\binom{t}{2}-|E|}$.

The function $f_{|S|}(x) = \mathbb{P}(B(|S|,p) < B(n-|S|-t,p)+x)$ is decreasing in $|S|$ so we have $f_s(x) \geq f_{|S|}(x)$. Let us supress the dependency on s by writing $f(x)$ instead of $f_s(x)$. We have

$$\mathbb{P}(S \text{ closed}) \leq \sum_E \mathbb{P}(E) \prod_{v \in T} f(d_E(v)). \tag{50}$$

The rest of the proof shall be spent bounding (50). The degree of vertices in T is heavily concentrated around pt, and we shall expand f around pt to show that (50) is not much larger than $f(pt)^t$.

We have by Corollary 5.4.3 that f is log-concave and so for any x and y with $f(y) \neq 0$,

$$f(x) \leq f(y) \left(\frac{f(y+1)}{f(y)} \right)^{x-y}.$$

Setting $y = \lceil pt \rceil \in \mathbb{N}$ we get

$$\begin{aligned} \mathbb{P}(S \text{ closed}) &\leq \sum_E \mathbb{P}(E) \prod_{v \in T} f(y) \left(\frac{f(y+1)}{f(y)} \right)^{d_E(v)-y} \\ &= \sum_E \mathbb{P}(E) f(y)^t \left(\frac{f(y+1)}{f(y)} \right)^{2|E|-ty} \end{aligned} \tag{51}$$

We have removed any dependence on E other than its size and so

$$\begin{aligned} \mathbb{P}(S \text{ closed}) &\leq \sum_{i=0}^{\binom{t}{2}} \binom{\binom{t}{2}}{i} p^i (1-p)^{\binom{t}{2}-i} f(y)^t \left(\frac{f(y+1)}{f(y)} \right)^{2i-ty} \\ &= \left(1 - p + p \left(\frac{f(y+1)}{f(y)} \right)^2 \right)^{\binom{t}{2}} \left(\frac{f(y)}{f(y+1)} \right)^{ty} f(y)^t. \end{aligned} \tag{52}$$

Setting $\frac{f(y+1)}{f(y)} = 1 + a$, we bound (52) using the inequalities $1 + w \leq e^w$ and

$(1+x)^{-1} \leq 1-x+x^2$ for $x \geq 0$ to get

$$\begin{aligned}
\mathbb{P}(S \text{ closed}) &\leq \left(1+2ap+a^2p\right)^{\frac{t^2}{2}} \left(\frac{1}{1+a}\right)^{pt^2} f(y)^t \\
&\leq \exp\left((2ap+a^2p)\frac{t^2}{2}+(a^2-a)pt^2\right) f(y)^t \\
&= \exp\left(\frac{3pa^2t^2}{2}\right) f(y)^t \qquad (53)
\end{aligned}$$

We have that

$$f(y+1) = f(y) + \mathbb{P}\left(B(s,p) = B(n-s-t,p)+y\right).$$

Let us write $z = \mathbb{P}\left(B(s,p) = B(n-s-t,p)+y\right)$ to ease up the notation. Thus $f(y+1) = f(y)+z$. By Proposition 5.4.13 applied with $N = \frac{n-t+T}{2}$, $S = \frac{n-2s-t+T}{2}$ and $T = \frac{\lceil pt \rceil}{p}$ and noting that $0 \leq T-t < p^{-1}$, we have

$$\begin{aligned}
z &< \frac{n-2s+\frac{1}{p}}{2\pi(1-p)n} \exp\left(-\frac{2p(\frac{n}{2}-s)^2}{(1-p)(n-t)} + o(1)\right) \\
&\quad + \frac{6}{\pi(n-2s)} \exp\left(-\frac{9p(n-2s)^2}{8(1-p)(n+\frac{1}{p})}\right) \\
&= \frac{\sqrt{\log d}}{2\pi(1-p)\sqrt{d}} \exp\left((-\frac{\log d}{2} + 2(\log\log d)^{1+\varepsilon})(1+\frac{t}{n}) + o(1)\right) \\
&\quad + \frac{6\sqrt{d}}{\pi n\sqrt{\log d}} \exp\left(-\frac{9\log d}{16} + \frac{9(\log\log d)^{1+\varepsilon}}{4} + o(1)\right). \qquad (54)
\end{aligned}$$

The second term in (54) is much smaller than the first so as $6 < 2\pi$ and $t\log d = o(n)$ we get

$$z < \frac{\sqrt{\log d}(\log d)^{2(\log\log d)^{\varepsilon}}}{6(1-p)d}.$$

We can rewrite $f(y)$ as

$$f(y) = 1 - \mathbb{P}\left(B(s,p) + B(n-s-t,(1-p)) \geq n-s-t+y\right),$$

We have the asymptotic relation $(p(n-2s)+1)(t+2s-n) = o(n\sqrt{np(1-p)})$ and so using Proposition 5.4.11 with $(N,S,h) = (\frac{n-t}{2}, \frac{n-2s-t}{2}, p(n-2s)+y-pt)$ we obtain

$$\begin{aligned} f(y) &< 1 - \frac{\sqrt{p(1-p)(n-t)}}{2\pi(p(n-2s)+1)} \exp\left(-\frac{(p(n-2s)+1)^2}{2p(1-p)(n-t)} - 4 - o(1)\right) \\ &< 1 - \frac{(\log d)^{2(\log\log)^{\varepsilon}}}{e^6\sqrt{d\log d}} \\ &< \exp\left(-\frac{(\log d)^{(\log\log d)^{\varepsilon}}}{e^6\sqrt{d}}\right), \end{aligned} \tag{55}$$

the second inequality follows from the same reasoning used in (54) and that $e^6 > 2\pi e^4$.

We can also apply Proposition 5.4.12 to get a lower bound on $f(y)$ of

$$f(y) > 1 - \frac{\sqrt{p(1-p)(n-t)}}{p(n-2s)} \exp\left(-\frac{p(n-2s)^2}{2(1-p)(n-t)} + 4\right) > \frac{1}{2},$$

here the bound on $1-f(y)$ is actually $o(1)$, being within a constant factor of the bound in (55).

We are now able to get a good upper bound on a,

$$a = \frac{z}{f(y)} < \frac{\sqrt{\log d}(\log d)^{2(\log\log d)^{\varepsilon}}}{3(1-p)d}.$$

Substituting these bounds into (53) we get

$$\mathbb{P}(S \text{ closed}) < \exp\left(\frac{p(\log d)^{4(\log\log d)^{\varepsilon}} n}{6(1-p)^2 d^2 \log d} - \frac{(\log d)^{(\log\log d)^{\varepsilon}}}{e^6\sqrt{d}}\right)^t.$$

The second term in the exponential is much larger than the first term and so

$$\begin{aligned}\mathbb{P}(S \text{ closed}) &< \exp\left(-\frac{(\log d)^{(\log\log d)^{\varepsilon}}}{2e^6\sqrt{d}}\right)^t \\ &< \exp\left(-\frac{n(\log d)^{(\log\log d)^{\varepsilon}-2}}{e^7\sqrt{d}}\right),\end{aligned}$$

as $t > \frac{2n}{e(\log d)^2}$. ■

We shall now bound the expected number of closed sets in this medium sized range that contain I_0, this is also a bound on the probability that I_0 is contained in such a medium sized closed set.

Proposition 5.2.4. *Assume that*

$$m = \frac{n}{2} - \frac{n\sqrt{\log d}}{2\sqrt{d}} + \frac{n\lambda \log\log\log d}{\sqrt{d\log d}}.$$

Then the expected number of closed sets in $G(n,p)$ of size between

$$\frac{n}{2} - \frac{n\sqrt{\log d}}{2\sqrt{d}} + \frac{n(\log\log d)^{1+\varepsilon}}{\sqrt{d\log d}} \quad \textit{and} \quad \frac{n}{2} + \frac{4n}{\sqrt{d}}$$

that contain [m] is $o(1)$.

