DISCRETE LOCAL CENTRAL LIMIT THEOREMS AND BOOLEAN FUNCTION COMPLEXITY MEASURES

BY JUSTIN GILMER

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Michael Saks

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January, 2015

ABSTRACT OF THE DISSERTATION

Discrete Local Central Limit Theorems and Boolean Function Complexity Measures

by JUSTIN GILMER

Dissertation Director: Michael Saks

This thesis consists of 6 chapters (the first being an introduction). Two chapters relate to local central limit theorems, and three chapters relate to various boolean function complexity measures. Although the problems studied in this work originate from different areas of mathematics, the methods used to attack these problems are unified in their probabilistic and combinatorial nature.

In Chapter 2 we prove a local central limit theorem for the number of triangles in the Erdos-Renyi random graph $G(n, p)$ for fixed $p \in (0, 1)$. In Chapter 6 we apply an existing local limit theorem for sums of independent random variables to estimate the density of a certain set of integers called *happy numbers*.

In Chapters 3, 4, and 5 we will investigate the general question of how large one complexity measure of boolean functions can be relative to another. In one case we present a probabilistic construction of family of boolean functions which show tight (in the sense that there is a matching upper bound) separation between two measures, namely *block sensitivity* and *certificate complexity*. We also give partial results for upper bounding one measure in terms of another. This includes a new approach to the well known *sensitivity conjecture* which asserts that the *degree* of any boolean function is bounded above by some fixed power of its *sensitivity*.

Acknowledgements

I am deeply grateful for all those that have helped and inspired me throughout my development as a mathematician. First off to Steve Davis, the best Calculus teacher a high school student could hope for. He was the first teacher to turn my interest in mathematics into a passion. At Washington University in St. Louis I had two terrific mentors in professors John Shareshian and Guido Weiss who helped mentor me during my time as a college student.

My deepest gratitude for my adviser Michael Saks, who was absolutely terrific. I am very thankful for his willingness, immediately upon completion of my oral exam, to invite me to collaborate with him and Srikanth. It was a great honor to work with him, and his brilliance never ceased to amaze me.

I am very lucky to have collaborated with Swastik. I still can't believe that my innocent question after his graph theory course turned into a two year collaboration resulting in a paper. He immediately saw the potential for proving such a result, and without his persistence we never would have cracked it open.

To Dr. Zeilberger I am thankful for his advice regarding my work on happy numbers. At the time I had no idea what to do with regards to publishing and he didn't hesitate to help me with it.

Several graduate students were very kind to have helped proofread some of my work, namely Kellen Myers and Simao Herdade. Many thanks as well to Pat Devlin for his help generating the necessary data for the happy numbers paper.

Much of this work was done while I was supported under grants NSF CCF-083727 and CCF-1219711. Without this financial support, much of this thesis would have never been completed.

Finally, I'd like to thank my parents for all of their love!

Acknowledgment of Journal Publications: Several of these chapters contain some overlap with journal versions of this work.

Chapter 2 was a joint work with Swastik Kopparty and is adapted from a journal version (see [GK]) which is, as of this writing, under referee review. Chapter 3, Section 3.3 was a joint work with Michael Saks and Srikanth Srinivasan and is part of a journal version (see [GSS]) which has yet to appear as of this writing. The work in Chapter 5 was a joint work with Michael Saks and Michal Koucky and is due to appear in [GKS]. Finally, Chapter 6 contains large overlap with the journal version [Gil13].

Table of Contents

Chapter 1

Introduction

Here we provide a bit more detail on what is covered in each chapter. Each also begins with an even more detailed introduction as well as historical background on each problem.

In Chapter 2 we study the random variable S_n which counts the number of triangles which appear in the random graph $G(n,p)$ where $p \in (0,1)$ is a fixed constant. We prove a *local* central limit theorem for S_n, that is we estimate $\mathbb{P}[S_n = k]$ to a $1 + o(1)$ multiplicative factor when k is close to the mean of S_n. This was a joint work with Swastik Kopparty.

In Chapter 3 we study the relationship between three complexity measures of Boolean functions, namely block sensitivity, certificate complexity, and degree. The main result in the chapter is a construction of an infinite sequence of functions for which the certificate complexity grows as a constant multiple of the square of the block sensitivity. This separation is tight up to a multiplicative factor. We also discuss attempts to separate degree from the other two measures. This was a joint work with Michael Saks and Srikanth Srinivasan.

Chapter 4 investigates the relationship between two randomized variants of the decision tree model of computation, namely zero-error and two-sided error randomized decision tree complexity, denoted respectively as $R(f)$ and $R_2(f)$. A natural approach for separating two such complexity measures is to start with a Boolean function f and iteratively compose the function with itself to generate an infinite sequence of functions where one complexity measure grows much faster relative to the second. The main result in this chapter shows that one cannot compose a monotone function to achieve polynomial separation between $R(f)$ and $R_2(f)$. Furthermore, if a certain conjecture

about Boolean functions holds, then composing any boolean function will not separate these measures.

Chapter 5 regards a novel approach to the well known sensitivity conjecture for Boolean functions. We study a certain cooperative two player communication game, and prove that a strong enough lower bound on the cost of this game implies the sensitivity conjecture. We establish such a lower bound for certain subset of protocols for this game. We also study two natural variations of the cost of this game. This was a joint work with Michael Saks and Michal Koucky.

In Chapter 6 we investigate the density of a sequence of integers called happy numbers. We use a probabilistic argument with an application of a local central limit theorem for sums of iid random variables to show that the asymptotic density of the set of happy numbers does not exist. In particular we show that the upper density of this set is at least .18 and the lower density is at most .12.

Chapter 2

A Local Central Limit Theorem for Triangles in a Random Graph

Acknowledgement of Journal Publication:

A large part of this chapter contains overlap with the journal version which has been submitted to *Randomized Structures and Algorithms* under the same title and (as of this writing) is under referee review. In some chapters we include more detail or extra proofs that do not appear in the journal version. It should be noted that the proof in Section 2.8 is due to Swastik Kopparty. However, the proof was adapted from an earlier idea of the author.

2.1 Introduction

We will work with the Erdos-Renyi random graph $G(n,p)$. Recall that $G(n,p)$ is the random undirected graph G on n vertices sampled by including each of the $\binom{n}{2}$ possible edges into G independently with probability p. Let S_n be the random variable equal to the number of triangles in $G(n,p)$. Let $\mu_n = \mathbb{E}[S_n] = p^3\binom{n}{3}$ and $\sigma_n = \sqrt{\mathbf{Var}[S_n]} = \Theta(n^2)$ (see the chapter Appendix for an exact calculation of σ_n). Our main result (Theorem 2.2) states that if p is a fixed constant in $(0,1)$, then the distribution of S_n is *pointwise* approximated by a discrete Gaussian distribution:

$$\Pr[S_n = k] = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-((k-\mu_n)/\sigma_n)^2/2} \pm o(1/n^2). \tag{2.1}$$

Thus, for every $k \in \mu_n \pm O(n^2)$, we determine the probability that $G(n,p)$ has exactly k triangles, up to a $(1+o(1))$ multiplicative factor.

2.1.1 Central Limit Theorems

The study of random graphs has over 50 years of history, and understanding the distribution of subgraph counts has long been a central question in the theory. When the edge probability p is a fixed constant in $(0, 1)$, there is a classical central limit theorem for the triangle count S_n (as well as for other connected subgraphs). This theorem says that for fixed constants a, b:

$$\left| \Pr\left[a \leq (S_n - \mu_n)/\sigma_n \leq b\right] - \int_a^b \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \right| = o(1),$$

(in other words, $(S_n - \mu_n)/\sigma_n$ converges in distribution to the standard Gaussian distribution). There are several proofs of the central limit theorem for subgraph counts, as well as some vast generalizations, known today.

The original proofs of the central limit theorem for triangle counts (and general subgraph counts) used the method of moments. This method is based on the fact for all distributions that are uniquely determined by their moments, the convergence of the moments of a sequence of random variables to the moments of the distribution implies convergence in distribution. Application of the moment method to subgraph statistics goes back to Erdos and Renyi's original paper [Erd60]. There were several papers in the 1980's (see [KR83] and [Kar84]) that used the moment method to understand, under increasingly general assumptions, when normalized subgraph counts converge in distribution to the Gaussian distribution. This line of work culminated with a paper by Ruciński [Ruc88] who completely characterized when normalized subgraph counts converge in distribution to the Gaussian distribution.

There are several other approaches to the central limit theorem for triangle counts (and general subgraph counts). Using Stein's method [Ste71], Barbour, Karoński and Ruciński [BKR89] obtained strong quantitative bounds on the error in the central limit theorem for subgraph counts. A technique from the asymptotic theory of statistics, known as U-statistics, was applied by Nowicki and Wierman [NW88] to obtain a central limit theorem for subgraph counts, although, in a slightly less general setting than the theorem of Ruciński. Janson [Jan92] used a similar method with several applications, including central limit theorems for the joint distribution of various graph statistics.

None of these techniques, however, seem to be quantitatively strong enough to estimate the point probability mass of the triangle/subgraph counts when the edge probability p is a constant.

2.1.2 Poisson Convergence

When the edge probability p is small enough (for example, $p \approx c/n$ for triangles), then there are classical results that give good estimates for $\Pr[S_n = k]$. In this regime, the distribution of the subgraph count S_n itself (i.e., without normalization) converges in distribution (and hence pointwise) to a Poisson random variable. Some of the work dedicated to understanding this probability regime goes back to the original paper of Erdos and Renyi [Erd60] who studied the distribution of counts of trees and cycles using the method of moments. Using Chen's [Che75] generalization of Stein's method to the Poisson setting, Barbour [Bar82] showed Poisson convergence for general subgraph counts. In the Poisson setting, the probability mass is concentrated in an interval of constant size and thus all results are "local" in the sense that they bound the point probability mass of these random variables.

For slightly larger $p \in [n^{-1}, O(n^{-(1/2)})]$ (this is the range of p where $\sigma_n = \Theta(\mu_n)$), Röllin and Ross [RR10] showed that the probability mass function for triangle counts (S_n) is close in the ℓ_∞ and total variation metrics to the probability mass function of a translated Poisson distribution (and hence a discrete Gaussian distribution), and asked whether a similar local limit law holds for larger p (See Remark 4.5 of that paper). Our result gives such a law for constant $p \in (0, 1)$ for the ℓ_∞ metric.

2.1.3 Subgraph counts mod q

Some more recent works studied the distribution of subgraph counts mod q. For example, Loebl, Matousek and Pangrac [LMP04] studied the distribution of S_n mod q in $G(n, 1/2)$. They showed that when $q \in (\omega(1), O(\log^{1/3} n))$, then for every $a \in \mathbb{Z}_q$, the probability that $S_n \equiv a \bmod q$ equals $(1 + o(1)) \cdot \frac{1}{q}$. Kolaitis and Kopparty [KK13] also studied this problem in $G(n, p)$ for fixed $p \in (0, 1)$. They showed that for every constant q, and every $a \in \mathbb{Z}_q$, the probability that $S_n \equiv a \bmod q$ equals $(1 + \exp(-n)) \cdot \frac{1}{q}$. This

latter result also generalizes to all connected subgraph counts, and to multidimensional versions for the joint distribution of all connected subgraph counts simultaneously. DeMarco, Kahn and Redlich [DKR14] extended these results of [KK13] to determine the distribution of subgraph counts mod q in $G(n,p)$ for all p. Many of these works use conditioning arguments that are similar to those used here.

2.1.4 Our result

The above lines of work:

1. the central limit theorem for triangle counts in $G(n,p)$ with p constant,
2. the Poisson local limit theorem for triangle counts in $G(n,p)$ with p close to n^{-1},
3. the uniform distribution of triangle counts mod q in $G(n,p)$ with p constant,

all strongly suggest the truth of our main theorem (Theorem 2.2): there is a local discrete Gaussian limit law for triangle counts in $G(n,p)$ with p constant.

The high level structure of our proof follows the basic Fourier analytic strategy behind the classical local limit theorem for the sums of i.i.d. integer valued random variables. To show that the distribution of $(S_n-\mu_n)/\sigma_n$ is close pointwise to the discrete Gaussian distribution (as in equation (2.1)), it suffices to show that their characteristic functions (Fourier transforms) are close in L_1 distance. Specifically, if we define $\psi_n(t) = \mathbb{E}[e^{it(S_n-\mu_n)/\sigma_n}]$, we need to show that:

$$\int_{-\pi\sigma_n}^{\pi\sigma_n} |\psi_n(t) - e^{-t^2/2}|\, dt = o(1).$$

The central limit theorem for triangle counts can be used to bound the above integral in the range $(-A, A)$ for any large constant A. By choosing A large enough, we can bound $\int_{A<|t|<\pi\sigma_n} |e^{-t^2/2}|\, dt$ by an arbitrarily small constant. We are thus reduced to showing that $\int_{A<|t|<\pi\sigma_n} |\psi_n(t)|\, dt = o(1)$. We achieve this using two different arguments. For $A < |t| < n^{0.55}$, we show that $|\psi_n(t)| < \frac{1}{t^{1+\delta}}$ using a conditioning argument, where we first reveal the edges in a set $F \subseteq \binom{[n]}{2}$, and count triangles according to how many edges they have in F. For $n^{0.55} < |t| < \pi\sigma_n$, we show that $|\psi_n(t)|$ is superpolynomially

small in t by another conditioning argument, where we partition the vertex set $[n]$ into two sets U and V, first expose all the edges within V, and then consider the increase to the total number of triangles that occurs when we expose the remaining edges.

We conjecture that a similar local discrete Gaussian limit law should hold for the number of copies of any fixed connected graph H in $G(n,p)$, for any p that lies above the threshold probability for appearance of H. It would also be interesting to understand the joint distribution of subgraph counts in $G(n,p)$ for several fixed connected graphs. It seems like there are many interesting questions here and much to be investigated.

2.2 Notation and Preliminaries

Let $[n]$ denote the set $\{1,2,\ldots,n\}$. For each positive integer n let K_n be the complete graph on the vertex set $[n]$. The Erdos-Renyi random graph $G(n,p)$ is the graph G with vertex set $[n]$, where for each $e \in \binom{[n]}{2}$, the edge e is present in G independently with probability p. For $e \in \binom{[n]}{2}$, let X_e denote the indicator for the event that edge e is present in G. For $E \subseteq \binom{[n]}{2}$, we will let $\{0,1\}^E$ denote the set of $\{0,1\}$-vectors indexed by E. Likewise $X_E \in \{0,1\}^E$ will be the random vector for which the value on coordinate e is the random variable X_e.

For the rest of this chapter $p \in (0,1)$ will be a universal fixed constant. All asymptotic notation will hide constants which may depend on p. We will use S_n to denote the number of triangles in $G(n,p)$ (thus $S_n \in [0, \binom{n}{3}]$). The mean of S_n is $p^3\binom{n}{3}$ and the variance (see Appendix) is $\sigma_n^2 = \Theta(n^4)$. We let R_n denote the normalized triangle count, $R_n \stackrel{\text{def}}{=} \frac{S_n - p^3\binom{n}{3}}{\sigma_n}$.

Fourier inversion formula for lattices: Let Y be a random variable that has support contained in the (shifted) discrete lattice $\mathcal{L} \stackrel{\text{def}}{=} \frac{1}{b}(\mathbb{Z} - a)$ for real numbers a, b. Let $\psi(t) \stackrel{\text{def}}{=} \mathbb{E}[e^{itY}]$ be the characteristic function of Y. Then for all $y \in \mathcal{L}$ it holds that

$$\Pr(Y = y) = \frac{1}{2\pi b} \int_{-\pi b}^{\pi b} e^{-ity}\psi(t)\, dt. \tag{2.2}$$

Throughout the chapter, for real numbers x we will use $\|x\|$ to denote the distance from x to the nearest integer. We will often apply the following easy bound.

Lemma 2.1. *Let B be a Bernoulli random variable that is 1 with probability p. Then:*

$$|\mathop{\mathbb{E}}_{B}[e^{i\theta B}]| \leq 1 - 8p(1-p) \cdot \|\theta/2\pi\|^2 .$$

Proof. Without loss of generality, we may assume that $\theta \in [-\pi, \pi]$. We first state two elementary inequalities:

$$\cos(t) \leq 1 - 8 \cdot \|t/2\pi\|^2 \qquad (\text{for } t \in [-\pi, \pi]) \tag{2.3}$$

and

$$\sqrt{1-t} \leq 1 - t/2 \qquad (\text{for } t \leq 1). \tag{2.4}$$

Then we have the following,

$$\begin{aligned}
|E[e^{i\theta b}]| &= |p + (1-p)e^{i\theta}| \\
&= \sqrt{p^2 + (1-p)^2 + 2p(1-p)\cos(\theta)} \\
&\leq \sqrt{p^2 + (1-p)^2 + 2p(1-p)\left(1 - 8 \cdot \|\theta/2\pi\|^2\right)} \qquad (\text{applying (2.3)}) \\
&= \sqrt{1 - 16p(1-p)\|\theta/2\pi\|^2} \\
&\leq 1 - 8p(1-p)\|\theta/2\pi\|^2 \qquad (\text{applying (2.4)}).
\end{aligned}$$

□

2.3 Main Result

We now give a formal statement of our main result.

Theorem 2.2 (Local limit law for triangles in $G(n,p)$)**.** *Let*

$$p_n(x) = \Pr(R_n = x) \quad \textit{for } x \in \mathbb{L}_n = \left\{ \frac{k - p^3\binom{n}{3}}{\sigma_n} : k \in \mathbb{Z} \right\}$$

and

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \textit{for } x \in (-\infty, \infty).$$

Then as $n \to \infty$,

$$\sup_{x \in \mathbb{L}_n} |\sigma_n p_n(x) - \mathcal{N}(x)| \to 0.$$

Equivalently, we have that for all n, for all $k \in \mathbb{Z}$,

$$\Pr[S_n = k] = \frac{1}{\sigma_n} \cdot \mathcal{N}\left(\frac{k - p^3 \cdot \binom{n}{3}}{\sigma_n}\right) + o\left(\frac{1}{n^2}\right),$$

(where the $o(1)$ term goes to 0 as $n \to \infty$, uniformly in k).

Proof. Let $\psi_n(t) = \mathbb{E}[e^{itR_n}]$. Then the Fourier inversion formula for lattices (equation 2.2) gives us

$$\sigma_n p_n(x) = \frac{1}{2\pi} \int_{-\pi\sigma_n}^{\pi\sigma_n} e^{-itx} \psi_n(t)\ dt.$$

The standard Fourier inversion formula (for $\mathbb{R}$), along with the well known formula for the Fourier transform of $\mathcal{N}$, gives us:

$$\mathcal{N}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-t^2/2}\ dt.$$

Therefore,

$$|\sigma_n p_n(x) - \mathcal{N}(x)| \leq \int_{-\pi\sigma_n}^{\pi\sigma_n} |\psi_n(t) - e^{-t^2/2}|\ dt + 2 \int_{\pi\sigma_n}^{\infty} e^{-t^2/2}\ dt \tag{2.5}$$

The second term goes to zero as n tends to infinity. Thus, it suffices to show that

$$\int_{-\pi\sigma_n}^{\pi\sigma_n} \left|\psi_n(t) - e^{-t^2/2}\right|\ dt \tag{2.6}$$

tends to 0.

Let $A > 0$ be a large constant to be determined later. We divide the integral into three regions:

- $R_1 = (-A, A)$
- $R_2 = (-n^{0.55}, -A) \cup (A, n^{0.55})$
- $R_3 = (-\pi\sigma_n, -n^{0.55}) \cup (n^{0.55}, \pi\sigma_n)$

The following three lemmas will help us bound the integral of $\left|\psi_n(t) - e^{-t^2/2}\right|$ in these three regions.

Lemma 2.3. *Let A be a fixed positive real number. Then*

$$\int_{-A}^{A} \left|\psi_n(t) - e^{-t^2/2}\right| \, dt \to 0$$

as $n \to \infty$.