Proof. Let S be a set of size s in our range, s can have at most $\frac{n\sqrt{\log d}}{\sqrt{d}}$ different values. For each possible value of s there are at most

$$\binom{n-m}{s-m} < \binom{n}{\frac{n\sqrt{\log d}}{\sqrt{d}}} < \left(\frac{e\sqrt{d}}{\sqrt{\log d}}\right)^{\frac{n\sqrt{\log d}}{\sqrt{d}}} < \exp\left(\frac{n(\log d)^{\frac{3}{2}}}{\sqrt{d}}\right)$$

possible closed sets. By Lemma 5.2.3 the expected number of closed sets is less than

$$\frac{n\sqrt{\log d}}{\sqrt{d}} \exp\left(\frac{n(\log d)^{\frac{3}{2}}}{\sqrt{d}} - \frac{n(\log d)^{(\log\log d)^{\varepsilon}-2}}{e^8\sqrt{d}}\right),$$

and this is $o(1)$ as $(\log\log d)^{\varepsilon}$ is unbounded. ■

Corollary 5.2.5. *If $\lambda > \frac{1}{2}$ then with high probability the $MB(n,p\,;m)$ process percolates.*

Proof. We have from Lemma 5.2.1 that with high probability $I_0 = [m]$ is contained in no closed set of size less than

$$\frac{n}{2} - \frac{n\sqrt{\log d}}{2\sqrt{d}} + \frac{n(\log\log d)^{2\lambda}}{e^8\sqrt{d\log d}}.$$

Proposition 5.2.4 applied to $\varepsilon = \lambda - \frac{1}{2}$ tells us that with high probability I_0 is contained in no closed set of size between

$$\frac{n}{2} - \frac{n\sqrt{\log d}}{2\sqrt{d}} + \frac{n(\log\log d)^{\lambda+\frac{1}{2}}}{\sqrt{d\log d}} \quad \text{and} \quad \frac{n}{2} + \frac{4n}{\sqrt{d}}.$$

We have from Lemma 5.2.2 that with high probability I_0 is contained in no closed set of size greater than

$$\frac{n}{2} + \frac{7n}{2\sqrt{d}},$$

and so with high probability I_0 is not contained in any closed set in $G(n,p)$ and hence percolates. ■

5.3 Lower Bound

In this section we shall show that if $\lambda < 0$ then with high probability the $MB(n,p\,;m)$ process does not percolate. In fact, as might be expected, we shall show that with high probability the $MB(n,p\,;m)$ process terminates with I only slightly larger than m. We shall do this by bounding the expected number of sets of some size that could be the first vertices to be infected.

We say that a set of vertices T percolates if all of its vertices will be infected eventually. For $T \subset I \setminus I_0$ we can order the vertices of $I_0 \cup T$ by the time they get infected. That is, take any order of T such that a vertex from I_j is infected before any vertex from $I_{j'}$

if $j < j'$. Notice that for each $v \in T$ the majority of its neighbours (in the whole graph) are in the set of its predecessors in this order. Our strategy will be to show that if $\lambda < 0$ then with high probability there is no percolating set T of a particular size and thus the $MB(n, p\,; m)$ process does not percolate.

Assume that $|I_0| = m$, set $t = |T|$ and denote by $E = E(T)$ the edge set of T. Write $d_E(i)$ for the degree within T of a vertex $i \in T$. Condition on the edge configuration E. We want to bound the probability that T percolates. To do so, we modify the infection rule within T so that the vertices inside T consider their neighbours in T to be already infected, regardless of their real state at any particular time step. The latter assumption only increases the probability and, more importantly, makes the events for vertices in T to be infected independent. This is because these events now only depend how many edges each vertex has to I_0 and $V(G)/(I_0 \cup T)$. We thus have

$$\mathbb{P}(T \text{ percolates}) \leq \sum_E \mathbb{P}(E) \prod_{i=1}^{t} \mathbb{P}(B(m,p) + d_E(i) \geq B(n-m-t,p)). \tag{56}$$

Denote $g(x) = \mathbb{P}\left(B(m,p) + x \geq B(n-m-t,p)\right)$. Due to the log-concavity of g (Corollary 5.4.3) we have for integer x, y that

$$g(x) \leq g(y) \frac{g^{x-y}(y+1)}{g^{x-y}(y)}$$

Using the latter inequality with $x = d_E(i)$ and $y = \lceil pt \rceil$, we can bound 56 by

$$\begin{aligned}
& \sum_E \mathbb{P}(E) \prod_{i=1}^{t} g(y) \left(\frac{g(y+1)}{g(y)} \right)^{d_E(i)-y} \\
= \; & \sum_E \mathbb{P}(E) g(y)^t \left(\frac{g(y+1)}{g(y)} \right)^{2|E|-ty} \\
= \; & \sum_{j=0}^{\binom{t}{2}} \binom{\binom{t}{2}}{j} p^j (1-p)^{\binom{t}{2}-j} g(y)^t \left(\frac{g(y+1)}{g(y)} \right)^{2i-ty} \\
= \; & \left(1 - p + p \left(\frac{g(y+1)}{g(y)} \right)^2 \right)^{\binom{t}{2}} \left(\frac{g(y)}{g(y+1)} \right)^{pt^2} g(y)^t.
\end{aligned}$$

Substituting $\frac{g(y+1)}{g(y)} = 1+a$ and the elementary inequality $1/(1+a) \leq 1-a+a^2$, we bound the latter expression by

$$\begin{aligned} & \left(1-p+p(1+a)^2\right)^{\binom{t}{2}} \left(1-a+a^2\right)^{ty} g(y)^t \\ \leq \quad & \exp\left((2ap+a^2p)\frac{t^2}{2}+(a^2-a)pt^2\right) g(y)^t \\ = \quad & \left(\exp\left(\frac{3pa^2t}{2}\right) g(y)\right)^t. \end{aligned} \tag{57}$$

We have by definition that $g(y)$ is equal to

$$g(y) = \mathbb{P}\left(X_1+X_2 \geq \mu_1+\mu_2+pn-2pm-pt-\lceil pt \rceil\right),$$

where $X_1 = B(m,p)$ with mean μ_1 and $X_2 = B(n-m-t,(1-p))$ with mean μ_2. Setting $t = \left\lfloor n(\log\log d)^\lambda/\sqrt{d\log d}\right\rfloor$ and using Proposition 5.4.12 with $N = \frac{n-t}{2}$, $S = \frac{n-2m-t}{2}$ and $h = p(n-2m-t)-y$ to bound $g(y)$, we obtain

$$\begin{aligned} g(y) \quad < \quad & \frac{\sqrt{p(1-p)(n-t)}}{pn-2pm-2pt-1} \exp\left(-\frac{(pn-2pm-2pt-1)^2}{2p(1-p)(n-t)}+4\right) \\ < \quad & \frac{e^4}{\sqrt{\log d}} \exp\left(-\frac{\log d}{2}+2\lambda\log\log\log d+O((\log\log d)^\lambda)\right) \\ < \quad & \left(\frac{e^5(\log\log d)^{2\lambda}}{\sqrt{d\log d}}\right), \end{aligned} \tag{58}$$

when $\lambda < 0$.

We can also bound $g(y)$ from below by Proposition 5.4.11

$$\begin{aligned} g(y) \quad > \quad & \frac{\sqrt{p(1-p)(n-t)}}{2\pi(pn-2pm-2pt)} \exp\left(-\frac{(pn-2pm-2pt)^2}{2p(1-p)(n-t)}-4-o(1)\right) \\ > \quad & \frac{1}{2\pi e^4\sqrt{\log d}} \exp\left(-\frac{\log d}{2}+2\lambda\log\log\log d+O((\log\log d)^\lambda)\right) \\ > \quad & \left(\frac{(\log\log d)^{2\lambda}}{e^6\sqrt{d\log d}}\right), \end{aligned} \tag{59}$$

when $\lambda < 0$.

By definition of g we have that

$g(y+1) = g(y) + \mathbb{P}(B(m,p) + y + 1 = B(n-m-t,p))$. Let us write

$u = \mathbb{P}(B(m,p) + y + 1 = B(n-m-t,p))$ for convenience. We shall now obtain an upper bound for u. Using Proposition 5.4.13 with $T = -\frac{y+1}{p}$, $N = \frac{n-t+T}{2}$ and $S = N - m$, we obtain

$$\begin{aligned} u &< \frac{\frac{n}{2} - m - 2t}{2\pi(1-p)(\frac{n}{2} - 2t - \frac{2}{p})} \exp\left(-\frac{2p(\frac{n}{2} - m - 2t - \frac{2}{p})^2}{(1-p)(n-t)} + o(1)\right) \\ &\quad + \frac{3}{\pi p(\frac{n}{2} - m - 2t - \frac{2}{p})} \exp\left(-\frac{9p(\frac{n}{2} - m - 2t - \frac{2}{p})^2}{8(1-p)(\frac{n}{2} - 2t - \frac{1}{p})}\right) \\ &< \frac{\sqrt{\log d}}{2\pi(1-p)\sqrt{d}} \exp\left(-\frac{\log d}{2} + 2\lambda \log\log\log d + o(1)\right) \\ &\quad + \frac{6\sqrt{d}}{\pi p n \sqrt{\log d}} \exp\left(-\frac{9\log d}{16} + \frac{9\lambda \log\log\log d}{4} + o(1)\right). \end{aligned} \tag{60}$$

The first term is much larger than the second and so we obtain the inequality

$$u < \frac{\sqrt{\log d}(\log\log d)^{2\lambda}}{\pi(1-p)d}. \tag{61}$$

We have that $a = \frac{z}{g(y)}$ and so from (59) and (61)

$$a < \frac{e^6 \log d}{\pi(1-p)\sqrt{d}} < \frac{e^5 \log d}{(1-p)\sqrt{d}}. \tag{62}$$

We can now bound the expression in (57) by

$$\begin{aligned} \mathbb{P}(T \text{ percolates}) &< \left(\exp\left(\frac{3pe^{10}(\log d)^2 n(\log\log d)^{\lambda}}{2(1-p)^2 d^2 \log d}\right) \frac{e^5(\log\log d)^{2\lambda}}{\sqrt{d\log d}}\right)^t \\ &< \left(\frac{e^6(\log\log d)^{2\lambda}}{\sqrt{d\log d}}\right)^t. \end{aligned} \tag{63}$$

The expected number of sets of size t that percolate is

$$\begin{aligned}\binom{n-m}{t}\mathbb{P}(T \text{ percolates}) &< \binom{n}{t}\left(\frac{e^6(\log\log d)^{2\lambda}}{\sqrt{d\log d}}\right)^t \\ &< \left(\frac{e^7 n(\log\log d)^{2\lambda}}{t\sqrt{d\log d}}\right)^t, \end{aligned} \tag{64}$$

because $\binom{n}{t} \leq \left(\frac{en}{t}\right)^t$. We chose $t = \left\lfloor \frac{n(\log\log d)^{\lambda}}{\sqrt{d\log d}} \right\rfloor$, and so the expected number of sets of size t that percolate is bounded above by

$$(e^7(\log\log d)^{\lambda})^t = o(1). \tag{65}$$

Therefore with high probability percolation does not occur for $\lambda < 0$.

5.4 Inequalities

We begin this section with some remarks on the log-concavity of the distribution function of the Binomial distribution. These results are standard, see for example [25], but we prove them for completeness.

Proposition 5.4.1. *The sum of independent Bernoulli random variables is log-concave, that is if X_i are independent Bernoulli random variables with means p_i, then for any k we have,*

$$\mathbb{P}(\sum_{i=1}^{n} X_i = k-1)\mathbb{P}(\sum_{i=1}^{n} X_i = k+1) \leq (\mathbb{P}(\sum_{i=1}^{n} X_i = k))^2.$$

Proof. We proceed by induction on n, with the base case $n = 1$ being trivial as one of the terms on the left hand side of the inequality is zero. Otherwise conditioning on X_{n+1}, and writing $f_{n,k} = \mathbb{P}(\sum_{i=1}^{n} X_i = k)$ we get,

$$
\begin{aligned}
f_{n+1,k-1}f_{n+1,k+1} &= (p_{n+1}f_{n,k-2}+(1-p_{n+1})f_{n,k-1})(p_{n+1}f_{n,k}+(1-p_{n+1})f_{n,k+1}) \\
&\leq (p_{n+1}f_{n,k-1}+(1-p_{n+1})f_{n,k})^2 \\
&= (f_{n,k})^2 \qquad (66)
\end{aligned}
$$

The inequality follows as $f_{n,k-1}f_{n,k+2} \leq f_{n,k}f_{n,k+1}$ is implied by the induction hypothesis. ■

Proposition 5.4.2. *The cumulative distribution of a discrete non-negative log-concave random variable X is log-concave, that is for all k,*

$$\mathbb{P}(X \leq k-1)\mathbb{P}(X \leq k+1) \leq (\mathbb{P}(X \leq k))^2.$$

Proof. Setting $r_i = \mathbb{P}(X = i)$ we get by Proposition 5.4.1,

$$(r_0 + \ldots + r_{k-1})r_{k+1} \leq (r_1 + \ldots + r_k)r_k + r_k r_0,$$

and so,

$$(r_0 + \ldots + r_{k-1})(r_0 + \ldots + r_{k+1}) \leq (r_0 + \ldots + r_k)^2.$$

■

When X is the sum of n independent Bernoulli random variables, we can rewrite $X = n - Y$, where Y is also the sum of n independent Bernoulli random variables, and so Proposition 5.4.2 is still true if we replace $\leq$, with $<$, $>$ or $\geq$.

Corollary 5.4.3. *The cumulative distribution of the sum or difference of independent binomial random variables is log-concave.*

Proof. Sums and differences of independent binomial random variables are also sums of

independent Bernoulli random variables plus a constant and so are log-concave. ■

A substantial part of this section is now taken up with providing tight bounds, up to a constant factor, on binomial probabilities and their sums.