Lemma 2.4. *There exists a sufficiently large constant $D = D(p)$ and $\delta > 0$ such that, for all t with $|t| \in (0, n^{0.55}]$,*

$$|\psi_n(t)| \leq D/|t|^{1+\delta}.$$

Lemma 2.5. *There exists a sufficiently large constant $D = D(p)$ such that, for all t with $|t| \in [n^{0.55}, \pi\sigma_n]$, it holds that*

$$|\psi_n(t)| \leq D/|t|^{50}.$$

We now proceed to bound (2.6).

By Lemma 2.3,

$$\int_{R_1} \left|\psi_n(t) - e^{-t^2/2}\right| \, dt \to 0,$$

for any fixed constant A.

For R_2 and R_3 we have the following,

$$\int_{R_2 \cup R_3} \left|\psi_n(t) - e^{-t^2/2}\right| \, dt \leq \int_{R_2 \cup R_3} |\psi_n(t)| \, dt + \int_{R_2 \cup R_3} \left|e^{-t^2/2}\right| \, dt$$

By Lemma 2.4 and Lemma 2.5, there exists constants $D = D(p), \delta > 0$ such that, $|\psi_n(t)| \leq \frac{D}{t^{1+\delta}}$ for all n and all t with $|t| \in (0, \pi\sigma_n]$. Therefore,

$$\int_{R_2 \cup R_3} \left|\psi_n(t) - e^{-t^2/2}\right| \, dt \leq \int_{R_2 \cup R_3} \left|\frac{D}{t^{1+\delta}}\right| \, dt + \int_{R_2 \cup R_3} \left|e^{-t^2/2}\right| \, dt.$$

Since $D/|t|^{1+\delta}$ and $e^{-t^2/2}$ both have finite integral over $(-\infty, -1) \cup (1, \infty)$, the last line above can be made smaller than any ϵ for large enough constant $A = A(\epsilon, p)$. □

2.4 Proof Sketch for Bounding $|\psi_n(t)|$

In this section we sketch with some more detail the strategy used to bound the characteristic function

$$\psi_n(t) \stackrel{\text{def}}{=} \mathbb{E}[e^{itR_n}].$$

As a warm up, suppose that R_n was the sum of n i.i.d random variables X_i. Then, by independence,

$$\psi_n(t) = \mathbb{E}\left[e^{it\sum_{i=1}^{n} X_i}\right] = \prod_{i=1}^{n} \mathbb{E}\left[e^{itX_i}\right] = \mathbb{E}\left[e^{itX_1}\right]^n.$$

Thus if $\left|\mathbb{E}[e^{itX_1}]\right|$ is bounded sufficiently far from 1, it would follow that $|\psi_n(t)|$ is small. Of course in our case R_n is the sum of dependent random variables, and one does not immediately have the expression decompose as a product. The idea that gets around this issue is to first reveal a subset F of the edges and then, conditioning on the values of the edges in F (assuming some nice event Λ occurs), show that the expectation is small. For certain choices of F the conditional expectation *does* decompose as a product, and thus the estimation becomes easier. If the good event Λ happens with high enough probability, then one has successfully bounded $\psi_n(t)$.

We now show an argument that bounds the $\psi_n(t)$ when $n^{1/2} \ll |t| \ll n$. For starters, suppose F was all the edges of $\binom{[n]}{2}$ except for a perfect matching M, and let X_F denote the indicator vector for the edges in F that appear in G. For $e = \{u, v\} \in M$ let C_e denote the number of paths of length 2 from u to v that appear in G (note any such path must consist of edges in F). Then conditioned on the value of X_F, the expectation becomes

$$\mathbb{E}\left[e^{it(C+\sum_{e\in M} C_e X_e)/\sigma_n}\right]$$

where C denotes the number of triangles that appear consisting only of edges in F. Note that C and the C_e are all constants conditioned on the value of X_F. Also, each $C_e = C_e(X_F)$ is a binomial random variable, and thus each is concentrated around np^2.

Thus for a "typical" value of X_F one has (roughly)

$$\begin{aligned}
\left|\mathbb{E}[e^{itR_n} \mid X_F]\right| &= \left|\mathbb{E}[e^{it(\sum_{e\in M} C_e X_e)/\sigma_n}]\right| \\
&\approx \left|\prod_{e\in M} \mathbb{E}[e^{itnp^2 X_e/\sigma_n]}\right| \\
&\leq \left(1 - 8p(1-p)\left\|\frac{tnp^2}{2\pi\sigma_n}\right\|^2\right)^{n/2} \qquad \text{(applying Lemma 2.1)} \\
&\approx \left(1 - 8p(1-p)\left(\frac{tnp^2}{2\pi\sigma_n}\right)^2\right)^{n/2} \qquad \text{(since } \sigma_n = \Theta(n^2) \text{ and } |t| \ll n\text{)} \\
&\approx \left(1 - \Theta(t^2/n^2)\right)^{n/2} \\
&\approx e^{-\Theta(t^2/n)}.
\end{aligned}$$

Thus if $|t| \gg n^{1/2}$ the above will be small.

In Section 2.6 we push the above analysis to cover the range where $t \leq n^{.55}$. There we instead let M be a bipartite subgraph obtained by taking a disjoint union of k perfect matchings, where k is chosen to depend on t. As above we first reveal all edges in $F \stackrel{\text{def}}{=} \binom{[n]}{2} - M$ and then condition on the value of X_F. We then count triangles according to how many edges are in M, letting C, Y, and Z denote the number of triangles with 0,1, and 2 edges in M respectively. As before C will be a constant conditioned on X_F, and $Y = \sum_{e\in M} C_e X_e$ is the sum of $nk/2$ independent random variables.

For k large enough, $\mathbb{E}[e^{itY/\sigma_n}]$ will be small conditioned on a "typical" X_F, and the analysis follows just as above because the expectation decomposes as a product. The difficulty now is Z is a degree 2 polynomial in the variables $\{X_e : e \in M\}$, and thus $\mathbb{E}[e^{it(Y+Z)/\sigma_n}]$ does not decompose as a product even after conditioning on X_F. However, by estimating the variance of Z we will find that tZ/σ_n will be tightly concentrated in an interval of size $o(1)$, whereas tY/σ_n will be roughly uniform mod 2π. Therefore, although Z has a complicated dependence on Y, $t(Y+Z)/\sigma_n$ will still be roughly uniform mod 2π and $|\psi_n(t)|$ will be small. It should be noted that we currently do not know how to prove a stronger bound than $|\psi_n(t)| \leq 1/t^{1+\delta}$ in this range (for contrast, the argument with a single perfect matching implies an exponentially small bound). This seems to be a major obstacle for obtaining a stronger quantitative local

limit law.

The above argument is not delicate enough to deal with arbitrary t with $|t| \gg n$. In Section 2.8, we use a different conditioning argument to bound $\psi_n(t)$ for all t with $|t| \in [n^{.55}, \pi\sigma_n]$. This argument is based on partitioning $\binom{[n]}{2}$ into two sets E and F, and then studying the difference between the number of triangles in the two random graphs (X_E, X_F) and (X'_E, X_F) (where X'_E is an independent copy of the random variable X_E).

2.5 Small $|t|$

In this section we prove Lemma 2.3.

Lemma 2.3 (restated). *Let A be a fixed positive real number. Then*

$$\int_{-A}^{A} \left| \psi_n(t) - e^{-t^2/2} \right| \, dt \to 0$$

as $n \to \infty$.

The proof essentially follows from the central limit theorem for triangle counts. We provide some of the details by applying a few standard results from probability theory regarding the method of moments. We begin with some additional preliminaries that we borrow (with minor changes) from Durrett's textbook "Probability: Theory and Examples" [Dur10].

For a random variable X, its *distribution function* is the function $F(x) \stackrel{\text{def}}{=} \Pr[X \leq x]$. A sequence of distribution functions is said to *converge weakly* to a limit F if $F_n(x) \to F(x)$ for all x that are continuity points of F. A sequence of random variables X_n is said to *converge in distribution* to a limit X_∞ (written $X_n \xrightarrow{d} X_\infty$) if their distribution functions converge weakly.

The moment method gives a useful sufficient condition for when a sequence of random variables converge in distribution.

Theorem 2.6. *Let X_n be a sequence of random variables. Suppose that $\mathbb{E}[X^k]$ has a limit μ_k for each k and*

$$\limsup_{k \to \infty} \mu_{2k}^{1/2k} / 2k < \infty;$$

then $X_n \xrightarrow{d} X_\infty$ *where* X_∞ *is the unique distribution with the moments* μ_k*.*

In the Appendix we provide the standard calculation that $\mathbb{E}[R_n^k] \to \mu_k$ for all k, where

$$\mu_k \stackrel{\text{def}}{=} \begin{cases} (k-1)!! & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases}$$

are the moments of $N(0,1)$. It is easy to check that these moments do not grow too quickly and thus the theorem implies the well known central limit theorem for triangle counts:

$$R_n \xrightarrow{d} N(0,1). \tag{2.7}$$

Durrett also provides a theorem that relates convergence in distribution to pointwise convergence of characteristic functions.

Theorem 2.7. Continuity Theorem. *Let* $X_n, 1 \le n \le \infty$*, be random variables with characteristic functions* ϕ_n*. If* $X_n \xrightarrow{d} X_\infty$*, then* $\phi_n(t) \to \phi_\infty(t)$ *for all* t*.*

Applying this with (2.7), we conclude that $\psi_n(t) \to e^{-t^2/2}$ for all t. To finish the proof of Lemma 2.3 we apply the dominated convergence theorem to conclude, for any fixed A, that

$$\int_{-A}^{A} \left| \psi_n(t) - e^{-t^2/2} \right| \, dt \to 0$$

as $n \to \infty$.

2.6 Intermediate $|t|$

In this section, we prove Lemma 2.4.

Lemma 2.4 (restated). *There exists a sufficiently large constant* $D = D(p)$ *and* $\delta > 0$ *such that, for all* t *with* $|t| \in (0, n^{0.55}]$*,*

$$|\psi_n(t)| = |\,\mathbb{E}[e^{itR_n}]| = |\,\mathbb{E}[e^{itS_n/\sigma_n}]| \le D/|t|^{1+\delta}.$$

Note that trivially $|\psi_n(t)| \leq 1$, and thus the lemma already holds for constant sized t. Thus we will assume that t and n are both bigger than a sufficiently large constant $D(p)$. To make the exposition simpler, we will assume n is even (however, the same argument can be easily seen to apply when n is odd).

We simplify notation by denoting R_n by R, S_n by S, and σ_n by σ. Partition $[n]$ into sets U, V both of size $n/2$ and let $P \subseteq \binom{[n]}{2}$ be the complete bipartite graph between vertex sets U, V. Let $k < \frac{n}{10^{10}}$ be a positive integer to be determined later. Let $M_1, \cdots, M_k \subseteq P$ be pairwise disjoint perfect matchings between U and V. Let $E = M_1 \cup M_2 \cup \cdots \cup M_k$, and let $F = \binom{[n]}{2} \setminus E$.

Recall that for the random graph $G \in G(n,p)$, we use X_e to denote the indicator for whether edge e appears in G. We also use X_E and X_F to denote the $\{0,1\}^E$-valued random variable $(X_e)_{e \in E}$ and the $\{0,1\}^F$-valued random variable $(X_e)_{e \in F}$ respectively. Let $C(X_F), Y(X_E, X_F)$ and $Z(X_E, X_F)$ be random variables that count the number of triangles in $G(n,p)$ which have 0,1, and 2 edges in E respectively (note that, by construction of E, no triangle may have all 3 edges in E). Thus we have $S = C(X_F) + Y(X_E, X_F) + Z(X_E, X_F)$.

We define:

$$\zeta = \mathop{\mathbb{E}}_{X_E, X_F}[Z(X_E, X_F)].$$

We now work towards bounding $\left|\mathbb{E}[e^{itS/\sigma}|]\right|$:

$$\left|\mathbb{E}[e^{itS/\sigma}]\right| = \left|\mathop{\mathbb{E}}_{X_E, X_F}[e^{it(C(X_F)+Y(X_E,X_F)+Z(X_E,X_F))/\sigma}]\right|.$$

By adding and subtracting the term $e^{it(C(X_F)+Y(X_E,X_F)+\zeta)}$ and applying the triangle inequality, the above is

$$\leq \left|\mathop{\mathbb{E}}_{X_E, X_F}[e^{it(C(X_F)+Y(X_E,X_F)+\zeta)/\sigma}]\right| + \mathop{\mathbb{E}}_{X_E, X_F}\left[\left|e^{it(Z(X_E,X_F))/\sigma} - e^{it\zeta/\sigma}\right|\right].$$

We bound each of the two terms separately in the following two lemmas. We will then use these lemmas to conclude the proof of Lemma 2.4.

Lemma 2.8.

$$\left|\mathop{\mathbb{E}}_{X_E, X_F}[e^{it(C(X_F)+Y(X_E,X_F)+\zeta)/\sigma}]\right| \leq e^{-\Theta(t^2 k/n)}.$$

Proof. We bound the above expectation by revealing the edges in two stages. We first reveal X_F, and show that with high probability over the choice of X_F, some good event occurs. We then show that whenever this good event occurs, the value of the above expectation over the random choice of X_E is small.

Formally, using the triangle inequality we get:

$$\left| \mathop{\mathbb{E}}_{X_E,X_F}[e^{it(C(X_F)+Y(X_E,X_F)+\zeta)/\sigma}] \right| = \left| \mathop{\mathbb{E}}_{X_F}\left[e^{it(C(X_F)+\zeta)/\sigma} \cdot \mathop{\mathbb{E}}_{X_E}[e^{it(Y(X_E,X_F))/\sigma}] \right] \right| \tag{2.8}$$

$$\leq \mathop{\mathbb{E}}_{X_F}\left[\left| \mathop{\mathbb{E}}_{X_E}[e^{it(Y(X_E,X_F))/\sigma}] \right| \right]. \tag{2.9}$$

For $e = \{u, v\} \in E$ and a vector $x_F \in \{0,1\}^F$, we let $Y_e(x_F)$ denote the number of paths of length 2 from u to v consisting entirely of edges $f \in F$ for which $(x_F)_f = 1$ [1]. In this way, for a given x_F, the random variable $Y(X_E, x_F)$ equals $\sum_{e \in E} Y_e(x_F) X_e$.

Define

$$L = \{x_F \in \{0,1\}^F \mid \text{for some } e \in E,\ Y_e(x_F) < np^2/2\}.$$

Let Λ denote the (bad) event that $X_F \in L$.

Claim 2.9.

$$\Pr_{X_F}[\Lambda] \leq e^{-\Theta(n)}.$$

Proof. Observe that for any given $e \in E$, the distribution of $Y_e(X_F)$ equals $\mathrm{Bin}(m_e, p^2)$, where m_e equals the number of paths of length 2 joining the endpoints of e, and consisting entirely of edges in F. Also note that we have $m_e \geq n - 2k \geq n(1 - 1/10^9)$.

By the Chernoff bound, we have:

$$\Pr[\mathrm{Bin}(m_e, p^2) < np^2/2] \leq e^{-np^2(1-p^2)/200}.$$

Taking a union bound over all $e \in E$, we get the claim. □

Next, we show that if we condition on Λ not occurring, then the desired expectation is small.

---

[1] This differs from the exposition in Section 2.4 (where E is a single perfect matching), in that some length-2 paths between u and v here may contain edges in E. We do not want to count those paths in $Y_e(x_F)$.

Claim 2.10. *For every $x_F \in \{0,1\}^F \setminus L$,*

$$\left| \mathop{\mathbb{E}}_{X_E \in \{0,1\}^E} \left[e^{itY(X_E, x_F)/\sigma} \right] \right| \leq e^{-\Theta(t^2 k/n)}.$$

Proof. Recall that $Y(X_E, x_F) = \sum_{e \in E} Y_e(x_F) X_e$. Thus we have:

$$\begin{aligned}
\left| \mathop{\mathbb{E}}_{X_E} \left[e^{itY(X_E, x_F)/\sigma} \right] \right| &= \left| \mathop{\mathbb{E}}_{X_E} \left[e^{it(\sum_{e \in E} Y_e(x_F) X_e)/\sigma} \right] \right| \\
&= \left| \prod_{e \in E} \mathbb{E} \left[e^{itY_e(x_F) X_e/\sigma} \right] \right| \quad \text{by the mutual independence of } (X_e)_{e \in E} \\
&\leq \prod_{e \in E} \left(1 - 8p(1-p) \left\| \frac{tY_e(x_F)}{2\pi\sigma} \right\|^2 \right) \quad \text{(applying Lemma 2.1)} \\
&= \prod_{e \in E} \left(1 - 8p(1-p) \cdot \left(\frac{tY_e(x_F)}{2\pi\sigma} \right)^2 \right) \\
&\leq \left(1 - 8p(1-p) \cdot \left(\frac{tnp^2}{4\pi\sigma} \right)^2 \right)^{nk/2} \quad \text{(since } x_F \in L\text{)}.
\end{aligned}$$

Recall that $\sigma = \sqrt{\frac{n(n-1)(n-2)(n-3)D}{2}}$ for some constant $D \leq 1$. Thus $\frac{tnp^2}{4\pi\sigma} \geq \frac{tp^2}{4\pi n}$. Therefore we may further bound the above expression by:

$$\begin{aligned}
&\leq \left(1 - 8p(1-p) \left(\frac{tp^2}{4\pi n} \right)^2 \right)^{nk/2} \\
&\leq e^{-\frac{t^2 p^5 (1-p) k}{\pi^2 n}} \\
&= e^{-\Theta(t^2 k/n)}
\end{aligned}$$

□

Going back to equation (6.5) we have

$$\begin{aligned}
\left| \mathop{\mathbb{E}}_{X_E, X_F} [e^{it(C(X_F) + Y(X_E, X_F) + \zeta)/\sigma}] \right| &\leq \mathop{\mathbb{E}}_{X_F} \left[\left| \mathop{\mathbb{E}}_{X_E} [e^{it(Y(X_E, X_F))/\sigma}] \right| \right] \\
&\leq \Pr_{X_F}[X_F \in L] + \max_{x_F \in \{0,1\}^F \setminus L} \left| \mathop{\mathbb{E}}_{X_E} [e^{it(Y(X_E, x_F))/\sigma}] \right| \\
&\leq e^{-\Theta(n)} + e^{-\Theta(t^2 k/n)} \quad \text{(applying claims 2.9 and 2.10)} \\
&\leq e^{-\Theta(t^2 k/n)}.
\end{aligned}$$

□

Lemma 2.11.

$$\mathop{\mathbb{E}}_{X_E,X_F}\left[\left|e^{it(Z(X_E,X_F))/\sigma}-e^{it\zeta/\sigma}\right|\right]\leq O\left(t^{3/2+\delta/2}\left(\frac{k}{n}\right)^{3/2}\right)+O\left(1/t^{1+\delta}\right)$$

Proof. Simplifying the expression we want to bound, we get:

$$\mathop{\mathbb{E}}_{X_E,X_F}\left[\left|e^{it(Z(X_E,X_F))/\sigma}-e^{it\zeta/\sigma}\right|\right]=\mathop{\mathbb{E}}_{X_E,X_F}\left[\left|e^{it(Z(X_E,X_F)-\zeta)/\sigma}-1\right|\right].$$

Thus proving the lemma reduces to proving a concentration bound: namely that $Z(X_E,X_F)$ is close to ζ with high probability. We will bound $\mathbf{Var}_{X_E,X_F}[Z(X_E,X_F)]$ and apply the Chebyshev inequality. This will give the desired concentration.

Let Δ' denote the set of triangles in K_n that have exactly 2 edges in E. For each $r\in\Delta'$, let $T_r(X_E,X_F)$ be the indicator for the triangle r appearing in G. For two triangles $r,s\in\Delta'$, write $r\sim s$ if r and s share an edge. Note for any $r\in\Delta'$ there are at most $6k$ triangles $s\in\Delta'$ for which $r\sim s$.