Proposition 5.4.4. *Suppose $pn \geq 1$ and $k = pn + h < n$, where $h > 0$. Set*

$$\beta = \frac{1}{12k} + \frac{1}{12(n-k)},$$

then $\mathbb{P}(B(n,p) = k)$ is at least

$$\frac{1}{\sqrt{2\pi p(1-p)n}} \exp\left(-\frac{h^2}{2p(1-p)n} - \frac{h^3}{2(1-p)^2n^2} - \frac{h^4}{3p^3n^3} - \frac{h}{2pn} - \beta\right).$$

Proof. This is Theorem 1.5 in [10], p. 12. ■

Corollary 5.4.5. *Suppose $p(1-p)n = \omega(n)$ and $k = pn + h$, where $0 < h = o\left((p(1-p)n)^{\frac{2}{3}}\right)$, then*

$$\mathbb{P}(B(n,p) = k) > \frac{1}{\sqrt{2\pi p(1-p)n}} \exp\left(-\frac{h^2}{2p(1-p)n} - o(1)\right)$$

Proof. For h in this range we have

$$\frac{h^3}{2(1-p)^2n^2} + \frac{h^4}{3p^3n^3} + \frac{h}{2pn} = o(1).$$

We also have that $k = \omega(n)$ and $n - k = \omega(n)$ and so the inequality follows from Proposition 5.4.4. ■

Proposition 5.4.6. *Suppose $pn \geq 1$ and $k \geq pn + h$, where $h(1-p)n \geq 3$. Then*

$$\mathbb{P}(B(n,p) = k) < \frac{1}{\sqrt{2\pi p(1-p)n}} \exp\left(-\frac{h^2}{2p(1-p)n} + \frac{h^3}{p^2n^2} + \frac{h}{(1-p)n}\right).$$

Proof. This is Theorem 1.2 of [10], p. 10. ■

Corollary 5.4.7. *Suppose* $p(1-p)n = \omega(n)$ *and* $k \geq pn+h$, *where* $1 < h = o((p(1-p)n)^{\frac{2}{3}})$, *then*

$$\mathbb{P}(B(n,p)=k) < \frac{1}{\sqrt{2\pi p(1-p)n}} \exp\left(-\frac{h^2}{2p(1-p)n} + o(1)\right).$$

Proof. For h in this range we have

$$\frac{h^3}{p^2n^2} + \frac{h}{(1-p)n} = o(1),$$

and so the inequality follows from Proposition 5.4.6, which can be applied as $h(1-p)n = \omega(n)$. ■

Proposition 5.4.8. *Suppose* $p(1-p)n = \omega(n)$ *and* $0 < h = o((p(1-p)n)^{\frac{2}{3}})$, *then*

$$\mathbb{P}(B(n,p) \geq pn+h) < \frac{\sqrt{p(1-p)n}}{\sqrt{2\pi}h} \exp\left(-\frac{h^2}{2p(1-p)n} + o(1)\right).$$

Proof. This proof follows that of Theorem 1.3 in [10]. For $m \geq pn+h$, we have

$$\frac{\mathbb{P}(B(n,p)=m+1)}{\mathbb{P}(B(n,p)=m)} \leq 1 - \frac{h+(1-p)}{(1-p)(pn+h+1)} = \lambda.$$

Hence

$$\mathbb{P}(B(n,p) \geq pn+h) \leq \frac{1}{1-\lambda}\mathbb{P}(B(n,p) = \lceil pn+h \rceil).$$

As $(1-\lambda)^{-1} < \frac{p(1-p)n}{h}(1+\frac{h}{pn}) < \frac{p(1-p)n}{h}e^{\frac{h}{pm}}$, we get from Proposition 5.4.6 that

$$\mathbb{P}(B(n,p) \geq pn+h) < \frac{\sqrt{p(1-p)n}}{h\sqrt{2\pi}} \exp\left(-\frac{h^2}{2p(1-p)n} + \frac{h}{p(1-p)n} + \frac{h^3}{p^2n^2}\right)$$

the last two terms in the exponent being $o(1)$, for $h = o(p(1-p)n)^{\frac{2}{3}}$. ■

Proposition 5.4.9. *Suppose* $p(1-p)n = \omega(n)$ *and* $(p(1-p)n)^{\frac{1}{2}} < h = o((p(1-p)n)^{\frac{2}{3}})$, *then*

$$\mathbb{P}(B(n,p) \geq pn+h) > \frac{\sqrt{p(1-p)n}}{h\sqrt{2\pi}} \exp\left(-\frac{h^2}{2p(1-p)n} - \frac{3}{2} - o(1)\right).$$

Proof. Due to the unimodality of the binomial distribution, we have that the probability density function of the binomial distribution is decreasing away from its mean and so,

$$\mathbb{P}(B(n,p) \geq pn+h) > \frac{p(1-p)n}{h}\mathbb{P}(B(n,p) = pn+h+\frac{p(1-p)n}{h}).$$

We can apply Corollary 5.4.5 as $h + \frac{p(1-p)n}{h} = o((p(1-p)n)^{\frac{2}{3}})$ and so it follows that

$$\mathbb{P}(B(n,p) \geq pn+h) > \frac{\sqrt{p(1-p)n}}{h\sqrt{2\pi}} \exp\left(-\frac{(h+\frac{p(1-p)n}{h})^2}{2p(1-p)n} - o(1)\right).$$

This is greater than the stated bound because $(h+\frac{p(1-p)n}{h})^2 \leq h^2 + 3p(1-p)n$. ■

We shall also want a weaker but more general bound than Proposition 5.4.8 due to Bernstein in [7].

Lemma 5.4.10. *Let* $X_1, \ldots, X_n$ *be independent zero-mean random variables. Suppose that* $|X_i| \leq M$, *then for all positive* t,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i > t\right) \leq \exp\left(-\frac{t^2}{2\sum \mathbb{E}(X_j^2) + \frac{2}{3}Mt}\right).$$

Proof. For a proof see [15]. ■

Proposition 5.4.11. *Suppose that* $p(1-p)N = \omega(N)$, *the inequality* $2(2p(1-p)N)^{\frac{1}{2}} < h = o((p(1-p)N)^{\frac{2}{3}})$ *holds and* $hS = o(N((p(1-p)N)^{\frac{1}{2}}))$. *For the independent random variables;* $X_1 = B(N-S,p)$, *with mean* μ_1 *and variance* σ_1^2,*; and* $X_2 = B(N+S,(1-p))$ *with mean* μ_2 *and variance* σ_2^2, *we have*

$$\mathbb{P}(X_1+X_2 \geq \mu_1+\mu_2+h) > \frac{\sqrt{2p(1-p)N}}{2\pi h}\exp\left(-\frac{h^2}{4p(1-p)N}-4-o(1)\right).$$

Proof. The conditions on S and h imply that $S=o(N)$. Set z and l equal to $\frac{2p(1-p)N}{h}$ and $\left\lfloor \frac{h}{\sqrt{2p(1-p)N}} \right\rfloor$ respectively. We can bound $\mathbb{P}(X_1+X_2 \geq \mu_1+\mu_2+h)$ from below by summing over the disjoint regions

$$\sum_{i=-l}^{l-1} \mathbb{P}\left(X_1 < \mu_1+\frac{h}{2}-iz, X_2 < \mu_2+\frac{h}{2}+(i+1)z, X_1+X_2 \geq \mu_1+\mu_2+h\right). \tag{67}$$

For each i there are at least $\lfloor z \rfloor(\lfloor z \rfloor - 1)/2$ pairs of integer values x_1, x_2, which X_1, X_2 can take while still satisfying all three relations in (67). We have that $h > 2lz$, and so if X_1, X_2 satisfy all three relation in (67), then $X_1 \geq \mu_1$ and $X_2 \geq \mu_2$. As we are only considering the region in which X_1, X_2 are larger than their means we can bound the sum in (67) from below by

$$\sum_{i=-l}^{l-1} \frac{\lfloor z \rfloor(\lfloor z \rfloor-1)}{2}\mathbb{P}\left(X_1 = \left\lceil \mu_1+\frac{h}{2}-iz \right\rceil\right)\mathbb{P}\left(X_2 = \left\lceil \mu_2+\frac{h}{2}+(i+1)z \right\rceil\right). \tag{68}$$

We have that $p(1-p)(N-S) = \omega(N)$ and $h+lz = o(p(1-p)(N-S))^{\frac{2}{3}}$, and so we can apply Corollary 5.4.5 to get that the quantity in (68) is at least

$$\sum_{i=-1}^{l-1} \frac{\lfloor z \rfloor(\lfloor z \rfloor-1)}{4\pi\sigma_1\sigma_2}\exp\left(-\frac{(\frac{h}{2}-iz+1)^2(N+S)+(\frac{h}{2}+(i+1)z+1)^2(N-S)}{2p(1-p)(N^2-S^2)}-o(1)\right).$$

Expanding this out, and noticing $\lfloor z \rfloor = z(1+o(1))$ and $(N-S)(N+S) = N^2(1+o(1))$ we get that the sum in (68) is at least

$$\sum_{i=-l}^{l-1} \frac{z^2}{4\pi p(1-p)N}\exp\left(-\frac{h^2N+2hz(N-S)+4i^2z^2N+(4i+2)z^2(N-S)+o(p(1-p)N^2)}{4p(1-p)(N^2-S^2)}-o(1)\right),$$

where the approximations for $\lfloor z \rfloor$ and σ_1, σ_2 have been taken care of in the $o(1)$ in the

exponential term. Using the bounds in the statement of the Proposition this is at least

$$\begin{aligned}&\frac{lz^2}{2\pi p(1-p)N}\exp\left(-\frac{h^2N}{4p(1-p)(N^2-S^2)}-4-o(1)\right)\\ &> \frac{\sqrt{2p(1-p)N}}{2\pi h}\exp\left(-\frac{h^2}{4p(1-p)N}-4-o(1)\right).\end{aligned} \tag{69}$$

The inequality following because $l > h/(2\sqrt{2p(1-p)N})$ and $hS = o(N(p(1-p)N)^{\frac{1}{2}})$. ■

Proposition 5.4.12. *Suppose that $p(1-p)N = \omega(N)$. Furthermore assume that*

$$2(2p(1-p)N)^{\frac{1}{2}} < h = o((p(1-p)N)^{\frac{2}{3}}) \quad \textit{and} \quad Sh = o(N(p(1-p)N)^{\frac{1}{2}}).$$

Then we have

$$\mathbb{P}(X_1+X_2 \geq \mu_1+\mu_2+h) < \frac{\sqrt{2p(1-p)N}}{h}\exp\left(-\frac{h^2}{4p(1-p)N}+4\right),$$

where $X_1 = B(N-S,p)$ with mean μ_1 and variance σ_1^2, and $X_2 = B(N+S,(1-p))$ with mean μ_2 and variance σ_2^2.

Proof. The conditions on S and h imply that $S = o(N)$. Set $z = \frac{2Np(1-p)}{h}$, and $l = \left\lfloor \frac{h^2}{4Np(1-p)} \right\rfloor$. We bound $\mathbb{P}(X_1+X_2 \geq \mu_1+\mu_2+h)$ from below by covering the region where this inequality holds by

$$\mathbb{P}(X_1+X_2 \geq \mu_1+\mu_2+h) < \tag{70}$$

$$\sum_{i+j\geq -1}^{-l\leq i,j\leq l-1}\left(\mathbb{P}\left(0 \leq X_1-\mu_1-\frac{h}{2}-iz < z\right)\mathbb{P}\left(0 \leq X_2-\mu_2-\frac{h}{2}-jz < z\right)\right) \tag{71}$$

$$+\mathbb{P}\left(X_1 \geq \mu_1+\frac{h}{2}+lz\right) \tag{72}$$

$$+\mathbb{P}\left(X_2 \geq \mu_2+\frac{h}{2}+lz\right). \tag{73}$$

We shall bound these three summands separately. Firstly for each i, j pair there are at most $\lceil z \rceil^2$ points inside the specified region and so the product inside the sum of (71) is at most

$$\lceil z \rceil^2 \mathbb{P}\left(X_1 = \left\lceil \mu_1 + \frac{h}{2} + iz \right\rceil\right) \mathbb{P}\left(X_2 = \left\lceil \mu_2 + \frac{h}{2} + jz \right\rceil\right).$$

Again because $h > 2lz$ we are only considering the range in which X_1 and X_2 are greater than their means. We have that $p(1-p)(N \pm S) = \omega(N)$ and $1 < h \pm lz = o(p(1-p)(N \pm S))^{\frac{2}{3}}$ and so we can apply Corollary 5.4.7 to get that the sum in (71) is at most

$$\sum_{i+j \geq -1}^{-l \leq i,j \leq l-1} \frac{\lceil z \rceil^2}{2\pi p(1-p)\sqrt{N^2 - S^2}} \cdot \exp\left(-\frac{\left(\frac{h}{2} + iz\right)^2 (N+S) + \left(\frac{h}{2} + jz\right)^2 (N-S)}{2p(1-p)(N^2 - S^2)} + o(1)\right). \tag{74}$$

This is equal to

$$\frac{z^2}{2\pi p(1-p)N} \exp\left(-\frac{h^2 N}{4p(1-p)(N^2 - S^2)} + o(1)\right) \sum_{i+j \geq -1}^{-l \leq i,j < l} \exp\left(-\frac{h(i+j)zN + hzs(i-j) + z^2 N(i^2 + j^2) + z^2 S(i^2 - j^2)}{2p(1-p)(N^2 - S^2)}\right). \tag{75}$$

We can bound the above by noting that $|i - j| \leq \sqrt{2}\sqrt{i^2 + j^2}$ and $|i^2 - j^2| \leq i^2 + j^2$. As we also have that $Np(1-p)/2 < z^2 l \leq Np(1-p)$, the sum appearing in 75 is at most

$$\sum_{i+j \geq -1}^{-l \leq i,j < l} \exp\left(-(i+j) + \sqrt{\frac{i^2 + j^2}{4l}} - \frac{i^2 + j^2}{4l} + o(1)\right). \tag{76}$$

A point (i,j) in the plane with integer coordinates and $\frac{i^2+j^2}{4l} - \sqrt{\frac{i^2+j^2}{4l}} < t$, also satisfies $|i-j| < \sqrt{21tl}$. Therefore the number of points (i,j) in the plane with integer coordinates and satisfying both $\frac{i^2+j^2}{4l} - \sqrt{\frac{i^2+j^2}{4l}} < t$, and $-1 \le i+j < t$ is at most $2(t+1)\sqrt{21lt}$. This allows us crudely bound (76) by

$$2\sqrt{21l}\sum_{t=1}^{\infty}(t+1)\sqrt{t}\exp\left(-(t-1)\right). \tag{77}$$

The latter sum is less than $50\sqrt{l}$ and so the sum in (71) is bounded above by

$$\frac{50\sqrt{p(1-p)N}}{h\pi}\exp\left(-\frac{h^2}{4p(1-p)N}+o(1)\right). \tag{78}$$

Secondly we bound the probability (72). As $l > \frac{h^2}{8Np(1-p)}$ we have that

$$\mathbb{P}\left(X_1 \ge \mu_1 + \frac{h}{2} + lz\right) < \mathbb{P}\left(X_1 \ge \mu_1 + \frac{3h}{4}\right).$$

By Proposition 5.4.8 we get that the quantity in (72) is at most

$$\begin{aligned} &\frac{4\sqrt{p(1-p)(N-S)}}{3h\sqrt{2\pi}}\exp\left(-\frac{9h^2}{32p(1-p)(N-S)}+o(1)\right) \\ < \quad &\frac{2}{3\sqrt{\pi}}\frac{\sqrt{2p(1-p)N}}{h}\exp\left(-\frac{h^2}{4p(1-p)N}\right). \end{aligned} \tag{79}$$

Similarly, the probability in (73) is at most

$$\frac{2}{3\sqrt{\pi}}\frac{\sqrt{2p(1-p)N}}{h}\exp\left(-\frac{h^2}{4p(1-p)N}\right). \tag{80}$$

As $\frac{50\sqrt{2}}{\pi} + \frac{4}{3\sqrt{\pi}} < e^4$ we get that the sum of our three bounds, (78), (79), and (80) is at most the stated bound. ■