We have:

$$\begin{aligned}\mathbf{Var}_{X_E,X_F}[Z(X_E,X_F)]&=\sum_{r\in\Delta'}\sum_{s\in\Delta'}\mathbf{Cov}_{X_E,X_F}[T_r(X_E,X_F),T_s(X_E,X_F)]\\&=\sum_{r\in\Delta'}\sum_{s\sim r}\mathbf{Cov}_{X_E,X_F}[T_r(X_E,X_F),T_s(X_E,X_F)]\\&\leq|\Delta'|\cdot|6k|\\&\leq 6nk^3\qquad(\text{since }|\Delta'|=n\tbinom{k}{2})\end{aligned}$$

Applying Chebyshev's inequality with $\lambda=\sqrt{6}\cdot n^{1/2}\cdot t^{1/2+\delta/2}\cdot k^{3/2}$ we have

$$\begin{aligned}\Pr_{X_E,X_F}[|Z(X_E,X_F)-\zeta|>\lambda]&<\frac{\mathbf{Var}_{X_E,X_F}[Z(X_E,X_F)]}{\lambda^2}\\&<1/t^{1+\delta}\end{aligned}$$

Recall that $\|x\|$ denotes the distance from real number x to the nearest integer. Let Λ be the (bad) event that $|Z(X_E,X_F)-\zeta|\geq\lambda$. Using the fact that for any real number

θ, $|e^{i\theta} - 1| \leq 2\pi \cdot \|\frac{\theta}{2\pi}\|$, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{X_E,X_F}\left[\left|e^{it(Z(X_E,X_F)-\zeta)/\sigma} - 1\right|\right] &\leq 2\pi \mathop{\mathbb{E}}_{X_E,X_F}\left[\left\|\frac{t(Z(X_E,X_F)-\zeta)}{2\pi\sigma}\right\|\right] \\
&\leq 2\pi \cdot \Pr[\Lambda^c] \cdot \frac{t\lambda}{2\pi\sigma} + 2\pi \cdot \Pr[\Lambda] \cdot \frac{1}{2} \\
&\leq \frac{t\lambda}{\sigma} + \pi \cdot \Pr[\Lambda] \\
&\leq \sqrt{6} \cdot t^{3/2+\delta/2} \cdot \frac{k^{3/2} \cdot n^{1/2}}{\sigma} + \frac{\pi}{t^{1+\delta}} \\
&\leq O\left(t^{3/2+\delta/2} \cdot \left(\frac{k}{n}\right)^{3/2}\right) + O\left(\frac{1}{t^{1+\delta}}\right). \quad \text{(since } \sigma = \Theta(n^2)\text{)}
\end{aligned}$$

This concludes the proof of Lemma 2.11. □

To conclude the proof of Lemma 2.4, we apply Lemma 2.8 and Lemma 2.11 to get the bound

$$|\mathbb{E}[e^{itS/\sigma}]| \leq e^{-\Theta(t^2k/1000n)} + O\left(t^{3/2+\delta/2} \cdot \left(\frac{k}{n}\right)^{3/2}\right) + O\left(1/t^{1+\delta}\right) \tag{2.10}$$

It only remains to check that k may be chosen as to make the right hand side of equation (6.13) bounded by $O(1/t^{1+\delta})$. Set $\delta = 0.01$, and observe that for $\Omega(1) < t < n^{0.55}$, we have the following two relations:

$$\frac{n\log^2(t)}{t^2} = O\left(\frac{n}{t^{5/3+\delta}}\right),$$

$$\frac{n}{t^{5/3+\delta}} = \omega(1).$$

Thus we may choose k to be an integer satisfying:

$$k = \Omega(n\log^2(t)/t^2) \quad \text{and} \quad k = O(n/t^{5/3+\delta}).$$

For such a k we have

$$e^{-\Theta(t^2k/n)} \leq O(1/t^{1+\delta})$$

and

$$t^{3/2+\delta/2}\left(\frac{k}{n}\right)^{3/2} = O(1/t^{1+\delta}).$$

This concludes the proof of Lemma 2.4.

2.7 A Motivating Argument

Recall that the proof sketch where we first reveal all but a single perfect matching does not work for all $t \gg n^{1/2}$. This was a technical difficulty which forced us to consider a new approach. Here we present a simpler argument which motivated the proof for large t contained in Section 2.8. We will sketch a proof that the random variable S_n is uniform mod k where $k = o(n)$. To make the argument simpler we will assume $p = 1/2$. Showing S_n is uniform mod k relates to bounding the characteristic function when $t/\sigma_n = 2\pi/k$ because for such t, $e^{itS_n\sigma_n}$ is supported over the k'th roots of unity.

Viewing $[n]$ as the set of vertices, let E be the set of edges which have both vertices in the set $\{1, 2, \cdots, n-1\}$. We first reveal all edges in E, and then study how many additional triangles appear when we reveal E^c. Let T be the number of triangles which appear containing an edge in E^c. Equivalently, T is the number of triangles which contain vertex n. We will show that with high probability over the randomness in X_E, the distribution of T conditioned on X_E is close to uniform mod k.

Note that T is equal to the number of edges which appear which have both vertices in the neighborhood of vertex n. Let

$$\vec{V} = \vec{V}(X_E) \stackrel{\text{def}}{=} (V_\emptyset, V_{\{1\}}, \cdots, V_S, \cdots, V_{[n-1]})$$

be the vector whose entries are indexed by subsets of $[n-1]$, where entry V_S is the number of edges which appear containing both vertices in the set S. Then the random variable $T = T(X_E, X_{E^c})$ can be viewed as the following: First reveal X_E, this determines the vector $\vec{V}$. Then reveal X_{E^c}, this chooses a random set S which is the neighborhood of vertex n. Then T is exactly the entry V_S in the vector $\vec{V}$. Note that the random set S chosen is a uniform subset of $[n-1]$ (since $p = 1/2$).

Thus to argue that T is uniform mod k it suffices to show, for each $j < k$, that w.h.p after we reveal X_E the fraction of entries V_S which are $\equiv j \bmod k$ is $1/k + o(1/k)$). We will use a second moment calculation to show this. Note that each entry $V_S = V_S(X_E) = \text{Bin}(\binom{|S|}{2}, 1/2)$. It is easy to show that when $|S| = \Omega(n)$ then V_S is close to uniform mod $k = o(n)$ (one can apply the local limit theorem for the Binomial random variable, or estimate these probabilities directly). Of course the entries V_S

are not mutually independent, however, most pairs $V_S, V_{S'}$ are *approximately* pairwise independent in the sense that

$$\mathbb{E}[V_S V_{S'}] \approx \mathbb{E}[V_S]\,\mathbb{E}[V_{S'}].$$

Fix $j \in \{0, 1, \cdots, k-1\}$ and for each $S \subseteq [n-1]$ let Z_S denote the indicator random variable that $V_S \equiv j \bmod k$. Let

$$Z = 1/2^{n-1} \sum_{S \subseteq [n-1]} Z_S.$$

Then $\mathbb{E}[Z] \approx 1/k$. This follows because each V_S is a binomial random variable and most sets $S \subset [n-1]$ have size $\Theta(n)$. Thus is remains to bound the variance of Z and apply Chebyshev's inequality. We say an ordered pair (S, S') is *good* if $|S - S'| \geq n/8$, otherwise the pair is *bad*. Note the fraction of pairs which are bad is exponentially small. Then,

$$\begin{aligned}\mathbb{E}[(Z - \mathbb{E}[Z])^2] &= 1/4^{n-1} \sum_{S \subseteq [n-1]} \sum_{S' \subseteq [n-1]} \mathbb{E}[(Z_S - \mathbb{E}[Z_S])(Z_{S'} - \mathbb{E}[Z_{S'}])] \\ &\leq \text{covariance of a good pair} + \text{fraction of bad pairs.}\end{aligned}$$

To finish this sketch, we will argue that the covariance $\mathbb{E}[(Z_S - \mathbb{E}[Z_S])(Z_{S'} - \mathbb{E}[Z_{S'}])]$ is exponentially small for a fixed good pair S, S'. First reveal all edges with both vertices in S', this determines the value of $Z_{S'}$. Conditioned on this first reveal, note that V_S is still $\text{Bin}(m, 1/2)$ where $m \geq \binom{n/8}{2}$ (because the pair S, S' is good). Thus the conditional distribution of V_S will still be close to uniform mod k (in fact, exponentially close). This means that $\mathbb{P}[Z_S = 1] \approx 1/k$. This finishes the proof sketch, the variance of Z using this argument can be seen to be exponentially small.

The above idea does not work to bound the characteristic function for all $t \gg n^{1/2}$ because it only applies when t/σ_n is a rational multiple of 2π and when $t/\sigma_n = \omega(1/n)$ (that is when $t = \omega(n)$). The argument presented in the next Section uses a similar conditioning on the edges, but requires a more delicate analysis of the error.

2.8 Big $|t|$

In this section we prove Lemma 2.5.

Lemma 2.5 (restated). *There exists a sufficiently large constant $D = D(p)$ such that, for all t with $|t| \in [n^{0.55}, \pi\sigma_n]$, it holds that*

$$|\mathbb{E}[e^{itR_n}]| = |\mathbb{E}[e^{itS_n/\sigma_n}]| \leq D/|t|^{50}.$$

The choice of 50 here is arbitrary, in fact the lemma will hold for any fixed constant in place of 50 (as long as $D(p)$ is chosen large enough). We only choose a large number here to remind the reader that the obstacle to a better quantitative local limit law lies in bounding $\psi_n(t)$ for $|t|$ in the range $(0, n^{.55}]$.

As in the previous section, since n is fixed we simplify notation by denoting S_n as S and σ_n as σ.

We will break down the proof into two different cases. Both cases will use a common framework, which we now set up.

Let $[n] = U \cup V$ be a partition of the vertices. Define $X_U = (X_e)_{e \in \binom{U}{2}}$. For every $x_U \in \{0,1\}^{\binom{U}{2}}$, we will show that:

$$\mathbb{E}[e^{itS/\sigma} | X_U = x_U] \leq O\left(\frac{1}{t^{50}}\right).$$

This will imply the desired bound.

From now on, we condition on $X_U = x_U$.

Let $E_U \subseteq \binom{U}{2}$ be the induced graph on U:

$$E_U = \left\{ \{u, u^*\} \in \binom{U}{2} \mid x_{\{u,u^*\}} = 1 \right\}.$$

Note that E_U is determined by x_U and is thus fixed.

For $u \in U$, let $A_u \in \{0,1\}^V$ denote the vector indicating the neighbors of u in V. Thus $A_u = (X_{\{u,v\}})_{v \in V}$.

Let $B \in \{0,1\}^{\binom{V}{2}}$ denote the adjacency vector of $G|_V$. Thus $B = \{X_e\}_{e \in \binom{V}{2}}$.

Note that all the entries of the A_u's and B are independent p-biased Bernoulli random variables. We will now express the number of triangles in G in terms of the A_u's and B (here $\langle \cdot, \cdot \rangle$ denotes the standard inner product over $\mathbb{R}$) :

- Let S_U denote the number of triangles in G with all three vertices in U (note that S_U is determined by x_U and is thus fixed).
- The expression $\sum_{\{u,u^*\} \in E_U} \langle A_u, A_{u^*} \rangle$ counts the number of triangles in G that have exactly two vertices in U.
- Let $P : \{0,1\}^V \to \{0,1\}^{\binom{V}{2}}$ denote the map defined by:

$$P(r)_{\{u,v\}} = r_u \cdot r_v.$$

  Then $\sum_{u \in U} \langle P(A_u), B \rangle$ counts the number of triangles in G that have exactly two vertices in V.
- Let $Q : \{0,1\}^{\binom{V}{2}} \to \mathbb{N}$ denote the map that sends an adjacency vector b to the number of triangles in the graph represented by b (that is the triangles whose vertices are contained in V).

  Thus $Q(B)$ counts the number of triangles in G with all three vertices in V.

Then we have the following expression for S in terms of the A_u's and B.

$$S = S_U + \sum_{u \in U} \langle P(A_u), B \rangle + \sum_{\{u,u^*\} \in E_U} \langle A_u, A_{u^*} \rangle + Q(B).$$

We now bound $\mathbb{E}[e^{itS/\sigma}]$.

$$\begin{aligned}
|\,\mathbb{E}[e^{itS/\sigma}]|^2 &= \left| \mathop{\mathbb{E}}_{(A_u)_{u \in U}, B} \left[e^{it(S_U + \sum_{u \in U} \langle P(A_u), B \rangle + \sum_{\{u,u^*\} \in E_U} \langle A_u, A_{u^*} \rangle + Q(B))/\sigma} \right] \right|^2 \\
&\leq \mathop{\mathbb{E}}_{B} \left[\left| e^{itQ(B)/\sigma} \cdot \mathop{\mathbb{E}}_{(A_u)_{u \in U}} \left[e^{it(\langle \sum_{u \in U} P(A_u), B \rangle + \sum_{\{u,u^*\} \in E_U} \langle A_u, A_{u^*} \rangle)/\sigma} \right] \right|^2 \right] \\
&\leq \mathop{\mathbb{E}}_{B} \left[\left| \mathop{\mathbb{E}}_{(A_u)_{u \in U}} \left[e^{it(\langle \sum_{u} P(A_u), B \rangle + \sum_{\{u,u^*\} \in E_U} \langle A_u, A_{u^*} \rangle)/\sigma} \right] \right|^2 \right]
\end{aligned}$$

$$= \mathop{\mathbb{E}}_{B} \mathop{\mathbb{E}}_{(A_u)_{u \in U}} \mathop{\mathbb{E}}_{(A'_u)_{u \in U}} \left[e^{it(\langle \sum_{u} P(A_u) - P(A'_u), B \rangle + \sum_{\{u,u^*\} \in E_U} \langle A_u, A_{u^*} \rangle - \sum_{\{u,u^*\} \in E_U} \langle A'_u, A'_{u^*} \rangle)/\sigma} \right].$$

(Where for each $u \in U$, A'_u is an independent copy of A_u).

$$= \mathop{\mathbb{E}}_{(A_u)_{u\in U}} \mathop{\mathbb{E}}_{(A'_u)_{u\in U}} \left[e^{it(\sum_{\{u,u^*\}\in E_U} \langle A_u, A_{u^*}\rangle - \sum_{\{u,u^*\}\in E_U} \langle A'_u, A'_{u^*}\rangle)/\sigma} \cdot \mathop{\mathbb{E}}_{B} \left[e^{it\langle \sum_u P(A_u) - P(A'_u), B\rangle/\sigma} \right] \right]$$

$$= \mathop{\mathbb{E}}_{(A_u)_{u\in U}} \mathop{\mathbb{E}}_{(A'_u)_{u\in U}} \left[e^{it(\sum_{\{u,u^*\}\in E_U} \langle A_u, A_{u^*}\rangle - \sum_{\{u,u^*\}\in E_U} \langle A'_u, A'_{u^*}\rangle)/\sigma} \cdot \mathop{\mathbb{E}}_{B} \left[e^{it\langle h_{\mathbf{A},\mathbf{A}'}, B\rangle/\sigma} \right] \right]$$

where $\mathbf{A} = (A_u)_{u\in U}$, $\mathbf{A}' = (A'_u)_{u\in U}$, and where $h_{\mathbf{A},\mathbf{A}'} \in \mathbb{Z}^{\binom{V}{2}}$ is given by:

$$h_{\mathbf{A},\mathbf{A}'} = \sum_{u\in U} (P(A_u) - P(A'_u)).$$

Observe that for each $e \in \binom{V}{2}$, $(h_{\mathbf{A},\mathbf{A}'})_e$ is distributed as the difference of two binomials of the form $B(|U|, p^2)$ (but the different coordinates of $h_{\mathbf{A},\mathbf{A}'}$ are not independent).

Our goal is to show that with high probability over the choice of $\mathbf{A}, \mathbf{A}'$, we have that:

$$C \stackrel{\text{def}}{=} \left| \mathop{\mathbb{E}}_{B} \left[e^{it\langle h_{\mathbf{A},\mathbf{A}'}, B\rangle/\sigma} \right] \right|$$

is small in absolute value. This will imply that $\mathbb{E}[e^{itS/\sigma}]$ is small, as desired.

We now achieve this goal for $|t| > n^{0.55}$ using two different arguments (to cover two different ranges of $|t|$), instantiating the above framework with different settings of $|U|$.

2.8.1 Case 1: $n^{1.001} \le |t| < \pi\sigma$

Suppose $n^{1.001} < |t| < \pi\sigma$. For this argument, we choose $|U| = 1$.

In this case, the coordinates of $h_{\mathbf{A},\mathbf{A}'}$ have the following joint distribution: Let $J \subseteq V$ be a random subset where each $v \in V$ appears independently with probability p. Let J' be an independent copy of J (think of J and J' as two independently chosen neighborhoods of the vertex u). Then the e coordinate of $h_{\mathbf{A},\mathbf{A}'}$ is 1 if $e \subseteq J - J'$, 0 if $e \subseteq J \cap J'$ or $e \subseteq J^c \cap (J')^c$, and -1 if $e \subseteq J' - J$. A Chernoff bound implies that with probability at least $1 - e^{-\Theta(n)}$ the symmetric difference of J and J' will have size at least $np(1-p)/2$. In such a case $h_{\mathbf{A},\mathbf{A}'}$ will have $\binom{np(1-p)/2}{2} = \Theta(n^2)$ non-zero coordinates. From now on we assume that $\mathbf{A}, \mathbf{A}'$ are such that this event occurs (and we call such an $\mathbf{A}, \mathbf{A}'$ "good").

Then we have:

$$
\begin{aligned}
C &= \left| \mathop{\mathbb{E}}_{B} \left[e^{it\langle \sum_u h_{\mathbf{A},\mathbf{A}'}, B\rangle/\sigma} \right] \right| \\
&= \left| \mathop{\mathbb{E}}_{B} \left[\prod_{e \in \binom{V}{2}} e^{it(h_{\mathbf{A},\mathbf{A}'})_e B_e/\sigma} \right] \right| \\
&= \left| \mathop{\mathbb{E}}_{B} \left[\prod_{e \in \binom{V}{2}, (h_{\mathbf{A},\mathbf{A}'})_e \neq 0} e^{it(h_{\mathbf{A},\mathbf{A}'})_e B_e/\sigma} \right] \right| \\
&= \left| \prod_{e \in \binom{V}{2}, (h_{\mathbf{A},\mathbf{A}'})_e \neq 0} \mathop{\mathbb{E}}_{B_e} \left[e^{it(h_{\mathbf{A},\mathbf{A}'})_e B_e/\sigma} \right] \right| \\
&\leq \prod_{e \in \binom{V}{2}, (h_{\mathbf{A},\mathbf{A}'})_e \neq 0} \left(1 - 8p(1-p) \cdot \left\| \frac{t \cdot |(h_{\mathbf{A},\mathbf{A}'})_e|}{2\pi\sigma} \right\|^2 \right) \quad \text{(by Lemma 2.1)} \\
&\leq \prod_{e \in \binom{V}{2}, (h_{\mathbf{A},\mathbf{A}'})_e \neq 0} \left(1 - 8p(1-p) \cdot \left(\frac{t}{2\pi\sigma} \right)^2 \right) \quad \text{(since } |(h_{\mathbf{A},\mathbf{A}'})_e| \in \{0, \pm 1\} \text{ and } |t| < \pi\sigma) \\
&\leq e^{-\frac{2p(1-p)t^2}{\pi^2\sigma^2} \cdot \Theta(n^2)} \quad \text{since } \mathbf{A}, \mathbf{A}' \text{ is good} \\
&\leq e^{-\Theta(t^2/n^2)} \quad (\text{since } \sigma = \Theta(n^2)).
\end{aligned}
$$

Now we use the fact that $t \geq n^{1.001}$ to conclude that $D \leq \exp(-\Theta(n^{0.002}))$.