Proposition 5.4.13. *Suppose that $p(1-p)N = \omega(N)$, that*

$\omega(N)(p(1-p)N)^{\frac{1}{2}} \leq pS = o((p(1-p)N)^{\frac{2}{3}})$ *and that* $T = o(N)$*, then*

$$\begin{aligned}\mathbb{P}(Z_1 = Z_2 + pT) \quad &< \quad \frac{S}{2\pi(1-p)N}\exp\left(-\frac{2pS^2}{(1-p)(2N-T)} + o(1)\right) \\ &\quad + \frac{3}{\pi pS}\exp\left(-\frac{9pS^2}{8(1-p)N}\right), \end{aligned} \tag{81}$$

where $Z_1 = B(N-S,p)$ *with mean* μ_1 *and variance* σ_1^2 *and* $Z_2 = B(N+S-T,p)$ *with mean* μ_2 *and variance* σ_2^2*.*

Proof. Let $\phi(i)$ be the probability that $Z_1 = Z_2 + pT = pN + i$, then

$$\phi(i) = \binom{N-S}{pN+i}\binom{N+S-T}{pN-pT+i} p^{p(2N-T)+2i}(1-p)^{(1-p)(2N-T)-2i}.$$

Denote the ratio between successive values of $\phi(i)$ by $\psi(i)$. We obtain

$$\begin{aligned}\psi(i) \quad &= \quad \frac{\phi(i+1)}{\phi(i)} = \frac{p^2((1-p)N-S-i)((1-p)(N-T)+S-i)}{(1-p)^2(pN+i+1)(p(N-T)+i+1)} \\ &= \quad \frac{\left(1-\frac{S+i}{(1-p)N}\right)\left(1+\frac{S-i}{(1-p)(N-T)}\right)}{\left(1+\frac{i+1}{pN}\right)\left(1+\frac{i+1}{p(N-T)}\right)}, \end{aligned} \tag{82}$$

and so ψ is a decreasing function of i. By noting that $e^{x-x^2} \leq (1+x) \leq e^x$, for $x \geq -\frac{1}{2}$, we can bound ψ for $i = o(p(1-p)N)$ between

$$\exp\left(\frac{pST-(2N-T)(i+1-p)}{p(1-p)N(N-T)} - \left(\frac{S+i}{(1-p)N}\right)^2 - \left(\frac{S-i}{(1-p)(N-T)}\right)^2\right)$$

and

$$\exp\left(\frac{pST-(2N-T)(i+1-p)}{p(1-p)N(N-T)} + \left(\frac{i+1}{pN}\right)^2 + \left(\frac{i+1}{p(N-T)}\right)^2\right).$$

Substituting in $i = \pm\frac{pS}{2}$, we get that

$$\begin{aligned}
\psi(\frac{pS}{2}) &< \exp\left(-\left(\frac{(2N-3T)S}{2(1-p)N(N-T)}\right)(1+o(1))\right) \\
&< \exp\left(-\frac{(2N-3T)S}{4(1-p)N(N-T)}\right) \\
&< 1-\frac{S}{3(1-p)N}
\end{aligned} \tag{83}$$

and

$$\begin{aligned}
\psi(-\frac{pS}{2}) &> \exp\left(\left(\frac{(2N+T)S}{2(1-p)N(N-T)}\right)(1+o(1))\right) \\
&> \exp\left(\frac{(2N+t)S}{4(1-p)N(N-T)}\right) \\
&> 1+\frac{S}{3(1-p)N}.
\end{aligned} \tag{84}$$

Therefore ψ is greater than 1 at $i = pN - \frac{pS}{2}$ and less than 1 at $i = pN + \frac{pS}{2}$. Consequently, the maximum value of ϕ occurs between these two values.

We have that

$$\phi(i) = \mathbb{P}(Z_1 = \mu_1 + pS + i)\mathbb{P}(Z_2' = \mu_2' + pS - i),$$

where $Z_2' = N + S - T - Z_2 = B(N+S-T,(1-p))$ with mean μ_2' and variance $(\sigma_2')^2$. By Corollary 5.4.7 we get that

$$\phi(i) < \frac{1}{2\pi\sigma_1\sigma_2'}\exp\left(-\frac{(pS+i)^2(N+S-T)+(pS-i)^2(N-S)}{2p(1-p)(N-S)(N+S-T)}+o(1)\right),$$

for $|i| \leq \frac{pS}{2}$. This is maximized when $i = \frac{pST-2pS^2}{2N-T}$ and there takes the value

$$\begin{aligned}
&\frac{1}{2\pi p(1-p)N}\exp\left(-\frac{pS^2((2N-T)^2-(T-2S)^2)}{2(1-p)(N-S)(N+S-T)(2N-T)}+o(1)\right) \\
&\qquad = \frac{1}{2\pi p(1-p)N}\exp\left(-\frac{2pS^2}{(1-p)(2N-T)}+o(1)\right).
\end{aligned} \tag{85}$$

We also obtain the bounds

$$\begin{aligned}\phi(\frac{pS}{2}) &< \frac{1}{2p(1-p)\pi N}\exp\left(-\frac{pS^2(10N+8S-9T)}{8(1-p)(N-S)(N+S-T)}+o(1)\right)\\ &< \frac{1}{2p(1-p)\pi N}\exp\left(-\frac{9pS^2}{8(1-p)N}\right)\end{aligned} \tag{86}$$

and

$$\begin{aligned}\phi(\frac{-pS}{2}) &< \frac{1}{2p(1-p)\pi N}\exp\left(-\frac{pS^2(10N-8S-T)}{8(1-p)(N-S)(N+S-T)}+o(1)\right)\\ &< \frac{1}{2p(1-p)\pi N}\exp\left(-\frac{9pS^2}{8(1-p)N}\right)\end{aligned} \tag{87}$$

Putting this all together we obtain

$$\begin{aligned}\mathbb{P}(Z_1=Z_2=pT) &< pS\max_i\phi(i)+\frac{1}{1-\psi(\frac{pS}{2})}\phi(\frac{pS}{2})+\frac{\psi(\frac{-pS}{2})}{\psi(\frac{-pS}{2})-1}\phi(\frac{-pS}{2})\\ &< \frac{S}{2\pi(1-p)N}\exp\left(-\frac{2pS^2}{(1-p)(2N-T)}+o(1)\right)\\ &\quad+\frac{3}{p\pi S}\exp\left(-\frac{9pS^2}{8(1-p)N}\right)\end{aligned} \tag{88}$$

■

We end with some propositions about the number of edges in and between sets in $G(n,p)$.