Taking into account the probability of $\mathbf{A}, \mathbf{A}'$ being good, we get:

$$|\,\mathbb{E}[e^{itS/\sigma}]|^2 < e^{-\Theta(n)} + e^{-\Theta(n^{0.002})} \ll \frac{1}{t^{100}},$$

as desired.

2.8.2 Case 2: $n^{0.55} \leq t < n^{1.01}$

Suppose $n^{0.55} < t < n^{1.01}$. For this argument, we choose $|U| = n/2$.

As before, we have:

$$C = \left| \mathop{\mathbb{E}}_{B} \left[\prod_{e \in \binom{V}{2}} e^{it(h_{\mathbf{A},\mathbf{A}'})_e B_e/\sigma} \right] \right|$$

Now for each $e \in \binom{V}{2}$, the distribution of $(h_{\mathbf{A},\mathbf{A}'})_e$ is the difference of two binomials of the form $Bin(|U|, p^2)$. Thus, we will typically have $(h_{\mathbf{A},\mathbf{A}'})_e$ around $\sqrt{|U|}$ in magnitude.

For each $e \in \binom{V}{2}$, let Λ_e be the following bad event (depending on $\mathbf{A}, \mathbf{A}'$): $|(h_{\mathbf{A},\mathbf{A}'})_e| \notin (|U|^{0.49}, |U|^{0.51})$. Let $\gamma = \Pr[\Lambda_e]$. By standard concentration and anti-concentration estimates for Binomial distributions, we have that $\gamma \leq 0.1$ (provided n is sufficiently large, depending on p).

Let Λ be the bad event that for more than $|V|^2/4$ choices of $e \in \binom{V}{2}$, the event Λ_e occurs.

Lemma 2.12. *There is a constant A such that for every k:*

$$\Pr[\Lambda] < \frac{k^{Ak}}{|V|^k}.$$

Proof. Let Z_e be the indicator variable for the event Λ_e. For each e, we have $\mathbb{E}[Z_e] = \gamma \leq 0.1$.

Note that if $e_1, \ldots, e_k$ are pairwise disjoint, then $Z_{e_1}, \ldots, Z_{e_k}$ are mutually independent.

Let $Z = \sum_{e \in \binom{V}{2}} (Z_e - \gamma)$. Note that $\mathbb{E}[Z] = 0$. We will show that $\mathbb{E}[Z^{2k}] \leq k^{O(k)} \cdot |V|^{3k}$. This implies that

$$\Pr[\Lambda] \leq \Pr[Z > |V|^2/8] \leq \Pr[Z^{2k} > (|V|^2/8)^{2k}] \leq \frac{\mathbb{E}[Z^{2k}]}{(|V|^2/8)^{2k}} \leq k^{O(k)} \frac{1}{|V|^k},$$

as desired.

It remains to show the claimed bound on $\mathbb{E}[Z^{2k}]$. We have:

$$\mathbb{E}[Z^{2k}] = \sum_{e_1, \ldots, e_{2k} \in \binom{V}{2}} \mathbb{E}[\prod_{j=1}^{2k} (Z_{e_j} - \gamma)].$$

We call a tuple $(e_1, \ldots, e_{2k}) \in \binom{V}{2}^{2k}$ *intersecting* if for every $i \in [2k]$, there exists $j \neq i$ with $e_j \cap e_i \neq \emptyset$. The key observation is the following: if $(e_1, \ldots, e_{2k})$ is not intersecting, then $\mathbb{E}[\prod_{j=1}^{2k} (Z_{e_j} - \gamma)] = 0$. To see this, suppose $(e_1, \ldots, e_{2k})$ is not intersecting because e_i does not intersect any other e_j. Then we have:

$$\mathbb{E}[\prod_{j=1}^{2k} (Z_{e_j} - \gamma)] = \mathbb{E}[Z_{e_i} - \gamma] \cdot \mathbb{E}[\prod_{j \neq i} (Z_{e_j} - \gamma)] = 0,$$

where the first equality follows from the independence property of the Z_e mentioned above.

Thus, $\mathbb{E}[Z^{2k}] \leq \sum_{(e_1,\ldots,e_{2k}) \text{ intersecting}} 1$. We conclude the proof by counting the number of intersecting tuples $(e_1, \ldots, e_{2k})$. Note that for every intersecting tuple $(e_1, \ldots, e_{2k})$, we have $\left|\bigcup_{j=1}^{2k} e_j\right| \leq 3k$. The number of intersecting tuples where every edge intersects exactly one other edge is $k^{\Theta(k)}n^{3k}$. Notice that every intersecting tuple that is not of this form has $\left|\bigcup_{j=1}^{2k} e_j\right| \leq 3k-1$. Thus the number of such intersecting tuples is at most $\binom{(3k)^2}{k} \cdot n^{3k-1} = k^{O(k)} \cdot n^{3k-1}$. Thus $\mathbb{E}[Z^{2k}]$ is at most $k^{O(k)} \cdot n^{3k}$, as desired. □

Now suppose Λ does not occur. Then we can bound C as follows:

$$\begin{aligned}
C &= \left| \mathop{\mathbb{E}}_{B} \left[\prod_{e \in \binom{V}{2}} e^{it(h_{\mathbf{A},\mathbf{A}'})_e B_e/\sigma} \right] \right| \\
&= \prod_{e \in \binom{V}{2}} \left| \mathop{\mathbb{E}}_{B_e} \left[e^{it(h_{\mathbf{A},\mathbf{A}'})_e B_e/\sigma} \right] \right| \\
&\leq \prod_{e \in \binom{V}{2}} \left| \left(1 - 8p(1-p) \cdot \left\| \frac{t(h_{\mathbf{A},\mathbf{A}'})_e}{2\pi\sigma} \right\|^2 \right) \right| \quad \text{(by Lemma 2.1)} \\
&\leq \prod_{e \in \binom{V}{2}|\neg\Lambda_e} \left| \left(1 - 8p(1-p) \cdot \left\| \frac{t(h_{\mathbf{A},\mathbf{A}'})_e}{2\pi\sigma} \right\|^2 \right) \right| \\
&= \prod_{e \in \binom{V}{2}|\neg\Lambda_e} \left| \left(1 - 8p(1-p) \cdot \left(\frac{t(h_{\mathbf{A},\mathbf{A}'})_e}{2\pi\sigma} \right)^2 \right) \right| \\
&\leq \prod_{e \in \binom{V}{2}|\neg\Lambda_e} \left| \left(1 - 8p(1-p) \cdot \left(\frac{t|U|^{0.49}}{2\pi\sigma} \right)^2 \right) \right| \quad \text{(since } |(h_{\mathbf{A},\mathbf{A}'})_e| \geq |U|^{0.49}\text{)} \\
&\leq e^{-\frac{|V|^2}{8} \cdot 8p(1-p) \cdot \left(\frac{t|U|^{0.49}}{2\pi\sigma} \right)^2}. \quad \text{(since } \Lambda \text{ did not occur)}
\end{aligned}$$

Now we use the fact that $|U| = |V| = n/2$, that $\sigma = \Theta(n^2)$ and that $n^{0.55} < t$. Thus $C \leq e^{-\Theta(n^{0.08})}$.

Thus, taking into account the probability of the bad event Λ, we get:

$$|\,\mathbb{E}[e^{itS/\sigma}]|^2 \leq O\left(\frac{k^{O(k)}}{n^k} \right) + e^{-\Theta(n^{0.08})} \ll \frac{1}{t^{100}},$$

(choosing $k = 200$), as desired.

2.9 Chapter Appendix

In this section we compute the moments of the random variable $Z_n \stackrel{\text{def}}{=} S_n - p^3\binom{n}{3}$.

Let Δ denote the set of $\binom{n}{3}$ triangles in K_n. For each $t \in \Delta$ denote X_t to be the indicator of the event that all edges in t appear. We write $t \sim t'$ if triangles t and t' share an edge. Note that if triangles t and t' do not share any edges, the random variables X_t and $X_{t'}$ are independent and

$$\mathbb{E}[(X_t - p^3)(X_{t'} - p^3)] = 0.$$

Lemma 2.13. *Let k be a positive integer. Let $C = C(p)$ be the constant $C(p) \stackrel{def}{=} \mathbb{E}[(X_t - p^3)(X_{t'} - p^3)]$ where t and t' are any two triangles that share exactly one edge. Then if k is odd*

$$\mathbb{E}[Z_n^k] = O(n^{2k-1})$$

and if k is even

$$\mathbb{E}[Z_n^k] = \frac{(n)_{2k} C^{k/2} (k-1)!!}{2^{k/2}} + O(n^{2k-1}).$$

Proof. We start with

$$E[Z_n^k] = \sum_{t_1 \in \Delta} \cdots \sum_{t_k \in \Delta} \mathbb{E}\left[\prod_{i=1}^{k} (X_{t_i} - p^3)\right].$$

We say an ordered tuple $(t_1, \cdots, t_k)$ of triangles is *intersecting* if for every i there is a $j \neq i$ for which $t_i \sim t_j$. Note that if $(t_1, \cdots, t_k)$ is not intersecting then there is an i for which the random variable X_{t_i} is independent with X_{t_j} for all $j \neq i$. Furthermore, for such a tuple

$$\mathbb{E}\left[\prod_{i=i}^{k} (X_{t_i} - p^3)\right] = 0.$$

We now split into cases based on the parity of k.

Case k is even:

Given an intersecting tuple we define its *skeleton* to be the subgraph of K_n obtained by taking the union of the triangles t_i. Let H be a graph on $2k$ vertices that consists of $k/2$ connected components, each component being the union of two triangles sharing a

single edge (although there are many such graphs H, note they are all isomorphic). We say a tuple $(t_1, \cdots, t_k)$ is *fully paired* if its skeleton is isomorphic to H. We first count the number of fully paired tuples by counting the number of copies of H that appear in K_n times the number of fully paired tuples whose skeleton is H.

To count the copies of H, first note that

$$\binom{n}{4}\binom{n-4}{4}\cdots\binom{n-2k+4}{4}\cdot\frac{1}{(k/2)!} = \frac{(n)_{2k}}{24^{k/2}(k/2)!}$$

counts the number of ways to choose the $k/2$ connected components. Within each component there are 6 choices of the shared edge of the two triangles, after which the two triangles are determined. Thus there are

$$\frac{(n)_{2k}6^{k/2}}{24^{k/2}(k/2)!} = \frac{(n)_{2k}}{2^k(k/2)!}$$

copies. For each copy there are $k!$ tuples whose skeleton is that copy. Thus the number of fully paired tuples is

$$\frac{(n)_{2k}k!}{2^k(k/2)!} = \frac{(n)_{2k}(k-1)!!}{2^{k/2}}.$$

For a fully paired tuples, the expression $\mathbb{E}\left[\prod_{i=1}^{k}(X_{t_i}-p^3)\right]$ splits as a product of the expectation of each connected component (which are pairwise independent). Thus,

$$\mathbb{E}\left[\prod_{i=1}^{k}(X_{t_i}-p^3)\right] = C^{k/2}. \tag{2.11}$$

We now quickly argue that the number of intersecting tuples that are not fully paired is $O(n^{2k-1})$. This follows because if a tuple is intersecting but not fully paired, than its skeleton consists of at most $2k-1$ vertices. There are $O(1)$ graphs on a given set of vertices, and given such a graph, there are $O(1)$ tuples whose skeleton is isomorphic to it (k is a constant). Thus there are

$$\sum_{i=3}^{2k-1} O(1)\binom{n}{i} = O(n^{2k-1}) \tag{2.12}$$

such intersecting graphs.

We then have the following calculation. Let P denote the set of fully paired tuples and Q denote the set of tuples that are intersecting but not fully paired.

$$\begin{aligned}\mathbb{E}(Z_n^k) &= \sum_{(t_1,\cdots,t_k)} \mathbb{E}\left[\prod_{i=1}^{k}(X_{t_i}-p^3)\right] \\ &= \sum_{(t_1,\cdots,t_k)\in P} \mathbb{E}\left[\prod_{i=1}^{k}(X_{t_i}-p^3)\right] + \sum_{(t_1,\cdots,t_k)\in Q} \mathbb{E}\left[\prod_{i=1}^{k}(X_{t_i}-p^3)\right] \\ &= \frac{(n)_{2k}C^{k/2}(k-1)!!}{2^{k/2}} + O(n^{2k-1}).\end{aligned}$$

Case k is odd:

Let Q denote the set of intersecting tuples. Note that if k is odd then there are no fully paired tuples of k triangles. Therefore $|Q| = O(n^{2k-1})$ and we have the following:

$$\begin{aligned}\mathbb{E}(Z_n^k) &= \sum_{(t_1,\cdots,t_k)} \mathbb{E}\left[\prod_{i=1}^{k}(X_{t_i}-p^3)\right] \\ &= \sum_{(t_1,\cdots,t_k)\in Q} \mathbb{E}\left[\prod_{i=1}^{k}(X_{t_i}-p^3)\right] \\ &= O(n^{2k-1}).\end{aligned}$$

□

Corollary 2.14. *Let* $\sigma_n^2 \stackrel{def}{=} \boldsymbol{Var}[S_n]$ *and let* $R_n \stackrel{def}{=} (S_n - p^3\binom{n}{3})/\sigma_n$. *Then* $\mathbb{E}[R_n^k] \to \mu_k$ *for all* k *fixed, where*

$$\mu_k = \begin{cases} (k-1)!! & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases}.$$

Chapter 3

Relationships between Block Sensitivity, Degree, and Certificate Complexity

Acknowledgment of a Journal Publication: Some parts of this chapter are closely related to work which is due (as of this writing) to appear in *Combinatorica* [GSS].

3.1 Overview of Complexity Measures

A *Boolean function* is a function $f : \{0,1\}^n \to \{0,1\}$. Throughout this chapter, we will study three complexity measures of boolean functions, namely *block sensitivity*, *degree*, and *certificate complexity*. We present the definitions of these measures as given in [BdW02], for these definitions let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function.

3.1.1 Sensitivity and Block Sensitivity

Sensitivity and block sensitivity both measure how sensitive f is to changes in the input x.

Given an input x and a set $B \subseteq [n]$ we use x^B to denote the input y where $y_i = x_i$ for $i \notin B$ and $y_i = x_i \oplus 1$ for $i \in B$ (where $\oplus$ indicates the mod-2 sum). A *block* for f at x is a set B for which $f(x) \neq f(x^B)$.

The *sensitivity of f at x*, $s_x(f)$ is the number of variables for which $f(x) \neq f(x^{\{i\}})$. Equivalently the sensitivity of f at x is the number of blocks of size 1. The *sensitivity of f* is $\max_x s_x(f)$.

The *block sensitivity of f at x*, denoted $bs_x(f)$ is the maximum number b for which there exists non-empty disjoint sets $B_1, B_2, \cdots, B_b$ which are blocks for f at x. The *block sensitivity of f*, $bs(f)$, is $\max_x bs_x(f)$. The *1-block sensitivity of f*, $bs_1(f)$, is

$\max_{x:f(x)=1} bs_x(f)$. The *0-block sensitivity* is defined similarly.

It is useful to have the notion of the *block hypergraph of f at x*, which is the hypergraph $\mathcal{H}$ consisting of all blocks for f at x. In this way the block sensitivity of f at x is the *packing number* of the hypergraph $\mathcal{H}$, denoted as $\nu(\mathcal{H})$.

3.1.2 Certificate Complexity

Certificate complexity intuitively measures, in the worst case, how many bits of an input x to a function must be given in order to fix the value of that function.

Let C be an assignment $C : S \to 0, 1$ of values to some subset S of the n variables of f. We say C is *consistent* with an input $x \in \{0,1\}^n$ if $x_i = C(i)$ for all $i \in S$.

For $b \in \{0,1\}$ a *b-certificate* for f is an assignment C for which $f(x) = b$ whenever x is consistent with C. The *size* of a certificate C is the cardinality of the set S.

The *certificate complexity of f at x*, denoted $C_x(f)$, is the size of the smallest certificate for f which is consistent with x. The *certificate complexity* of f is $C(f) \stackrel{\text{def}}{=} \max_x C_x(f)$. The *1-certificate complexity of f* is $C_1(f) \stackrel{\text{def}}{=} \max_{x:f(x)=1} C_x(f)$, the *0-certificate complexity of f*, denoted $C_0(f)$, is defined similarly.

In terms of the block hypergraph $\mathcal{H}$ of f at x, the certificate complexity of f at x is exactly the *vertex cover number* of $\mathcal{H}$, which is denoted $\tau(\mathcal{H})$. Along these lines it is useful to define a *witness of f at x* which is a subset W of the variables for which, if we fix the input x on W then f becomes a constant function. Equivalently, a witness is a set W for which $W \cap B \neq \emptyset$ for all blocks B (otherwise known as a *hitting set* for the block hypergraph).

3.1.3 Degree

A *multilinear polynomial* is a polynomial $p : \{0,1\}^n \to \{0,1\}$ of the form $p(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i$. It is well known that every boolean function f has a unique representation as a multilinear polynomial, the *degree* of f, $deg(f)$, is the degree of this polynomial.

3.1.4 Decision Tree Complexity

A *decision tree* is a rooted binary tree T where each internal node is labelled with a variable x_i and every leaf is labelled with a 0 or 1. One evaluates a tree T on an input x as follows: Start at the root. If the root is a leaf then stop and output the binary label on the leaf. Otherwise take the variable labelled on the root and query the bit x_i in the input, if it is 0 then recursively evaluate the left subtree, if it is 1 then evaluate the right subtree.

We say a decision tree *computes* f is its output equals $f(x)$ for all $x \in \{0,1\}^n$. There are many possible decision trees which compute a given function f. The *decision tree complexity of* f, $D(f)$, is the smallest depth of any tree T which computes f.

3.1.5 Boolean Function Composition and the Critical Exponent

There is a large body of work dedicated to proving bounds relating one complexity measure in terms of another. For example, Nisan showed that for every boolean function f, $C(f) \leq bs^2(f)$. Given an ordered pair of complexity measures (m_1, m_2) we define the *critical exponent of the pair* (m_1, m_2) to be the infimum of all r for which there exists constants a, b such that $m_1(f) \leq a m_2(f)^r + b$ for all Boolean functions f. For example, Nisan's result implies that the critical exponent for the pair (C, bs) is at most 2. Later in this chapter we will give two different constructions which show that Nisan's bound is tight, the critical exponent for (C, bs) is exactly 2.

In order to lower bound the critical exponent for a pair of measures (m_1, m_2), one must construct and infinite sequence of functions $(f_n)_1^\infty$ for which $m_1(f_n) = \Omega(m_2(f_n)^r)$. A natural way to construct such a sequence of functions is to use *boolean function composition*. Given boolean functions $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^m \to \{0,1\}$ their composition $f \circ g : \{0,1\}^{nm} \to \{0,1\}$ is defined as $f \circ g \stackrel{\text{def}}{=} f(g, g, g, \cdots, g)$ where there are n copies of g, each evaluating a separate input of m variables. To show separation between two measures one often starts with base function f and iteratively composes it to obtain a sequence of functions $(f^k)_1^\infty$, where $f^k \stackrel{\text{def}}{=} f \circ f^{k-1}$. If one can show that $m_1(f^k)$ grows significantly faster than $m_2(f^k)$, then one has shown separation

between two measures m_1 and m_2. One of the constructions given later in this chapter uses boolean function composition.

In order to analyze the growth rate of a sequence $m(f^k)$ it helps to define the following limit: For a Boolean function complexity measure m let

$$m^{\lim}(f) \stackrel{\text{def}}{=} \limsup m(f^k)^{1/k}.$$

Note that if $m_1^{\lim}(f) = a$ and $m_2^{\lim}(f) = b$ with $a > b$ then there exists a sequence of functions for which $m_1(f) = \tilde{\Omega}(m_2(f)^{a/b})$, where $\tilde{\Omega}(\cdot)$ hides logarithmic factors. In particular, such a sequence implies the critical exponent for the pair (m_1, m_2) is at least a/b.