Proposition 5.4.14. *Suppose that $p(1-p)n=\omega(n)$. If n is large enough, then for all $t>\frac{n}{5}$ we have that with probability at least $1-4^{-t}$, every set in $G(n,p)$ of size t has at most $p\binom{t}{2}+2t\sqrt{p(1-p)t}$ edges.*

Proof. The expected number of sets of size t with more than $p\binom{t}{2}+2t\sqrt{p(1-p)t}$ edges is

$$\binom{n}{t}\mathbb{P}\left(B\left(\binom{t}{2},p\right)\geq p\binom{t}{2}+2t\sqrt{p(1-p)t}\right).$$

By Lemma 5.4.10 and the fact that $\binom{n}{t} \leq (\frac{en}{t})^t$, this expectation is at most

$$(5e)^t \exp\left(-\frac{4p(1-p)t^3}{2p(1-p)\binom{t}{2} + \frac{4t\sqrt{p(1-p)t}}{3}}\right).$$

As $\sqrt{p(1-p)t} = \omega(n)$, we have that if n is large enough, then for all $t > \frac{n}{5}$ we have

$$2p(1-p)\binom{t}{2} + \frac{4t\sqrt{p(1-p)t}}{3} \leq 1.001p(1-p)t^2.$$

Substituting this in we have that the expected number of sets of size t with more than $p\binom{t}{2} + 2t\sqrt{p(1-p)t}$ edges is at most,

$$\exp\left(t(\log 5 + 1) - \frac{4p(1-p)t^3}{p(1-p)t^2}(1+o(1))\right) < 4^{-t}.$$

■

Proposition 5.4.15. *Suppose that $p(1-p)n = \omega(n)$. If n is large enough then for all t in the range $\frac{n}{5} < t \leq \frac{n}{2}$ we have that with probability at least $1-4^{-t}$ every set in $G(n,p)$ of size t has at least $pt(n-t) - 3t\sqrt{p(1-p)(n-t)}$ edges between it and its complement.*

Proof. The expected number of sets T of size t with less than $pt(n-t) - 3t\sqrt{p(1-p)(n-t)}$ edges between T and $[n] \setminus T$ is

$$\binom{n}{t}\mathbb{P}\left(B(t(n-t),(1-p)) \geq (1-p)t(n-t) + 3t\sqrt{p(1-p)(n-t)}\right).$$

By Lemma 5.4.10 and the fact that $\binom{n}{t} \leq (\frac{en}{t})^t$, this expectation is at most

$$(5e)^t \exp\left(-\frac{9p(1-p)t^2(n-t)}{2p(1-p)t(n-t) + \frac{4t\sqrt{p(1-p)(n-t)}}{3}}\right).$$

As $\sqrt{(n-t)p(1-p)} = \omega(n)$, we have that if n is large enough, then for all t in the

range $\frac{n}{5} < t \leq \frac{n}{2}$,

$$2p(1-p)t(n-t) + \frac{4t\sqrt{p(1-p)(n-t)}}{3} \leq \frac{9}{2}p(1-p)t(n-t).$$

Substituting this in we have that the expected number of sets T with a small number of edges between T and $[n] \setminus T$ is

$$\exp\left(t(\log 5 + 1) - \frac{9t}{2}(1+o(1))\right) < 4^{-t}.$$

■

Proposition 5.4.16. *Suppose that $p(1-p)n \geq 4\log n$. If n is large enough, then for all $t \leq \frac{n}{5}$ we have that with probability at least $1 - n^{-\frac{t}{120}}$, for every set T in $G(n,p)$ of size t there are at least twice as many edges between T and $[n] \setminus T$ as there are in T.*

Proof. The expected number of sets T of size t such that there are less than twice as many edges between T and $[n] \setminus T$ as there are in T is

$$\binom{n}{t}\mathbb{P}\left(B\left(t(n-t),p\right) < 2B\left(\binom{t}{2},p\right)\right).$$

We can rewrite this as,

$$\binom{n}{t}\mathbb{P}\left(2B\left(\binom{t}{2},p\right) - pt(t-1) - B(t(n-t),p) + pt(n-t) > pt(n-2t+1)\right).$$

By Lemma 5.4.10, this is at most

$$\binom{n}{t}\exp\left(-\frac{(pt(n-2t+1))^2}{2p(1-p)t(n+t-2) + \frac{4pt(n-2t-1)}{3}}\right). \tag{89}$$

For $t < \frac{n}{24}$, using the inequality $\binom{n}{t} \leq n^t$ we have that the quantity in (89) is at most

$$n^t \exp\left(-\frac{pt(\frac{11n}{12})^2}{\frac{10n}{3}}\right) < n^{-\frac{t}{120}}.$$

For $t \geq \frac{n}{24}$, using the inequality $\binom{n}{t} \leq \left(\frac{en}{t}\right)^t$ we have that the quantity in (89) is less than,

$$\left(\frac{en}{t}\right)^t \exp\left(-\frac{pt(\frac{3n}{5})^2}{\frac{10n}{3}}\right) < \left(\frac{24e}{n^{\frac{2}{5}}}\right)^t < n^{-\frac{t}{120}}.$$

■

Corollary 5.4.17. *Suppose that $pn \geq \log n$. If n is large enough, then for all t satisfying $n^{\frac{24}{25}} \leq t \leq \frac{n}{5}$, we have that with probability at least $1 - n^{\frac{-t}{120}}$, for every set T in $G(n,p)$ of size t, there are at least twice as many edges between T and $[n] \setminus T$ than there are in T.*

Proof. By the exact same reasoning as in Proposition 5.4.16 the expected number of sets T of size t with less than twice as many edges between T and $[n] \setminus T$ than there are in T is at most

$$\left(\frac{en}{t}\right)^t \exp\left(-\frac{pt\left(\frac{3n}{5}\right)^2}{\frac{10n}{3}}\right) < \left(\frac{e}{n^{\frac{17}{250}}}\right)^t < n^{-\frac{t}{120}}.$$

■

Proposition 5.4.18. *For every fixed ε and $p \geq \frac{(1+\varepsilon)\log n}{n}$ with high probability the minimal degree of $G(n,p)$ is greater than 8.*

Proof. The expected number of vertices with degree at most 8 is equal to

$$\begin{aligned} n\mathbb{P}(B(n-1,p) \leq 8) &= n\sum_{i=0}^{8}\binom{n-1}{i}p^i(1-p)^{n-1-i} \\ &\leq n\left(\binom{n-1}{8}p^8(1-p)^{n-9}\left(1+\frac{9(1-p)}{p(n-9)}+\left(\frac{9(1-p)}{p(n-9)}\right)^2+\dots\right)\right) \\ &\leq \frac{9n^9}{8!}p^8(1-p)^{n-9}. \end{aligned} \tag{90}$$

These inequalities follow as $\max_{i\leq 8}\mathbb{P}(B(n-1,p)=i)$ occurs when $i=8$ and so $\mathbb{P}(B(n-1,p)\leq 8)\leq 9\mathbb{P}(B(n-1,p)=8)$. The last line of (90) is maximised over $0\leq p\leq 1$ when $\frac{p}{8}=\frac{1-p}{n-9}$, that is when $p=\frac{8}{n-1}$. So for p in our range, (90) is maximised when $p=\frac{(1+\varepsilon)\log n}{n}$. Therefore

$$\begin{aligned} n\mathbb{P}(B(n-1,p)\leq 8) &\leq \frac{9n^9(1+\varepsilon)^8(\log n)^8}{8!n^8}e^{-\frac{(n-9)(1+\varepsilon)\log n}{n}} \\ &\leq \frac{(\log n)^8}{n^{\varepsilon}}. \end{aligned} \tag{91}$$

■

Proposition 5.4.19. *Suppose that* $(1+\varepsilon)\log n\leq pn\leq 5\log n$. *If* n *is large enough, then for all* t *satisfying* $t\leq n^{\frac{29}{30}}$, *we have that with probability at least* $1-n^{-\frac{t}{120}}$, *every set in* $G(n,p)$ *of size* t *has at most* $2t$ *edges.*

Proof. The expected number of sets T in $G(n,p)$ of size t with at least $2t$ edges is

$$\binom{n}{t}\mathbb{P}\left(B\left(\binom{t}{2},p\right)\geq 2t\right)=\binom{n}{t}\sum_{i=2t}^{n}\binom{\binom{t}{2}}{i}p^i(1-p)^{\binom{t}{2}-i} \tag{92}$$

By carefully bounding the summands in (92) for $i=2t$ and $i=2t+1$, we shall get a good bound on the total sum.

$$\binom{\binom{t}{2}}{2t}p^{2t}(1-p)^{\binom{t}{2}-2t}<\left(\frac{ep(t-1)}{4}\right)^{2t}<\left(\frac{5et\log n}{4n}\right)^{2t}.$$