3.1.6 Known Relationships Between Complexity Measures and Outline for the Chapter

Table 3.1.6 gives the best known bounds relating the complexity measures discussed so far. Each entry of the table gives the known bounds on the critical exponent for the pair (m_1, m_2) where m_1 is the measure on the row and m_2 is the measure on the column. For example, the entry $[\log_3(6), 2]$ on row $bs(f)$ and column $deg(f)$ indicates that it is known that $bs(f) = O(deg(f)^2)$ for all functions, and there exists an infinite sequence of functions for which $bs(f) = \Omega(deg(f)^{\log_3(6)})$. Although our definition of critical exponent ignores logarithmic factors, it turns out that for all pairs of measures given in the table, there is an upper bound (or lower bound) which determines a, b and r such that $m_1(f) \leq am_2(f)^r + b$ for all f (or $m_1(f) \geq am_2(f)^r + b$).

Table 3.1: Best known bounds for the smallest r s.t. $m_1(f) = O(m_2(f)^r)$ for all f.

	$bs(f)$	$C(f)$	$D(f)$	$deg(f)$
$bs(f)$	1	1	1	$[\log_3(6), 2]$
$C(f)$	2^*	1	1	$[\log_3(6), 3]$
$D(f)$	$[2, 3]$	2	1	$[\log_3(6), 3]$
$deg(f)$	$[2, 3]$	2	1	1

The rest of this chapter involves improving some of the bounds in this table. In Section 3.2 we investigate a potential construction of a family of functions which show separation between degree and other complexity measures (hoping to improve the $log_3(6)$

lower bounds in the table). We will prove that this construction cannot be used to separate $deg(f)$ and $bs(f)$. The question on whether or not it can separate $C(f)$ or $D(f)$ reduces to a question on weakly intersecting set systems due to Furedi which we were unable to solve.

In the last section we will construct an infinite family of functions for which $C(f) = \Omega(bs(f)^2)$. This gives a tight bound on the best exponent, and improves on the previous best known construction, due to Aaronson [Aar08], which gave $C(f) = \Omega(bs(f)^{\log_{4.5}(5)})$. This was a joint work with Michael Saks and Srikanth Srinivasan.

3.2 Investigating a Certain Family of Functions with Low Degree

In this section we explore a potential way of constructing functions which have low degree but $m(f)$ is large for some measure $m(\cdot) \in \{bs(\cdot), C(\cdot), D(\cdot)\}$. To do this we construct a large family of functions which all have low degree, and then hope to show that there is some function in this family for which another complexity measure is large. Unfortunately, it turns out that for all functions f in this particular family, $bs(f) \leq 2deg(f)$, and thus the construction cannot separate block sensitivity and degree. The question on whether or not functions in this family show large separation between $C(f)$ and $deg(f)$ reduces to a question due to Furedi on sets of weakly intersecting set systems. Now to define the family of functions. Given a sequence S of pairs of subsets of $[n]$,

$$S = (A_1, B_1), (A_2, B_2), \cdots, (A_m, B_m),$$

we say that S is *weakly intersecting* if the following hold:

- For all $i \neq j$, either $A_i \cap B_j \neq \emptyset$ or $A_j \cap B_i \neq \emptyset$.
- For all i, $A_i \cap B_i = \emptyset$.
- For all i, $A_i \neq \emptyset$.

Given a weakly intersecting sequence S define f_S to be the boolean function

$$f_S = \sum_i \prod_{j \in A_i} x_j \prod_{k \in B_i} (1 - x_k).$$

Note that the weakly intersecting property guarantees that f_S is boolean valued, since at most one of the summands $\prod_{j \in A_i} x_j \prod_{k \in B_i} \bar{x_k}$ can evaluate to 1 on any input. The assumption that $A_i \neq \emptyset$ for all i only added to make $f_S(\vec{0}) = 0$ (but is otherwise not required).

Proposition 3.1. *Let $S = (A_1, B_1), \cdots, (A_m, B_m)$ be weakly intersecting. Let $\mathcal{H}_A$ be the hypergraph consisting of the sets $A_1, \cdots, A_m$. Then the boolean function $f \stackrel{def}{=} f_S$ satisfies:*

1. $deg(f) \leq \max_i(|A_i| + |B_i|)$
2. $bs(f) \geq \nu(\mathcal{H}_A)$
3. $C(f) \geq \tau(\mathcal{H}_A)$

where $\nu(\mathcal{H})$ and $\tau(\mathcal{H})$ respectively denote the packing number and vertex cover number of $\mathcal{H}$.

Proof. The first part is trivial.

For the second part for each index i let $x^{(i)}$ be the input which is set to 1 on the indices in A_i and set to 0 on the indices in A_i^c. Then $f(x^{(i)}) = 1$ for all i. Since $f(\vec{0}) = 0$ this means that the sets A_i are all blocks for f at the all zero input. Thus the block sensitivity of f at $\vec{0}$ is the packing number of the hypergraph $\mathcal{H}_A$.

For the third part, since each set A_i is a block for f at $\vec{0}$, any witness W for f at $\vec{0}$ must intersect A_i for all i. In particular W must be a hitting set for the hypergraph $\mathcal{H}_A$. Thus $\tau(\mathcal{H}_A) \leq |W|$ which proves the proposition. □

Thus to separate degree with block sensitivity or certificate complexity, its enough to find a weakly intersecting system $S = (A_1, B_1), \cdots, (A_m, B_m)$ for which $|A_i|$ and $|B_i|$ are all small relative to $\nu(\mathcal{H}_A)$ or $\tau(\mathcal{H}_A)$.

Remark: The above proposition focuses on the all zero input. One may wonder whether or not there could be another input x for which $bs_x(f)$ or $C_x(f)$ is large, however in such a case we can complement some variables to obtain a function g for which $deg(g) = deg(f)$ and $bs_{\vec{0}}(g) = bs_x(f)$. It is easy to see as well that such a $g = g_{S'}$

for another weakly intersecting set system S'. In particular, we can separate degree and block sensitivity (or certificate complexity) for such a function if and only if one can find a weakly intersecting system S which separates $|A_i| + |B_i|$ and $\nu(\mathcal{H}_A)$ (or $\tau(\mathcal{H}_A)$).

Lemma 3.2. *Let $f = f_S$ for any weakly intersecting system $S = (A_1, B_1), \cdots, (A_m, B_m)$. Then $bs(f) \leq 2deg(f)$.*

Proof. Assume without loss of generality that the sets A_i are pairwise disjoint, (one can throw away all pairs (A_j, B_j) where A_j is not in the maximum sized packing of $\mathcal{H}_A$ and arrive a smaller sequence S' which is weakly intersecting and $\nu(\mathcal{H}_{A'}) = \nu(\mathcal{H}_A)$). Consider the digraph on vertex set $[m]$, where vertex i points to j if $A_i \cap B_j$. Since S is weakly intersecting, this digraph is a tournament, possibly with edges that point both directions. In particular, there must be a vertex (call it j) with indegree at least $m/2$. Then $B_j \cap A_i \neq \emptyset$ for at least $m/2$ sets A_i. Since the A_i are all disjoint, it follows that $B_j \geq m/2$. Thus $bs(f) = m$ and $deg(f) \geq m/2$. □

I was unable to determine whether or not such a function exists which separates degree and certificate complexity, the corresponding question for weakly intersecting set systems appeared in [Für88].

3.3 Achieving Quadratic Separation Between $C(f)$ and $bs(f)$

In this section we prove the following

Theorem 3.3. *For infinitely many $n \in \mathbb{N}$ there is a function $f : \{0,1\}^{n^2} \to \{0,1\}$ such that $bs(f) \leq bs^*(f) = O(n)$ and $C(f) = \Omega(n^2)$.*

We do this first with a probabilistic construction and second with a construction using Boolean function composition. The latter construction is slightly weaker in that the bound it implies ignores logarithmic factors.

3.3.1 A Probabilistic Construction

We start with the following

Proposition 3.4. *Let g be a non-constant boolean function and $f = OR_n \circ g$. Then for either complexity measure $m \in \{C, bs\}$ we have:*

$$\begin{aligned} m_1(f) &= m_1(g) \\ m_0(f) &= n \cdot m_0(g). \end{aligned}$$

Proof. Let I be the index set for the variables of g, so $J = [n] \times I$ is the index set for the variables of f. For $i \in [n]$, write J_i for the index subset $\{i\} \times I$.

First we show $m_1(f) = m_1(g)$. The function g is a subfunction of f (i.e., can be obtained from f by restricting some variables) so $m_1(f) \geq m_1(g)$ for each of the above complexity measures m. For the reverse inequality, we argue that $C_1(f) \leq C_1(g)$, the argument for block sensitivity is similar. Let $\alpha \in g^{-1}(1)$ be an input for which $C_\alpha(g)$ is maximum. Construct an input β for f by fixing the variables in J_n according to α and for each $i \in [n-1]$ fix the variables in J_i to some input y for g such that $g(y) = 0$. It is easy to check that $C_1(f) \leq C_\beta(f) = C_\alpha(g) = C_1(g)$.

Next we show that $m_0(f) = n \cdot m_0(g)$. For this, write an assignment to the variables of f as $\alpha^1, \ldots, \alpha^n$ where each α^i is an assignment to the variables of g. We have $f(\alpha^1, \ldots, \alpha^n) = 0$ if and only if $g(\alpha^1) = \cdots = g(\alpha^n) = 0$. It is easy to check that for each of the measures m under consideration, if $g(\alpha^1) = \cdots = g(\alpha^n) = 0$ then $m_{\alpha^1,\ldots,\alpha^n}(f) = m_{\alpha^1}(g) + \cdots + m_{\alpha^n}(g)$. Thus an input in $f^{-1}(0)$ that maximizes $m_{\alpha^1,\ldots,\alpha^n}(f)$ is one for which $\alpha^1 = \cdots = \alpha^n = \alpha$, where α satisfies $m_0(g) = m_\alpha(g)$. This gives $m_0(f) = n \cdot m_0(g)$.

□

We will now construct a sequence of n-variate functions g_n (for n sufficiently large) such that $C_0(g_n) = \Omega(n)$ and $bs_0(g_n) = O(1)$. If we then define $f_n = OR_n \circ g_n$ we may apply Proposition 3.4 to conclude that

$$C(f_n) \geq C_0(f_n) = n \cdot C_0(g_n) = \Omega(n^2),$$

while

$$bs(f_n) \leq \max(bs_0(f_n), bs_1(f_n)) \leq \max(nbs_0(g_n), bs_1(g_n)) = O(n).$$

This will imply Theorem 3.3.

Let us write $\delta(x, y)$ to denote the Hamming distance between $x, y \in \{0,1\}^n$. We define $g = g_n : \{0,1\}^n \to \{0,1\}$ as follows (we view n as being sufficiently large). Choose $x_1, \ldots, x_N \in \{0,1\}^n$ uniformly at random (with replacement) with $N = 2^{n/50}$. We set $g(x_i) = 1$ for each i, and $g(x) = 0$ otherwise.

Claim 3.5. *With high probability, for all i, j distinct $\delta(x_i, x_j) \geq \frac{n}{100}$.*

Proof. Let $A_{i,j}$ denote the event $\delta(x_i, x_j) < \frac{n}{100}$. Let x be a fixed point in $\{0,1\}^n$ and $B(x, r)$ denote the Hamming ball of radius r and center x. Then $|B(x, r)| = \sum_{i=0}^{r} \binom{n}{i}$. Thus we have

$$B\left(x, \frac{n}{100}\right) < 2\binom{n}{n/100} \leq 2(100e)^{n/100} < 2^{n/10}.$$

These inequalities imply that

$$\mathbf{P}(A_{i,j}) = \frac{B\left(x, \frac{n}{100}\right)}{2^n} < 2^{-9n/10}.$$

By the union bound the hypothesis fails with probability at most

$$2^{-9n/10}\binom{N}{2} = o(1).$$

□

If the hypothesis of the claim holds and $g(x) = 0$, then all but possibly one of the blocks for g at x will have size at least $\frac{n}{200}$. Thus, at most 200 blocks can be packed and $bs_0(g) \leq 200$. Likewise, this bound on the size of blocks implies that $bs_0^*(g) \leq 200$.

We now argue that all sufficiently large subcubes of $\{0,1\}^n$ will contain a 1 of g almost surely.

Claim 3.6. *With high probability, $C_0(g) \geq \frac{n}{100}$.*

Proof. It is enough to show that every subcube of co-dimension $\frac{n}{100}$ will contain a y such that $g(y) = 1$. For each S which is a subcube of co-dimension $\frac{n}{100}$, denote A_S as the event $g(x) = 0$ for all $x \in S$. Then

$$\mathbf{P}(A_S) \leq (1 - 2^{-n/100})^N < \exp(-\frac{N}{2^{n/100}}) = \exp(-2^{n/100})$$

There are $\binom{n}{n/100}2^{n/100} < 2^{2n}$ subcubes of co-dimension $\frac{n}{100}$. Thus by union bound the hypothesis fails with probability at most

$$\exp(-2^{n/100})2^{2n} = o(1).$$

□

We have shown, for sufficiently large n, that with high probability a random function g satisfies $bs_0^*(g) \leq 200$ and $C_0(g) \geq \frac{n}{100}$. Thus for each n sufficiently large, there exists a function g_n with this property.

3.3.2 A Construction Using Iterated Composition

In this section we construct a function f on n variables for which $C^{lim}(f) \geq \frac{n}{2}$ and $(bs)^{\lim}(f) \leq 4\sqrt{n}$. Thus, for any $\epsilon > 0$, we may choose n large enough to conclude that the critical exponent is at least $2 - \epsilon$ (and thus the exponent is 2). It is highly nontrivial to analyze the growth rate of each complexity measure when we compose the function f. Luckily, Michael Saks and Srikanth Srinivasan, gave a (somewhat simple) characterization of $bs^{\lim}(f)$ and $C^{\lim}(f)$ for any f. We will first need some additional definitions in order to understand their characterization.

Some Additional Preliminaries

We will need to define another complexity measure of Boolean functions. Recall that, given a function f and input x a *witness* for f at x is a subset W of the variables for which, if we fix the input to agree with x then f becomes a constant. The measure $C_x(f)$ was then defined to be the size of the smallest witness for f at x. Equivalently, W is a witness if $W \cap B \neq \emptyset$ for all blocks B for f at x. A *fractional witness* for f at x is weighting of the input variables $W : [n] \to [0, 1]$ such that for each block B for f at x,

$$\sum_{j \in B} W(j) \geq 1.$$

The size of a fractional witness W is $\sum_{j \in [n]} W(j)$. This is the natural generalization of witnesses, in that a witness is just a fractional witness where the function W is

required to be Boolean valued. The *fractional certificate complexity of f at x*, denoted $C^*_x(f)$, is the size of the smallest fractional witness for f at x. The *fractional certificate complexity of f*, denoted $C^*(f)$, is $\max_x C^*_x(f)$.

Consider the following definitions:

- An *input pair for f* is a pair of inputs (α_0, α_1) where $f(\alpha_0) = 0$ and $f(\alpha_1) = 1$. Such an α_0 is called a *0-input* for f and α_1 is a *1-input*.

- Given an input pair (α_0, α_1) a *witness pair for f at the pair* (α_0, α_1) is a pair (W_0, W_1) where W_0 is a witness for f at α_0 and W_1 is a witness for f at α_1.

- A *fractional witness pair* is defined analogously as a witness pair.

- Given a witness W for f at input x, we define its *profile* $\vec{p}(x, W)$ to be the vector $\vec{p}(x, W) \stackrel{\text{def}}{=} (p_0, p_1)$ where p_0 is the number of elements $i \in W$ such that $x_i = 0$ and p_1 is the number of $i \in W$ such that $x_i = 1$.

- Given a fractional witness W^* for f at input x its *profile* is the vector $\vec{p}(x, W^*) = (p_0, p_1)$ where $p_0 = \sum_{i:x_i=0} W^*(i)$ and $p_1 = \sum_{i:x_i=1} W^*(i)$.

- Given a witness pair (W_0, W_1) for f at input pair (α_0, α_1) its *profile matrix* is the matrix
$$M(\alpha_0, \alpha_1, W_0, W_1) \stackrel{\text{def}}{=} \begin{bmatrix} \vec{p}(\alpha^0, W_0) \\ \vec{p}(\alpha^1, W_1) \end{bmatrix}.$$
Note that profile matrices are 2 by 2 matrices with non-negative entries.

- The *profile matrix* of a fractional witness pair (W^*_0, W^*_1) at input pair (α_0, α_1) is defined analogously.

- Given a 2 by 2 matrix M with non-negative entries we define $\rho(M)$ to be its largest eigenvalue. It follows from Perron-Frobenious theory that for M with non-negative entries, $\rho(M) \geq 0$.

Michael Saks and Srikanth Srinivasan showed that $bs^{\lim}(f) = (C^*)^{\lim}(f)$ for all f and then gave the following characterizations for $(C^*)^{\lim}(f)$ and $C^{\lim}(f)$:

Theorem 3.7. *For any boolean function f,*

$$bs^{\lim}(f) = (C^*)^{\lim}(f) = \max_{(\alpha_0,\alpha_1)} \min_{(W_0,W_1)} \rho(M(\alpha_0,\alpha_1,W_0,W_1),$$

where the max is over all input pairs (α_0,α_1) for f and the min is over all witness pairs (W_0,W_1) for f at (α_0,α_1).

Theorem 3.8. *For any boolean function f,*

$$C^{\lim}(f) = \max_{(\alpha_0,\alpha_1)} \min_{(W_0^*,W_1^*)} \rho(M(\alpha_0,\alpha_1,W_0^*,W_1^*),$$

where the max is over all input pairs (α_0,α_1) for f and the min is over all fractional witness pairs (W_0^*,W_1^*) for f at (α_0,α_1).

The Construction

We will now use the above characterization to analyze $(C^*)^{\lim}(f)$ and $C^{\lim}(f)$ for the following function. In what follows we use $|x|$ to denote the hamming weight of an input $x \in \{0,1\}^n$.

Let d,k,n be positive integers such that $n \geq k \geq d$, $d \mid k$, and $k \mid n$. We define $f : \{0,1\}^n \to \{0,1\}$ to be the following boolean function on n variables:

View the n indices of the input x as being divided into $\frac{n}{k}$ disjoint groups, with each group containing k indices. We set $f = 1$ if and only if $|x| \geq d$ and all the 1's in x can be found in a single group. Note that $f(x) = 1$ implies $|x| \leq k$.

Although f itself shows no separation between $bs(f)$ and $C(f)$, the key is that both the zero and one certificate complexity for f are large, while the zero block sensitivity is small. Also, any 1-assignment for f contains many 0 indices.

In the following analysis, we assume n is an even perfect square and set $k := 2\sqrt{n}$ and $d := \sqrt{n}$. We wish to bound $C^{lim}(f)$ and $(C^*)^{\lim}(f)$ using Theorems 3.7 and 3.8.

Claim 3.9. *For the boolean function $f : \{0,1\}^n \to \{0,1\}$ defined above we have:*

$$(C^*)^{\lim}(f) \leq 4\sqrt{n}.$$

Proof. We proceed by showing that for any input pair (α^0,α^1) for f, we can find a pair of fractional witnesses (W_0^*,W_1^*) such that profile matrix $M(\alpha_0,\alpha_1,W_0^*,W_1^*)$ has both

eigenvalues less than $4\sqrt{n}$. For each possible α_0 we exhibit a small fractional witness W_0^*.

Case 1, $\alpha_0 = (0, 0, \ldots, 0)$:

Here we choose $W_0^* := (\frac{1}{d}, \frac{1}{d}, \cdots, \frac{1}{d})$. It follows that W_0^* is a fractional witness as each block for this input has size at least d. The profile vector $\vec{p}(\alpha_0, W_0^*) = (\frac{n}{d}, 0)$.

Case 2, $|\alpha_0| = j$, and all 1's in α_0 appear in the same group:

Note this means that $j < d$ as α_0 is a 0-input. Let X_1 be the set of indices for α_0 which are 1's, let G_1 be the group which contains X_1. Pick an $s \in X_1$, we define a fractional hitting set W_0^* to assign weight 1 to s, weight 1 to all indices in $G_1 \setminus X_1$, and weight 0 otherwise. To see that W_0^* is indeed a hitting set, note that if B is a block for α_0, then either $B \subseteq G_1$ or $X_1 \subset B$. If $X_1 \subset B$, then $s \in B$ and it has been assigned weight 1. If $B \subseteq G_1$ then B must contain an index in G_1 which is labeled 0 as $|\alpha_0| < d$, this index was assigned weight 1 by W_0^*. Thus W_0^* is a fractional witness and the profile vector $\vec{p}(\alpha^0, W_0^*) = (k - j, 1) \leq (k, 1)$.

Case 3, At least two different groups in α_0 contain 1's:

Let G_1, G_2 be two distinct groups containing 1's. Let X_1, X_2 be the set of indices which are assigned 1 by α_0 in G_1, G_2 respectively. Then if B is a block for α_0, either $X_1 \subseteq B$ or $X_2 \subseteq B$. We define W_0^* to assign weight 1 to an index in X_1 and in index in X_2. This will be a fractional witness, and the profile vector $\vec{p}(\alpha_0, W_0^*) = (0, 2)$. This concludes the analysis of each possible 0 assignment.

The 1-inputs α_1:

If α_1 is a 1 assignment then $|\alpha_1| \geq d$ and all the 1's appear in a single group, call it G_1. In this case we define W_1^* to assign weight 1 to all indices outside G_1, and weight 1 to d indices in G_1 which are assigned 1 by α_1. This will be a fractional witness as any block must contain a 0 index outside of G_1 or leave less than d 1's inside of G_1 after flipping the indices in B. Here the profile vector $\vec{p}(\alpha_1, W_1^*) = (n - k, d)$.

If M, M' are 2×2 matrices with non-negative entries, and $M \leq M'$ entry by entry, then $\rho(M) \leq \rho(M')$. Considering this along with the 3 cases of 0 assignments above, bounding $(C^*)^{\lim}(f)$ reduces to bounding the largest eigenvalues of the following

matrices:

$$\begin{bmatrix} \frac{n}{d} & 0 \\ n-k & d \end{bmatrix} \quad \begin{bmatrix} k & 1 \\ n-k & d \end{bmatrix} \quad \begin{bmatrix} 0 & 2 \\ n-k & d \end{bmatrix}$$

We set $k = 2\sqrt{n}$ and $d = \sqrt{n}$. For such k, d the second matrix has the largest eigenvalue of the three and has largest eigenvalue less than $4\sqrt{n}$. □

Claim 3.10.

$$C^{\lim}(f) \geq \frac{n}{2}.$$

Proof. To prove this we choose an input pair (α_0, α_1) for which all witness pairs (W_0, W_1) satisfy $\rho\left(M(\alpha_0, \alpha_1, W_0, W_1)\right) \geq \frac{n}{2}$. We set $\alpha_0 \stackrel{\text{def}}{=} (0, 0, \cdots, 0)$ and α_1 to have exactly d 1's in the first group, and be identically 0 in every other group.

Any witness for α_0 must contain $k-d+1$ indices in each group, thus must have size at least $\frac{n}{k}(k-d+1)$. It follows that any minimum sized witness W_0 yields the profile vector $\vec{p}(\alpha_0, W_0) = (\frac{n}{k}(k-d+1), 0)$.

Likewise, any witness W_1 for α_1 must contain all 1 indices (there are d of them), and contain all the 0 indices outside the unique group containing the 1's. Thus any minimal profile vector $\vec{p}(\alpha_1, W_1) = (d, n-k)$. The claim then reduces to looking at the maximum eigenvalue of the matrix

$$A = \begin{bmatrix} \frac{n}{k}(k-d+1) & 0 \\ n-k & d \end{bmatrix}.$$

When $k = 2\sqrt{n}$ and $d = \sqrt{n}$ this matrix has an eigenvalue larger than $\frac{n}{2}$.

□

Chapter 4

Bounds for Randomized Decision Tree Complexity

We will consider two natural probabilistic models of computation which extend the decision tree model of computation discussed in the previous chapter, zero-error and two-sided-error randomized decision tree complexity.

4.1 Introduction and Definitions

A *randomized decision tree* is a distribution $\mu(T)$ of decision trees. Given a decision tree T, its *cost* on an input x, $\text{cost}(T, x)$, is the number of variables which are queried when T evaluates x. The *cost* of a randomized decision tree $\mu(T)$ on an input x is defined as,

$$\text{cost}(\mu, x) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{T\sim\mu} \text{cost}(T, x).$$

The *zero-error randomized decision tree complexity* of a boolean function f is defined as

$$R(f) \stackrel{\text{def}}{=} \min_{\mu} \max_{x} \text{cost}(\mu, x)$$

where the minimum is over all distributions which are supported on decision trees that compute f.

We say a randomized decision tree *computes f with two-sided error* if for all inputs x its output is equal to $f(x)$ with probability at least $2/3$. The *two-sided-error randomized decision tree complexity* of a function f is defined as

$$R_2(f) \stackrel{\text{def}}{=} \min_{\mu(T)} \max_{x} \text{cost}(\mu(T), x),$$

where the minimum is over all randomized decision trees which compute f with two-sided error.

It is natural to wonder what bounds can be proven relating $R(f)$ and $R_2(f)$ for general functions f. Clearly $R_2(f) \leq R(f)$ for all f, but how much smaller can $R_2(f)$ be? The following trivial algorithm shows that $R_2(f) \leq R(f)/3$ for all f: With probability $1/3$ output 0, with probability $1/3$ output 1, and with probability $1/3$ run the optimal zero-error algorithm A and output the result. So far, we have not been able to find a two-sided error algorithm for a function f which beats the the above trivial algorithm.

Santha [San95] proved that the trivial algorithm is best possible for a certain class of boolean functions called "balanced read-once AND-OR trees f". In Section 4.3, we mention an approach which could possibly show that $R(f) = O(R_2(f)\log(n))$ for all functions monotone functions f, however we were unable to make much progress in this direction.

4.2 Using Boolean Function Composition to Separate $R(f)$ and $R_2(f)$

As discussed in Chapter 3, one potential method for constructing an example with large separation is to use boolean function composition. In this Section however, we prove a theorem which suggests that boolean function composition cannot be used for this purpose.

Theorem 4.1. *For any boolean function f,*

$$R^{\lim}(f) \leq \max(R_2^{\lim}(f), C^{\lim}(f)).$$

It was shown by Nisan [Nis91] that for every boolean function f,

$$bs(f) \leq 3R_2(f).$$

Also it is well known that $bs(f) = C(f)$ for all monotone functions, thus for monotone functions $R_2^{\lim}(f) \geq C^{\lim}(f)$. Therefore as a Corollary to (4.1) we get

Corollary 4.2. *For every monotone boolean function f it holds that,*

$$R^{\lim}(f) = R_2^{\lim}(f).$$

In particular, one cannot achieve polynomial separation between $R(f)$ and $R_2(f)$ by composing a monotone boolean function.

It is believed that $R_2(f) = \Omega(C(f))$ for all functions. If this is true than indeed Theorem 4.1 implies that boolean function composition cannot be used to obtain polynomial separation between $R(f)$ and $R_2(f)$ for *any* function.

To prove (4.1) we will first prove the following

Lemma 4.3. *For any two boolean functions f and g, where f is a function on n variables, it holds that*

$$R(f \circ g) \leq R_2(g)n\log(R(g)n^2) + C(f)R(g) + 1.$$

Proof. Let k be the number of variables for the function g. Let A be an optimal two-sided error algorithm for g (so that $R_2(g)$ is the worst case cost of A on any input). Likewise let B be an optimal zero-sided error algorithm for g. We use A and B to construct a zero error randomized algorithm for $f \circ g$. The algorithm is as follows: Let $X = (X_1, \cdots, X_n) \in \{0,1\}^{nk}$ be an arbitrary input to the function $f \circ g$, where each $X_i \in \{0,1\}^k$. Let $y = (g(X_1), \cdots, g(X_n)) \in \{0,1\}^n$. Our algorithm first runs A on each of the inputs X_i $48\log(R(g)n^2)$ times. Let y_i' be the majority of the outputs given by running A on X_i. By chernoff bound

$$\mathbb{P}(y_i' \neq y_i) \leq e^{-(3/4)48\log(R(g)n^2)(2/3-1/2)^2} = \frac{1}{R(g)n^2}.$$

By a union bound,

$$\mathbb{P}(\exists i \text{ s.t.} y_i \neq y_i') \leq \frac{1}{nR(g)}.$$

Let $S \subseteq [n]$ be a set such that fixing y' on S is a certificate for f. Note we may choose S such that $|S| \leq C(f)$. The algorithm then runs B on X_i for each i in S. In the event that $y' = y$, we will find a certificate for $f \circ g$. If, after running B on each X_i for $i \in S$, we still haven't found a certificate we simply run B on the remaining X_i. Conditioning on $y' = y$, the above algorithm has average cost of at most $R_2(g)n\log(R(g)n^2) + C(f)R(g)$. Conditioning on $y' \neq y$ then it may have running time up to $R_2(g)n\log(R(g)n^2) + nR(g)$. Therefore the worst case expected cost of the above algorithm is at most $R_2(g)n\log(R(g)n^2) + C(f)R(g) + 1$ □

We are now ready to prove the main theorem in this section.

Theorem 4.4. *For any boolean function f,*

$$R^{\lim}(f) \leq \max(R_2^{\lim}(f), C(f)).$$

Proof. Let $L_1 = \limsup R(f^k)^{1/k}$ and $L_2 = \limsup R_2(f^k)^{1/k}$. If $L_1 = L_2$ we are done, so assume $L_1 > L_2$ (note it is trivial that for any f, $R(f) \geq R_2(f)$). It now suffices to show that for any $\epsilon > 0$, $L_1 \leq C(f) + \epsilon$. Let $\epsilon > 0$ be fixed and pick $\epsilon' > 0$ such that $\frac{L_2+\epsilon'}{L_1-\epsilon'} < 1$ and $\epsilon' < \epsilon/3$, (since $L_1 > L_2$ this can be done). Pick K' such that $k > K'$ implies that (recall n is a constant here)

$$\left(\frac{L_2 + \epsilon'}{L_1 - \epsilon'}\right)^{k-1} (k+2)n\log(n) < \frac{\epsilon}{3}.$$

Recall that the sequence $R(f^k)$ is monotone increasing. We may also assume that $R(f^k) \to \infty$, as otherwise we would have $L_1 = 0$. Therefore, we may pick K'' such that $k > K''$ implies that $\frac{1}{R(f^{k-1})} < \epsilon/3$. Finally, let $K = \max(K', K'')$. We first prove the following

Claim 4.5. *There exists $k > K$ such that the following hold:*

- $R(f^k) \geq (L_1 - \epsilon')R(f^{k-1})$
- $R(f^k) \geq (L_1 - \epsilon')^k$.

Proof. **Case 1:** For all $k > K$ it holds that $R(f^k) \geq (L_1 - \epsilon')^k$.

Then since $\limsup R(f^k)^{1/k} = L_1$ it follows that there must exist $k > K$ such that $R(f^k) \geq (L_1 - \epsilon')R(f^{k-1})$, otherwise the limit would be strictly smaller. This k satisfies the claim.

Case 2: There exists some $k_0 > K$ such that $R(f^k) < (L_1 - \epsilon')^{k_0}$.

In this case let k be the smallest $k > k_0$ such that $R(f^k) \geq (L_1 - \epsilon')^k$. Then for this k, $R(f^{k-1}) < (L_1 - \epsilon')^{k-1}$ and therefore

$$\frac{R(f^k)}{R(f^{k-1})} \geq (L_1 - \epsilon').$$

This k then satisfies the conditions of the claim. □

Now for the k guaranteed by the claim and applying Lemma (4.3) we have

$$R(f^{k-1})(L_1 - \epsilon) \leq R(f^k) \leq R_2(f^{k-1})n\log(R(f^{k-1})n^2) + C(f)R(f^{k-1}) + 1.$$

Note that $R(f^{k-1}) \leq n^{k-1}$ trivially (its bounded by the number of variables of f). Simplifying, the above implies that

$$L_1 - \epsilon' \leq \frac{R_2(f^{k-1})}{R(f^{k-1})}(k+2)n\log(n) + C(f) + \frac{1}{R(f^{k-1})}.$$

By the properties of k we have

$$L_1 - \epsilon' \leq \left(\frac{L_2 + \epsilon'}{L_1 - \epsilon'}\right)^{k-1}(k+2)n\log(n) + C(f) + \frac{1}{R(f^{k-1})}.$$

Which implies

$$L_1 \leq C(f) + \epsilon/3 + \epsilon/3 + \epsilon/3.$$

This concludes the proof.

Remark: A corollary of the above argument yields the result

$$R^{\lim}(f) \leq \max(R_2^{\lim}(f), C^{\lim}(f)).$$

To see this we pick a large $k = k(\epsilon)$ such that $C(f^k)^{1/k} < C^{\lim}(f) + \epsilon$. We then set $F = f^k$ and apply the theorem on F to get

$$R^{\lim}(F) \leq \max(R_2^{\lim}(F), C(F)).$$

Then since $m^{\lim}(F) = \left(m^{\lim}(f)\right)^k$ for any complexity measure (which follows from the definition of $m^{\lim}(\cdot)$), we have

$$R^{\lim}(f)^k \leq \max(R_2^{\lim}(f)^k, (C^{\lim}(f) + \epsilon)^k).$$

This implies

$$R^{\lim}(f) \leq \max(R_2^{\lim}(f), (C^{\lim}(f) + \epsilon)).$$

This holds for any $\epsilon > 0$ so the claim follows. □

4.3 $R_2(f)$ for read once AND-OR trees

Santha [San95] showed that $R_2(f) = R(f)/3$ for all balanced read once AND-OR trees (see their paper for a definition of balanced). In this section we propose an approach for showing that $R(f) \leq O(\log(n)R_2(f))$ for all read once AND-OR trees. This would improve the current best known bound which is $R(f) \leq R_2(f)^2$.

Given a boolean function f, we say an input x is *critical* if it satisfies the following property: View f as a boolean formula with AND and OR gates all of fan-in 2. An input x is critical if upon evaluating the formula, no AND gate has both inputs set to 0, and no OR gate has both inputs set to 1. These inputs seem intuitively harder than non-critical inputs in that algorithms evaluating f on a non-critical input are more likely to determine the value of an AND (OR) gate by having one of its sub trees evaluate to a 0 (1), thus not having to query variables in the other sub tree.

Recall that the $R(f)$ is defined as $\min_{\tilde{A}} \max_{x}$ where the minimum is over all randomized algorithms (i.e. distributions over decision trees computing f) and the max is over all inputs x. We define the *critical zero error randomized decision tree complexity*, denoted $R_c(f)$, to be $\min_{\tilde{A}} \max_{x \text{ critical}} cost(\tilde{A}, x)$. Trivially, $R_c(f) \leq R(f)$, however it seems reasonable to believe that the hardest input for the best randomized algorithm should be critical. If this were true then we would have $R_c(f) = R(f)$. The difficulty in proving this statement though is that there exist algorithms for which the hardest input is not critical, however such algorithms seem to be very inefficient and have very high cost.

For a function f and input x a *minimal block* is a block B for f at x such that there is no block B' for f at x which satisfies $B' \subset B$. Note that critical inputs for read once AND-OR trees satisfy that their minimal blocks all have size 1.

Remark: A natural generalization for arbitrary boolean functions is to define an input x to be critical if all of the minimal blocks for f at x have size 1. I believe it may be possible to show that the hardest input for the best randomized algorithm for a monotone boolean function should be critical. The intuition is similar in that non-critical inputs may allow the algorithm to terminate prematurely because there will be multiple certificates which may be found by the algorithm, whereas critical inputs have

a unique certificate.

The main result in this section is

Theorem 4.6. *For any boolean function f,*

$$R_c(f) = O(R_2(f)\log(n).$$

Remark: We also conjecture that for any read-once AND-OR tree (or possibly any monotone boolean function), $R_c(f) = R(f)$. In particular the hardest distribution can be assumed to be supported over critical inputs. If this is true then Theorem 4.6 would imply that $R(f) = O(R_2(f)\log(n)$ for all read-one AND-OR trees (or monotone boolean functions). Despite much effort though we were unable to prove this statement.

Proof. Let A be the best randomized two-sided error algorithm. Our randomized zero error algorithm for f is to independently run A $\Theta(\log(n))$ times, and if a certificate has still not been found, query the remaining variables in some arbitrary order. We will show that the probability that a certificate has not been found after these repeated trials is at most $1/n$. Thus the cost of our zero error algorithm in the average case is at most $\Theta(\log(n))R_2(f) + n \cdot (1/n)$.

Let x be any critical input for the function f. Since x is critical, all the min-blocks for f at x have size 1. Furthermore if our algorithm queries all of the minimal blocks, then a certificate for f will have been found.

Since A outputs the correct value of $f(x)$ with probably $2/3$, the algorithm A must query each min-block for f at x with probability at least $1/3$ (otherwise it would not be able to distinguish the inputs x and $x \oplus e_B$. Thus running A k times, the probability that there is a min-block which is not queried is at most (by a union bound) $(2/3)^k n$. Thus if $k = log_{2/3}(n^2) = \Theta(\log(n))$, the probability a certificate is not found is at most $1/n$. This concludes the proof. □

Chapter 5

A New Approach To the Sensitivity Conjecture

Acknowledgement of a Journal Publication: This chapter will be very similar to the journal version of this work [GKS], however it includes an extra section that discusses "two-stage protocols". This was a joint work with Michael Saks and Michal Koucky. Although the entire chapter should be considered as a joint work, some parts use ideas which are due entirely to Michael Saks or Michal Koucky and we have indicated when appropriate when this is the case.

5.1 Introduction

In this chapter we discuss a novel approach to the sensitivity conjecture. This conjecture asserts that there exists a constant k so that $deg(f) = O(s(f)^k)$ for all boolean functions f (the complexity measures $deg(f)$ and $s(f)$ were defined in Chapter 3, however we will redefine them in Section 5.1.3).

5.1.1 A Communication Game

The focus of this work is a somewhat unusual cooperative two player communication game. The game is parameterized by a positive integer n and is denoted G_n. Alice receives a permutation $\sigma = (\sigma_1, \ldots, \sigma_n)$ of $[n] = \{1, \ldots, n\}$ and a bit $b \in \{0, 1\}$ and communicates to Bob in a very restricted way (which will be described momentarily). Bob receives the message from Alice and then outputs a subset J of $[n]$ that is required to include σ_n, the last element of the permutation. The cost to Alice and Bob is the size of the set $|J|$.

The communication from Alice to Bob is constrained as follows: Alice has a memory vector $\mathbf{v}$ consisting of n cells which we will refer to as *locations*, where each location v_ℓ

is either empty, denoted by $v_\ell = *$, or is set to 0 or 1. Initially all locations are empty. Alice gets the input as a data stream $\sigma_1, \ldots, \sigma_n, b$ and is required to fill the cells of $\mathbf{v}$ in the order specified by σ. After receiving σ_i for $i < n$, Alice fills location σ_i with 0 or 1. Upon receiving σ_n and b, Alice writes b in location σ_n.

Once $\mathbf{v}$ is filled, Bob inspects $\mathbf{v}$ and outputs the subset J.

Given a protocol Π for this game, the cost of the protocol $c(\Pi)$ is the maximum of the output size $|J|$ over all inputs $\sigma_1, \ldots, \sigma_n, b$.

For example, consider the following protocol. Let $k = \lceil \sqrt{n} \rceil$. Alice and Bob fix a partition of the locations of $\mathbf{v}$ into k blocks each of size at most k. Alice fills $\mathbf{v}$ as follows: When σ_i arrives, if σ_i is the last location of its block to arrive then fill the entry with 1 otherwise fill it with 0.