We also obtain

$$\frac{\binom{\binom{t}{2}}{2t+1}p^{2t+1}(1-p)^{\binom{t}{2}-2t-1}}{\binom{\binom{t}{2}}{2t}p^{2t}(1-p)^{\binom{t}{2}-2t}}=\frac{p(\binom{t}{2}-2t)}{(1-p)(2t+1)}\leq pt<\frac{1}{2}.$$

Therefore

$$\begin{aligned}
\binom{n}{t}\mathbb{P}\left(B\left(\binom{t}{2},p\right)\geq 2t\right) &\leq \binom{n}{t}2\left(\frac{5et\log n}{4n}\right)^{2t} \\
&\leq 2\frac{e^{3t}25^t(\log n)^{2t}t^t}{16^t n^t} \\
&\leq \left(\frac{C(\log n)^2}{n^{\frac{1}{30}}}\right)^t, \qquad (93)
\end{aligned}$$

and so the expected number of set T in $G(n,p)$ of size t with at least $2t$ edges is at most $n^{-\frac{t}{120}}$. ■

REFERENCES

[1] AIZENMAN, M., AND LEBOWITZ, J. L. Metastability effects in bootstrap percolation. *Journal of Physics A: Mathematical and General 21*, 19 (1988), 3801.

[2] ANDERSON, I. *Combinatorics of finite sets*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1987.

[3] BALOGH, J., BOLLOBÁS, B., AND MORRIS, R. Majority bootstrap percolation on the hypercube. *Combinatorics, Probability and Computing 18*, 1-2 (2009), 17–51.

[4] BALOGH, J., AND PITTEL, B. G. Bootstrap percolation on the random regular graph. *Random Structures & Algorithms 30*, 1-2 (2007), 257–286.

[5] BENTKUS, V. An inequality for large deviation probabilities of sums of bounded i.i.d. random variables. *Liet. Mat. Rink. 41*, 2 (2001), 144–153.

[6] BENTKUS, V. On measure concentration for separately Lipschitz functions in product spaces. *Israel J. Math. 158* (2007), 1–17.

[7] BERNSTEIN, S. Sur une modification de linéqualité de tchebichef. *Annal. Sci. Inst. Sav. Ukr. Sect. Math. I* (1924), 38–49.

[8] BOBKOV, S. G., AND CHISTYAKOV, G. P. Bounds for the maximum of the density of the sum of independent random variables. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 408*, Veroyatnost i Statistika. 18 (2012), 62–73, 324.

[9] BOLLOBÁS, B. *Combinatorics*. Cambridge University Press, Cambridge, 1986. Set systems, hypergraphs, families of vectors and combinatorial probability.

[10] BOLLOBÁS, B. *Random graphs*, vol. 73. Cambridge university press, 2001.

[11] BOLLOBAS, B., AND RIORDAN, O. Percolation on self-dual polygon configurations. In *An Irregular Mind*. Springer, 2010, pp. 131–217.

[12] BRASCAMP, H. J., LIEB, E. H., AND LUTTINGER, J. M. A general rearrangement inequality for multiple integrals. *J. Functional Analysis 17* (1974), 227–237.

[13] BROADBENT, S. R., AND HAMMERSLEY, J. M. Percolation processes i. crystals and mazes. In *Proc. Cambridge Philos. Soc* (1957), vol. 53, pp. 629–41.

[14] CHALUPA, J., LEATH, P., AND REICH, G. Bootstrap percolation on a bethe lattice. *Journal of Physics C: Solid State Physics 12*, 1 (1979), L31.

[15] CRAIG, C. C. On the tchebychef inequality of bernstein. *The Annals of Mathematical Statistics 4*, 2 (1933), 94–102.

[16] DZINDZALIETA, D., JUŠKEVIČIUS, T., AND ŠILEIKIS, M. Optimal probability inequalities for random walks related to problems in extremal combinatorics. *SIAM J. Discrete Math. 26*, 2 (2012), 828–837.

[17] ERDÖS, P. On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc. 51* (1945), 898–902.

[18] ESSEEN, C. G. On the Kolmogorov-Rogozin inequality for the concentration function. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 5* (1966), 210–216.

[19] HALÁSZ, G. Estimates for the concentration function of combinatorial number theory and probability. *Period. Math. Hungar. 8*, 3-4 (1977), 197–211.

[20] HARDY, G. H., LITTLEWOOD, J. E., AND PÓLYA, G. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.

[21] HARRIS, T. E. A lower bound for the critical probability in a certain percolation process. In *Mathematical Proceedings of the Cambridge Philosophical Society* (1960), vol. 56, Cambridge Univ Press, pp. 13–20.

[22] HOEFFDING, W. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc. 58* (1963), 13–30.

[23] JANSON, S., ŁUCZAK, T., TUROVA, T., AND VALLIER, T. Bootstrap percolation on the random graph $g\_\{n,p\}$. *The Annals of Applied Probability 22*, 5 (2012), 1989–2047.

[24] KATONA, G. Intersection theorems for systems of finite sets. *Acta Math. Acad. Sci. Hungar 15* (1964), 329–337.

[25] KEILSON, J., AND GERBER, H. Some results for discrete unimodality. *Journal of the American Statistical Association 66*, 334 (1971), 386–389.

[26] KESTEN, H. A sharper form of the Doeblin-Lévy-Kolmogorov-Rogozin inequality for concentration functions. *Math. Scand. 25* (1969), 133–144.

[27] KESTEN, H. The critical probability of bond percolation on the square lattice equals 1/2. *Communications in Mathematical Physics 74*, 1 (1980), 41–59.

[28] KLEITMAN, D. J. On a combinatorial conjecture of Erdős. *J. Combinatorial Theory 1* (1966), 209–214.

[29] KLEITMAN, D. J. On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors. *Advances in Math. 5* (1970), 155–157 (1970).

[30] KOLMOGOROV, A. Sur les propriétés des fonctions de concentrations de M. P. Lévy. *Ann. Inst. H. Poincaré 16* (1958), 27–34.

[31] LEADER, I., AND RADCLIFFE, A. J. Littlewood-Offord inequalities for random variables. *SIAM J. Discrete Math. 7*, 1 (1994), 90–101.

[32] LITTLEWOOD, J. E., AND OFFORD, A. C. On the number of real roots of a random algebraic equation. III. *Rec. Math. [Mat. Sbornik] N.S. 12(54)* (1943), 277–286.

[33] MILNER, E. C. A combinatorial theorem on systems of sets. *J. London Math. Soc. 43* (1968), 204–206.

[34] NEUMANN, J. V. *Theory of Self-Reproducing Automata*. University of Illinois Press, Champaign, IL, USA, 1966.

[35] ROGOZIN, B. A. An estimate of the concentration functions. *Teor. Verojatnost. i Primenen 6* (1961), 103–105.

[36] ROGOZIN, B. A. An estimate for the maximum of the convolution of bounded densities. *Teor. Veroyatnost. i Primenen. 32*, 1 (1987), 53–61.

[37] SZTENCEL, R. On boundedness and convergence of some Banach space valued random series. *Probab. Math. Statist. 2*, 1 (1981), 83–88.

[38] TUCKER, H. G. On quasi-convergence of series of independent random variables. *Proc. Amer. Math. Soc. 16* (1965), 435–439.

[39] WIERMAN, J. C. Bond percolation on honeycomb and triangular lattices. *Advances in Applied Probability* (1981), 298–313.

[40] WIERMAN, J. C. A bond percolation critical probability determination based on the star-triangle transformation. *Journal of Physics A: Mathematical and General 17*, 7 (1984), 1525.