Notice that if $b = 1$ then the final vector $\mathbf{v}$ will have a single 1 in each block. If $b = 0$ then $\mathbf{v}$ will have a unique all 0 block.

Bob chooses J as follows: if there is an all 0 block, then J is set to be that block, and otherwise J is set to be the set of locations containing 1's. It is clear that $\sigma_n \in J$ and so this is a valid protocol. In all cases the size of J will be at most k and so the cost of the protocol is $\lceil \sqrt{n} \rceil$. We will refer to this protocol as the AND-OR protocol. In Section 5.2.1 we remark on this protocol's connection to the boolean function

$$\text{AND-OR}(x) = \bigwedge_{i=1}^{\sqrt{n}} \bigvee_{j=1}^{\sqrt{n}} x_{ij}.$$

Let us define $C(n)$ to be the minimum cost of any protocol for G_n. We are interested in the growth rate of $C(n)$ as a function of n. In particular, we propose:

Question 5.1. *Is there a $\delta > 0$ such that $C(n) = \Omega(n^\delta)$?*

5.1.2 Connection to the Sensitivity Conjecture

Why consider such a strange game? The motivation is that the game provides a possible approach to the well known *sensitivity conjecture* from boolean function complexity.

Recall that the sensitivity of an n-variate boolean function f at an input $\mathbf{x}$, denoted $s_f(\mathbf{x})$, is the number of locations ℓ such that if we flip the bit of $\mathbf{x}$ in location ℓ then

the value of the function changes. (Alternatively, this is the number of neighbors of $\mathbf{x}$ in the hamming graph whose f value is different from $f(\mathbf{x})$.) The sensitivity of f, $s(f)$, is the maximum of $s_f(\mathbf{x})$ over all boolean inputs $\mathbf{x}$.

The degree of a function f, $deg(f)$, is the smallest degree of a (real) polynomial p in variables $x_1, \ldots, x_n$ that agrees with f on the boolean cube.

Conjecture 5.2. *(The Sensitivity Conjecture) There is a $\delta > 0$ such that for any boolean function f, $s(f) \geq \Omega(deg(f)^{\delta})$.*

An easy argument (given in Section 5.2) connects the cost function $C(n)$ of the game G_n to the sensitivity conjecture:

Proposition 5.3. *For any boolean function on n variables, $s(f) \geq C(deg(f))$.*

In particular, an affirmative answer to Question 5.1 would imply the sensitivity conjecture.

5.1.3 Background on the Sensitivity Conjecture

Sensitivity and degree belong to a large class of complexity measures for boolean functions that seek to quantify, for each function f, the amount of knowledge about individual variables needed to evaluate f. Other such measures include decision tree complexity and its randomized and quantum variants, certificate complexity, and block sensitivity. The value of such a measure is at most the number of variables. There is a long line of research aimed at bounding one such measure in terms of another. For measures a and b let us write $a \leq_r b$ if there are constants C_1, C_2 such that for every total boolean function f, $a(f) \leq C_1 b(f)^r + C_2$. For example, the decision tree complexity of f, $D(f)$, is at least its degree $deg(f)$ and thus $deg \leq_1 D$. It is also known [Mid04] that $D \leq_3 deg$. We say that a is *polynomially bounded* by b if $a \leq_r b$ for some $r > 0$ and that a and b are *polynomially equivalent* if each is polynomially bounded by the other.

The measures mentioned above, with the notable exception of sensitivity, are known to be polynomially equivalent. For example, in relating block sensitivity, $bs(f)$, to

degree Nisan and Szegedy [NS94] show that $bs(f) \leq_2 deg(f)$. In the other direction, the bound $deg(f) \leq_3 bs(f)$ follows from a result in [BBC+01]. For a survey on many of these results, see [BdW02]. The sensitivity conjecture asserts that $s(f)$ is polynomially equivalent to all of the measures mentioned in this section, and for this, it suffices to show that it is polynomially related to $deg(f)$.

There are a number of equivalent formulations of the sensitivity conjecture. For instance [GL92] give a graph theoretic formulation by exploring a different relationship between sensitivity and degree than what is presented here. The same graph theoretic question also appeared somewhat earlier in [CFGS88], however, sensitivity of boolean functions was only mentioned as a related problem and no direct connection was given. For a good survey of many other variations of the sensitivity conjecture, see [HKP11].

The sensitivity conjecture perhaps more commonly appears as a question on the relationship between sensitivity and block sensitivity. For example, Nisan and Szegedy [NS94] asked specifically if $bs(f) = O(s^2(f))$ for all functions, and as of this writing no counterexample has been given. The best known bound relating sensitivity to another measure was given by Kenyon and Kutin [KK04]. They proved that $bs(f) \leq \frac{e}{2\pi} e^{s(f)} \sqrt{s(f)}$ for all boolean functions.

5.1.4 Outline of the Chapter

In Section 5.2 we prove that a positive answer to Question 5.1 would imply the sensitivity conjecture. We also describe how protocols relate adverserial methods for proving that boolean functions are evasive (that is have decision tree complexity $D(f) = n$). At the end of the section we prove that it suffices to answer Question 5.1 for a special subset of protocols called *order oblivious protocols*.

In Section 5.3 we present three stronger variants of Question 5.1. We then show that for two of these variants, there are protocols that give negative answers to the questions, and suggest that Question 5.1 has a negative answer as well. However, these protocols satisfy a property called monotonicity and in Section 5.4 we prove an $\Omega(n^{1/2})$ lower bound on the cost of any monotone protocol, which shows that any protocol that gives a negative answer to Question 5.1, must look quite different from the two protocols that

refuted the strengthenings. In the same section we prove a rather weak lower bound for a special class of protocols called assignment oblivious protocols. Finally, in Section 5.5 we give the construction of the lowest cost protocol that we know, whose cost is lower than that of the AND-OR protocol by a constant factor.

5.2 Connection between the Sensitivity Conjecture and the Game

In this section we prove Proposition 5.3, which connects the sensitivity conjecture with the two player game described in the introduction.

We will use $\mathbf{e}_\ell$ to denote the assignment in $\{0,1\}^n$ that is 1 in location ℓ and 0 elsewhere. Given two assignments $\mathbf{v}, \mathbf{w} \in \{0,1\}^n$ we will use $\mathbf{v} \oplus \mathbf{w}$ to denote the assignment for which each coordinate is the mod-2 sum of the corresponding coordinates in $\mathbf{v}$ and $\mathbf{w}$.

Recall that Alice's strategy gives the mapping from the input permutation σ and bit b to a boolean vector $\mathbf{v}$ and Bob's strategy maps the vector $\mathbf{v}$ to a subset of locations in $\mathbf{v}$. We first observe that for each strategy for Alice there is a canonical best strategy for Bob. For a permutation σ, we let $\Pi_A(\sigma)$ denote the vector Alice writes down after receiving $\sigma_1, \cdots, \sigma_{n-1}$ (so the location σ_n is still labeled with a $*$). Thus $\Pi_A(\sigma)$ can be viewed as an edge in the *hamming graph* $\mathbb{H}_n$ whose vertex set is $\{0,1\}^n$, with two vertices adjacent if they differ in one coordinate. The *edge set* $E(\Pi)$ of a protocol Π is the set of edges $\Pi_A(\sigma)$ over all permutations σ. This defines a subgraph of $\mathbb{H}_n$. Given Alice's output $\mathbf{v}$, the possible values for σ_n are precisely those locations ℓ that satisfy $(\mathbf{v}, \mathbf{v} \oplus \mathbf{e}_\ell)$ is an edge in $E(\Pi)$. Thus the best strategy for Bob is to output this set of locations. It follows that $c(\Pi)$ is equal to the maximum vertex degree of the graph $E(\Pi)$.

Proposition 5.3 will therefore follow by showing the following: Given a boolean function with degree n and sensitivity s, there is a strategy Π for Alice for the game G_n such that the graph $E(\Pi)$ has maximum degree at most s.

We need a few preliminaries. A *subfunction* of a boolean function f is a function g obtained from f by fixing some of the variables of f to 0 or 1. Note it is clear that if

g is a subfunction of f then $s(f) \geq s(g)$. We say a function has *full degree* if $deg(f)$ is equal to the number of variables of f. We start by recalling some well known facts.

Lemma 5.4. *For any boolean function f there exists a subfunction g on $deg(f)$ variables that has full degree.*

Proof. If p is the (unique) multilinear real polynomial that agrees with f on the boolean cube, then p contains a monomial $\prod_{\ell \in S} x_\ell$ where $|S| = deg(f)$. Let g be the function obtained by fixing the variables in $[n] \setminus S$ to 0. Then g is a function on $deg(f)$ variables that has full degree. □

Lemma 5.5. *Given a function f with full degree and a location ℓ, there exists a bit b such that the function obtained from f by fixing $x_\ell = b$ is also of full degree.*

Proof. The polynomial (viewed as a function from $\{0,1\}^n \to \{0,1\}$) for f may be written in the form $p_1(x_1, x_2, \cdots, \cancel{x_\ell}, \cdots, x_n) + x_\ell p_2(x_1, x_2, \cdots, \cancel{x_\ell}, \cdots, x_n)$. Here the cross-through notation indicates that the variable x_ℓ is not an input to the polynomial. If p_1 has a non zero coefficient on the monomial $\prod_{k \neq \ell} x_k$, then we set $x_\ell = 0$ and the resulting function will have full degree. For the other case, note p_2 must have a non zero coefficient on $\prod_{k \neq \ell} x_k$ because f has full degree. Thus, setting $x_\ell = 1$ will work. □

We remark that the argument in the above lemma is essentially the same as the standard argument that the decision tree complexity of any function f is at least $deg(f)$.

We are now ready to prove Proposition 5.3.

Proof. Given the function f, let g be a subfunction on $deg(f)$ variables with full degree. We will construct a protocol Π that satisfies $E(\Pi) \subseteq E(g)$, where $E(g)$ denotes the set of sensitive edges for the function g, i.e. the edges of $\mathbb{H}_n$ whose endpoints are mapped to different values by g. This will imply that $c(\Pi) \leq s(g) \leq s(f)$, and thus prove the proposition. As Alice receives $\sigma_1, \sigma_2, \cdots, \sigma_n$, she fills in $\mathbf{v}$ in such a way so that the function f restricted to the partial assignment written on $\mathbf{v}$ remains a full degree function, which is possible by Lemma 5.5.

Note that after Alice writes a bit in location σ_{n-1}, the function g restricted to $\mathbf{v}$ is now a non-constant function of one variable, and thus the edge $\Pi_A(\sigma)$ is a sensitive edge for the function g. This implies that $E(\Pi) \subseteq E(g)$.

□

Remark: To summarize, the reduction above shows that a degree n Boolean function having sensitivity s can be converted into a strategy for Alice for the game G_n of cost at most s. We don't know whether this connection goes the other way, i.e., we can't rule out the possibility that the answer to Question 5.1 is negative (there is a very low cost protocol for G_n) but the sensitivity conjecture is still true.

5.2.1 Connection to Decision Tree Complexity

We note the connection between protocols Π for the game G_n and boolean functions on n variables for which $D(f) = n$ (sometimes referred to as *evasive* functions). A common method for showing that a function is evasive is to use an *adversary argument*. For example, consider the evasive function

$$\text{AND-OR}(\mathbf{x}) = \bigwedge_{i=1}^{\sqrt{n}} \bigvee_{j=1}^{\sqrt{n}} x_{ij}.$$

To show this function is evasive we simulate the computation of some decision tree on an input $\mathbf{x}$, except when the tree queries a variable x_{ij} the adversary will respond either 0 or 1 in such a way as to keep the value of the function on the input $\mathbf{x}$ unknown until all variables are queried. For the AND-OR function, take the adversary that always answers 0 as long as some other variable in the corresponding OR block remains undetermined, otherwise it answers 1. This adversary is exactly Alice's part of the AND-OR protocol described in the introduction. For more examples of adversary arguments see [LY02].

Every evasive function by definition admits an adversary argument which in turn defines a protocol Π. In fact a function f is evasive if and only if there exists a protocol Π for which $E(\Pi) \subseteq E(f)$ (recall $E(f)$ is the set of sensitive edges of the function f). This work explores the question, can we use the inherent structure of an arbitrary adversary

(or protocol) to exhibit a lower bound on sensitivity? We provide some limited evidence that this may be possible by proving lower bounds for restricted classes of protocols Π (see Section 5.4).

5.2.2 Order Oblivious Protocols

In the game G_n, at each step $i < n$, the value written by Alice at location σ_i may depend on her knowledge up to that step, which includes both the sequence $\sigma_1, \cdots, \sigma_i$ and the partial assignment already made to v at locations $\sigma_1, \ldots, \sigma_{i-1}$. A natural way to restrict Alice's strategy is to require that the bit she writes in location σ_i depend only on σ_i and the current partial assignment to v but not on the order in which $\sigma_1, \ldots, \sigma_{i-1}$ arrived. A protocol satisfying this restriction is said to be *order oblivious*. The following easy proposition shows that it suffices to answer Question 5.1 for order oblivious protocols.

Proposition 5.6. *Given any protocol* Π *there exists an order oblivious protocol* Π' *such that* $E(\Pi') \subseteq E(\Pi)$. *In particular,* $c(\Pi') \leq c(\Pi)$.

Proof. First some notation. Given a permutation σ let $\sigma_{\leq k}$ denote the prefix of the first k elements of σ. We let $\Pi_A(\sigma_{\leq k})$ denote the partial assignment written on $\mathbf{v}$ after Alice has been streamed $\sigma_1, \cdots, \sigma_k$.

We give a canonical way of obtaining an order oblivious protocol Π' from Π. We define Π' in steps, where step k refers to what Alice does when she is streamed σ_k. For step 1, when σ_1 arrives, she writes according to what Π does for that value of σ_1. In order to define step $k+1$, assume Π' is defined for the first k steps. Assume as well that it satisfies for every permutation σ, there is a permutation τ of $\sigma_1, \cdots, \sigma_k$ so that $\Pi_A(\tau) = \Pi'_A(\sigma_{\leq k})$.

Suppose σ_{k+1} arrives and the current state of the vector is $\mathbf{v} \stackrel{\text{def}}{=} \Pi'(\sigma_{\leq k})$. Note from $\mathbf{v}$ Alice can deduce the set of the first k elements of σ (it is the set of locations not labeled with a *). Alice then considers all permutations τ of $\sigma_1, \cdots, \sigma_k$ such that $\Pi_A(\tau) = \Pi'_A(\sigma_{\leq k})$ and picks the lexicographically smallest permutation (call it τ^*) in that set and writes on location σ_{k+1} according to what Π does after τ^*. Note that the

bit written on location σ_{k+1} does not depend on the relative order of $\sigma_1, \sigma_2, \cdots, \sigma_k$. Using this strategy, Alice maintains the invariant that for every permutation σ, there is a permutation τ of $\sigma_1, \cdots, \sigma_k$ so that $\Pi(\tau) = \Pi'(\sigma_{\leq k})$.

Thus, by construction, Π' is assignment oblivious. Also for any permutation σ there is a permutation τ for which $\Pi_A(\tau) = \Pi'_A(\sigma)$. This implies that $E(\Pi') \subseteq E(\Pi)$. □

5.3 Stronger Variants of Question 1

In this section we propose three natural variants of Question 5.1, and refute two of these variants by exhibiting and analyzing some specific protocols.

The cost function $c(\Pi)$ of a protocol is defined based on the worst case over all choices of $\sigma_1, \ldots, \sigma_n, b$. Alternatively, it is natural to evaluate a protocol based on the average size of the set Bob outputs, where the average is taken over a random permutation $\sigma_1, \ldots, \sigma_n$ and a random bit b. We call this the *expected cost of* Π and denote it by $\tilde{c}(\Pi)$. Let $\tilde{C}(n)$ denote the minimum expected cost of a protocol for G_n.

Question 5.7. *Is there a* $\delta > 0$ *such that* $\tilde{C}(n) = \Omega(n^\delta)$*?*

An affirmative answer to this question would give an affirmative answer to Question 5.1.

We point out that it is well known that the natural probabilistic version of the sensitivity conjecture, where sensitivity is replaced by average sensitivity (where the average is taken uniformly over $\{0,1\}^n$) is trivially false (for example, for the OR function). For contrast, consider the protocol Π where Alice writes a 0 at each step. This protocol is closely related to the OR function in that Alice's part of this protocol is exactly the adversary argument used to prove that OR is evasive. Note also that $E(\Pi)$ is exactly the set of sensitive edges for the OR function. However, the average cost $\tilde{c}(\Pi)$ is $n/2$ whereas the average sensitivity of the OR function is $o(1)$. We currently know of no protocol Π for which $\tilde{c}(\Pi) = o(\sqrt{n})$.

We also remark that an analog of Proposition 5.6 holds for the cost function $\tilde{c}(\Pi)$, and therefore it suffices to answer the question for order oblivious protocols. (The proof of the analog is similar to the proof of Proposition 5.6, except when modifying

the protocol τ^* is not selected to be the lexicographically smallest permutation in the indicated set, but rather the permutation in the indicated set that minimizes the expected cost conditioned on the first k steps.)

There is another natural variant of Question 5.1 based on average case. When we run a fixed protocol Π on a random permutation σ and bit b, we can view the vector $\mathbf{v}$ produced by Alice as a random variable. Let $\tilde{h}(\Pi)$ be the conditional entropy of σ_n given $\mathbf{v}$; intuitively this measures the average number of bits of uncertainty that Bob has about σ_n after seeing $\mathbf{v}$. It is easy to show that this is bounded above by $\log(c(\Pi))$. Let $\tilde{H}(n)$ be the minimum of $\tilde{h}(\Pi)$ over all protocols Π for G_n. The analog of Question 5.1 in this setting asks whether there is a positive constant δ such that $\tilde{H}(n) = \Omega(\delta \log(n))$? An affirmative answer to this would imply an affirmative answer to Question 5.1, however it turns out that the answer to this new question is negative. The following construction is due to Michal Kocky.

Theorem 5.8. *There is an order oblivious protocol* Π *for* G_n *such that* $\tilde{h}(\Pi) = O(\log\log(n))$.

Remark: Earlier we showed one can transform any protocol into an order oblivious protocol with smaller cost. However, it is not clear whether or not this transformation can increase $\tilde{h}$. Instead, we directly provide an example of an order oblivious protocol for which $\tilde{h}(\Pi)$ is small.

Proof. Before defining the protocol Π we need some setup. Let $k = \lceil \log(n) \rceil$ and associate each integer $\ell \in [n]$ to its binary expansion, viewed as a vector $\mathbf{b}(\ell) \in \mathbb{F}_2^k$. Note that $0 \notin [n]$, and thus each vector $\mathbf{b}(\ell)$ is nonzero. Let $t > k$ be an integer (which we'll choose to be $\log^2(n)$) and for each $S \subseteq [n]$ of size t, let $Z(S)$ be a maximal subset of S such that $\sum_{\ell \in Z(S)} \mathbf{b}(\ell)$ is the 0 vector. Observe that by maximality, $Z(S) \geq |S| - k$ (otherwise $S \setminus Z(S)$ would have a linearly dependent subset which we could add to $Z(S)$). Finally let $\mathcal{H} = \{Z(S) : S \in \binom{[n]}{t}\}$.

Given $T \in \mathcal{H}$ and a partial assignment π, we say T is *compatible* with π if $\pi_i \in \{1, *\}$ for all $i \in T$. The protocol Π is defined as follows. For $i \neq n$ Alice writes a 0 on location σ_i unless doing so makes all $T \in \mathcal{H}$ not compatible with the resulting partial assignment written on $\mathbf{v}$, otherwise she writes a 1.

There is a simpler version of this protocol where Alice writes 1 on location i if and only if $i \in Z(S)$ where S is the set of the last t locations of σ. The cost of this protocol is easier to analyze, but it is not order oblivious. Here we instead analyze the order oblivious protocol you obtain if Alice writes a 0 as long as she remains consistent with some partial assignment in the order sensitive protocol.

We note two properties of Π. First, Alice will write a 0 on the first $n - t$ streamed locations. To see this, let $S(\sigma)$ denote the set of the last t elements of σ. Then $Z(S(\sigma))$ will be compatible with $\mathbf{v}$ for the first $n - t$ steps. We also have:

Claim 5.9. *There is a unique set $F \in \mathcal{H}$ that is compatible with the partial assignment $\Pi_A(\sigma)$.*

Proof. Recall that $\Pi_A(\sigma)$ will have a $*$ in location σ_n. Suppose that there are two sets F_1, F_2 that are compatible with $\Pi_A(\sigma)$ and let T be their symmetric difference. First suppose $T - \{\sigma_n\}$ is non-empty and pick $i \in T - \{\sigma_n\}$. Then when location i arrived, Alice could have written a 0 since one of F_1 or F_2 would remain compatible. This contradicts the construction of the protocol. Now suppose that $T = \{\sigma_n\}$. In this case, since $\sum_{\ell \in F_1} \mathbf{b}(\ell) = \sum_{\ell \in F_2} \mathbf{b}(\ell) = \vec{0}$, the vector $\mathbf{b}(\sigma_n)$ must be the zero vector. This is also impossible because we defined the protocol to have all $\mathbf{b}(\ell)$ non-zero. □

We will refer to the set promised by Claim 5.9 as the *final set* and denote it as $F(\sigma)$.

We now obtain an upper bound on the conditional entropy of σ_n given $\mathbf{v}$. Let L be the random variable that is 1 if $\sigma_n \in F(\sigma)$ and 0 otherwise. We have:

$$\begin{aligned} H(\sigma_n|\mathbf{v}) &\leq H(\sigma_n, L|\mathbf{v}) \\ &= H(L|\mathbf{v}) + H(\sigma_n|\mathbf{v}, L) \\ &\leq 1 + H(\sigma_n|\mathbf{v}, L) \\ &= 1 + H(\sigma_n|\mathbf{v}, L = 1)\mathbb{P}L = 1 \\ &\quad + H(\sigma_n|\mathbf{v}, L = 0)\mathbb{P}L = 0 \end{aligned}$$

We first bound the second term. Note that given $L = 1$ we have that σ_n is in the final set $F(\sigma)$ and that Bob can deduce $F(\sigma)$ given the vector $\mathbf{v}$. To see this, let W

be the set of locations ℓ for which $\mathbf{v}$ is set to 1 and let $\Gamma = \sum_{\ell \in W} \mathbf{b}(\ell)$. If Γ is $\vec{0}$, then $F(\sigma)$ must be the set of locations that are set to 1. Otherwise Γ will be equal to $\mathbf{b}(\ell^*)$ for some unique ℓ^*, and $F(\sigma)$ is then the set of locations set to 1 union ℓ^*. In either case, the number of possible values for σ_n is no more than t and so the second term is at most $H(\sigma_n|\mathbf{v}, L = 1) \leq \log(t)$.

To bound the third term we first show the following:

Claim 5.10. *The probability that $L = 0$ is at most k/t.*

Remark: This claim is very easy to see for the order sensitive version mentioned earlier ($L = 0$ is exactly the event that $\sigma_n \in S - Z(S)$). The fact that it still works for the order oblivious version seems quite intuitive because Alice writing some additional 0's should only help the probability. For completeness, we provide a rigorous proof of this below.

Proof. Recall that $L = 0$ means that $\sigma_n \in S \setminus F(\sigma)$. As before let $\sigma_{\leq j}$ denote the prefix of the first j elements of σ and let $T(\sigma_{\leq j})$ denote the set of the first j elements of σ. Given a prefix τ of length $n - l$ we let $M(\tau)$ denote $\max_E |T(\tau) - E|$ where the max is over all sets E that are compatible with $\Pi_A(\tau)$. For integers l and m let $f(l, m)$ denote $\min_\tau (\mathbb{P}L = 0|\sigma_{\leq n-t} = \tau)$ where the minimum is over all prefixes τ of length $n - \ell$ for which $M(\tau) = m$. We will show that $f(\ell, m) \leq m/\ell$ for all ℓ, m. In particular, since every $Z(S)$ has size at least $t - k$, showing that $f(t, k) \leq k/t$ will prove the claim. We proceed by induction on $\ell + m$. As a base case, it is easy to see that if $m = 0$ the probability is 0, and if $\ell = m$ then the probability is 1.

Let τ be any prefix of length $n-\ell$ for which $M(\tau) = m$ and suppose that $\sigma_{\leq(n-\ell)} = \tau$. Note that if Alice writes a 0 next, then $M(\sigma_{\leq(n-\ell+1)}) \leq M(\sigma_{\leq(n-\ell)}) - 1$. Also if Alice writes a 1 next, then $M(\sigma_{\leq(n-\ell+1)}) = M(\sigma_{\leq(n-\ell)})$. Let p denote the probability that Alice will write a 0 on location $\sigma_{n-\ell+1}$. Then $p \geq m/\ell$ (if there is exactly one set T that is compatible then $p = m/\ell$ and with additional sets the probability only increases).

Thus

$$\begin{aligned} f(\ell, m) &\leq \mathbb{P}L = 0 | \sigma_{\leq n-\ell} = \tau \\ &\leq pf(\ell-1, m) + (1-p)f(\ell-1, m-1). \\ &\leq \frac{m}{\ell}\frac{m-1}{\ell-1} + \frac{\ell-m}{\ell}\frac{m}{\ell-1} \qquad \text{(by the I.H.)} \\ &= m/\ell \end{aligned}$$

□

Note that trivially $H(\sigma_n|\mathbf{v}, L=0) \leq \log(n)$, thus the claim implies that the third term is at most $\log(n) \cdot \frac{k}{t}$. By choosing $t = \log^2(n)$ the second term is $O(\log\log(n))$ and the third term is $O(1)$. □

For our last variant, suppose Alice can communicate to Bob with a ternary alphabet instead of a binary alphabet. We will show that Question 5.1 is false in this setting. The setup is the same as before: Alice is streamed a permutation σ, only when σ_i arrives she may write a 0,1, or 2 on location σ_i in $\mathbf{v}$. When $b \in \{0,1,2\}$ arrives she is forced to write b at location σ_n. Bob sees $\mathbf{v}$ and has to output a set J which must contain σ_n. The cost is the maximum size of J for any σ and b. The following construction is due to Michal Koucky and Michael Saks.

Theorem 5.11. *There is a protocol* Π *using a ternary alphabet that has cost* $O(\log(n))$.

Proof. Let $t < n$ be a parameter to be chosen later (we will end up showing that the cost is less than t).

Alice begins by writing 0 on the first $n-t$ locations streamed to her. After this, Alice writes only 1's and 2's (as described below). Clearly if the final input b is not 0, Bob will see exactly t locations that are not labeled a 0 and know the last t elements. Consider then the case that $b = 0$. We'll show that Alice can write the 1's and 2's in such a way that Bob can then determine σ_n exactly. In what follows, a binary string will refer to a string of 1's and 2's.

Consider the graph defined on t element sets where two sets are joined if they have symmetric difference 2. The degree of this graph is trivially less than n^2 so it has a

proper coloring with at most n^2 colors.

Now let us encode each of these colors by a binary string of length t. Write $E(c)$ for the encoding of color c. We want our encoding to have the following property: for any two colors c, d if you delete any single bit from the encoding of $E(c)$ (which leaves a $t-1$ bit string) and delete any single bit from the encoding of $E(d)$ then they are still different.

Claim 5.12. *There is such an encoding for $t = 5\log(n)$.*

Proof. Consider the graph defined on binary strings of length t, where two strings s_1, s_2 are joined if there is a way of deleting a symbol from s_1 and a symbol from s_2 to arrive at the same string of length $t-1$. The degree of this graph is trivially less than $2t^2$, thus there is a proper coloring with at most $2t^2$ colors. Thus there is a color class of size at least $\frac{2^t}{2t^2}$ strings. If $t > 5\log(n)$ then there is a color class of size at least n^2. Picking n^2 strings in this color class will give us the desired encoding $E(c)$. □

After Alice writes the first $(n-t)$ 0's, she knows the final t positions denoted $j_1 < \ldots < j_t$. She determines the color c of that set and the encoding $E(c)$. She then writes the bits of $E(c)$ in the positions $j_1, \ldots, j_t$ (writing the bits in this order and not in the σ order of the last t elements).

If $b = 0$, Bob only sees $t-1$ of the bits. However, by the property of the encoding, this is enough to recover $E(c)$ and therefore c. Furthermore, knowing c and $t-1$ out of the last t elements, the property of the coloring allows Bob to recover the missing element, which is σ_n. This concludes the construction. □

5.4 Lower Bounds for Restricted Protocols

In the previous section we formulated two stronger variants of Question 5.1 that turned out to be false. This may suggest that the original question is also false. In this section however, we will prove a lower bound which implies that any counterexample to Question 5.1 will need to look quite different from the two protocols provided in the last section.

An order oblivious protocol can be specified by a sequence of maps $A_1, \cdots, A_n$ where each A_i maps partial assignments on the set $[n]$ to a single bit. When location σ_i arrives, the bit Alice writes is $A_{\sigma_i}(\mathbf{v})$. For partial assignments α and β, we say that β is an *extension* of α, denoted as $\beta \geq \alpha$, if β is obtained by starting from α and possibly fixing more variables. An order oblivious protocol is *monotone* if each of the maps $A_1, \cdots, A_n$ are monotone with respect to the extension partial order. That is, if $\beta \geq \alpha$ are partial assignments, then $A_i(\beta) \geq A_i(\alpha)$ for each i. As a remark, when running the protocol there may be assignments that are never written on $\mathbf{v}$, however defining each A_i to have domain all partial assignments is still valid and simplifies notation.

Both the AND-OR protocol described in the introduction and the protocol constructed in Theorem 5.8 are examples of monotone protocols. This definition easily generalizes to protocols on alphabets of size k, in which case the ternary protocol given in the previous section can be seen to be monotone. Our main result in this section is that monotone protocols on binary alphabets have cost $\Omega(\sqrt{n})$. In particular, Question 5.1 is true for such protocols. For the rest of the chapter, all protocols will be on binary alphabets.

Before proving the theorem we'll need some new definitions. Recall that an edge $e \in \mathbb{H}_n$ may be written as a vector in $\{0, 1, *\}^n$ for which $e_\ell = *$ on exactly one location ℓ. We call this location ℓ the *free location* of that edge. We say two edges e, e' *collide* if $e_\ell = e'_\ell$ for all ℓ that is not a free location of either edge. Equivalently, two edges collide if they share at least one vertex (each edge collides with itself). Both of the lower bounds in this section will follow by finding an edge $e \in E(\Pi)$ that collides with m other edges in $E(\Pi)$. This implies at least one of the vertices in e has degree at least $m/2$ in the graph $E(\Pi)$, which in turn lower bounds the cost of the protocol.

Finally, given a permutation σ we will use $\ell <_\sigma k$ to denote that the element ℓ comes before the element k in σ.

Theorem 5.13. *All monotone protocols have cost* $\Omega(\sqrt{n})$.

Proof. Let Π be a monotone protocol.

For a permutation σ denote by $\text{bump}_k(\sigma)$ the permutation obtained from σ by

"bumping" the element k to the end of σ and maintaining the same relative order for the rest of σ. For example, $\text{bump}_1(321654) = 326541$.

We let $w(\sigma)$ denote the vector $\Pi_A(\sigma)$ with the entries sorted in σ order. In other words, $w(\sigma)$ is the vector defined by $w(\sigma)_i = (\Pi_A)_{\sigma_i}$. Our proof follows by repeated application of the following:

Claim 5.14. *Let σ be any permutation and let τ be obtained from σ by performing some sequence of bumps on σ. Suppose that τ and $m < n$ satisfies the following:*

- *The elements $\tau_1, \tau_2, \cdots, \tau_m$ were never bumped.*
- *Alice originally wrote a 0 on the locations $\tau_1, \cdots, \tau_m$, that is $\Pi_A(\sigma)_{\tau_i} = 0$ for all $i \leq m$.*

Then $\Pi_A(\tau)_{\tau_i} = 0$ for all $i \leq m$. Equivalently, $w(\tau)$ begins with m 0's.

Proof. The claim follows easily by induction on i. Suppose we have already shown that $w(\tau)$ begins with $(i-1)$ 0's. Let $\mathbf{v}(\sigma, k)$ denote the partial assignment written on $\mathbf{v}$ just before Alice receives the index k (here the reader should take care to distinguish this from the partial assignment just before Alice receives σ_k). Consider the partial assignment $\mathbf{v}(\tau, \tau_i)$. It follows from the first assumption and the inductive hypothesis that $\mathbf{v}(\sigma, \tau_i)$ is an extension of $\mathbf{v}(\tau, \tau_i)$. Thus, since Alice originally wrote a 0 on location τ_i, by monotonicity she continues to write a 0 on that location when being streamed τ (that is $\Pi_A(\tau)_{\tau_i} = 0$). □

Let σ be the permutation for which $w(\sigma)$ is lexicographically smallest.

Claim 5.15. *$w(\sigma)$ consists of a string of 0's followed by a string of 1's, followed by a single *.*

Proof. Suppose for contradiction that there is a a 0 that comes after a 1, and let k be the least index such that $w(\sigma)_k = 1$ and $w(\sigma)_{k+1} = 0$. Let τ be obtained from σ by bumping all of the locations ℓ for which $\ell <_\sigma k$ and $\Pi_A(\sigma)_\ell = 1$. Let m denote the number of locations ℓ for which $\ell <_\sigma k$ and $\Pi_A(\sigma)_\ell = 0$. Then by Claim 5.14, $w(\tau)$ begins with $(m+1)$ 0's. This contradicts the choice of σ □

Let $n-t$ be the number of initial 0's in $w(\sigma)$ and $t-1$ be the number of 1's. For k between 1 and n, let $\tau^{(k)} = \text{bump}_k(\sigma)$. Let x be the assignment obtained from $\Pi_A(\sigma)$ by setting location σ_n (which is a *) to 1.

Claim 5.16. *The edges $\Pi_A(\tau^{(k)})$ and $\Pi_A(\sigma)$ intersect at the input x for all k among the last t elements of σ. In particular x has degree at least t in the graph $E(\Pi)$.*

Proof. Fix k among the last t elements of σ. Clearly $w(\tau^{(k)})$ has the first $n-t$ bits 0, and so by the choice of σ all other locations in $w(\tau^{(k)})$ must be labeled 1. Thus $w(\tau^{(k)}) = w(\sigma)$. This means that the edges $\Pi_A(\sigma)$ and $\Pi_A(\tau^{(k)})$ agree at all locations except for σ_n and σ_k (which are the free location of the edges respectively). Since $\Pi_A(\sigma)_{\sigma_k} = \Pi_A(\tau^{(k)})_{\sigma_n} = 1$, the two edges meet at x. □

To conclude the proof of the theorem we will find an assignment y that has degree at least $(n-t)/(t+1)$ in the graph $E(\Pi)$.

Claim 5.17. *For k among the first $n-t$ elements of σ, $w(\tau^{(k)})$ has the first $n-t-1$ bits equal to 0, and has at most one 0 among the next t bits (and last bit *).*

Proof. The fact that the first $n-t-1$ bits of $w(\tau^{(k)})$ are labeled 0 follows by directly by Claim 5.14.

Suppose for contradiction that there are at least 2 0's among the next t locations and denote the locations of the first and second 0 to be ℓ_1 and ℓ_2 respectively. Take all of the locations that are labeled 1 in $\Pi_A(\tau^{(k)})$ and bump them to the end and let this new permutation be ρ. Once again by applying Claim 5.14 we have $\Pi_A(\rho)_{\ell_1} = \Pi_A(\rho)_{\ell_2} = 0$. Thus $w(\rho)$ has the first $n-t+1$ locations set to 0 which contradicts the choice of σ. □

Now classify each of the first $n-t$ elements of σ into $t+1$ types $n-t, \ldots, n$. Element k is of type n if $w(\tau^{(k)})$ has t 1's. Otherwise $w(\tau^{(k)})$ has $(t-1)$ 1's, and the type of k is equal to the index j between $n-t$ and $n-1$ such that $w(\tau^{(k)})_j = 0$.

Some type occurs at least $m \stackrel{\text{def}}{=} (n-t)/(t+1)$ times, call it j^*, and let $k_1, k_2, \cdots, k_m$ be the m elements that are type j^*. For $1 \le i \le m$ let $y^{(i)}$ be the assignment obtained by taking the edge $\Pi_A(\tau^{(k_i)})$ and assigning the $*$ to 0.

Claim 5.18. *The assignments $y^{(i)}$ are all equal.*

Proof. By the definition of the bump operation the permutations $\tau^{(k_i)}$ all have the same elements at positions $n-t, n-t+1, \cdots, n-1$ (they have the same suffix with the exception of the last element). Since they are all of the same type it follows that the $y^{(i)}$ all agree on locations in the set $\{\tau^{(k_1)}(j) \mid j \in n-t, \cdots, n-1\}$. For all other locations, each $y^{(i)}$ is set to 0, thus they are the same assignment. □

Therefore there are m distinct edges in the graph $E(\Pi)$ that are incident with the assignment $y \stackrel{\text{def}}{=} y^{(1)}$. Thus y has degree at least $m = (n-t)/(t+1)$. This implies that cost of Π is at least $\max(t, (n-t)/(t+1)) = \Omega(\sqrt{n})$. □

As demonstrated by the AND-OR protocol, Theorem 5.13 is tight up to a constant factor. We remark that the monotone protocols we consider here seem to have no general connection to the class of monotone boolean functions, and our result for monotone protocols seems to be unrelated to the easy and well known fact that the sensitivity conjecture is true for monotone functions.

We conclude this section with a lower bound for a second class of protocols. Although the lower bound is only logarithmic, we point out that proving a logarithmic lower bound for all protocols with a strong enough constant would imply new bounds relating degree and sensitivity.

For a permutation σ let $S_k(\sigma)$ denote the set of elements ℓ that satisfy $\ell <_\sigma k$. For example, if $\sigma = 321654$ then $S_1(\sigma) = \{2, 3\}$. We say a protocol is *assignment oblivious* if the bit written by Alice in location k only depends on the set $S_k(\sigma)$. Such protocols can be described by a collection of n hypergraphs $H_1, H_2, \cdots, H_n$, where each H_ℓ is a hypergraph with vertex set $[n] \setminus \{\ell\}$. When k arrives, Alice writes a 1 if and only if the set $S_k(\sigma)$ is in H_k.

Theorem 5.19. *Every assignment oblivious protocol Π has $c(\Pi) \geq \log_2(n)/2$.*

Proof. Let Π be an assignment oblivious protocol.

Given a permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ and $k \in [n]$ we define $\text{swap}_k(\sigma)$ to be the permutation obtained by swapping the positions of the elements k and σ_n within σ and

keeping every other element in the same place. For example, $\text{swap}_3(654321) = 654123$. The lemma will follow by constructing a permutation σ such that that $\Pi_A(\sigma)$ and $\Pi_A(\text{swap}_k(\sigma))$ collide for each $k \in \{\sigma_{n-1}, \cdots, \sigma_{n-\lceil \log_2(n) \rceil}\}$

We build up such a σ in a greedy manner. We start with setting $\sigma_{n-1} = 1$. With σ_{n-1} fixed, the bit Alice writes in location 1 is completely determined by σ_n (and does not depend on the values we later choose for $\sigma_1, \cdots, \sigma_{n-2}$). This holds by the assignment oblivious property and because $S_1(\sigma) = \{\ell : \ell \neq 1, \sigma_n\}$. Let R_1 be the locations ℓ for which setting $\sigma_n = \ell$ results in Alice writing a 1 in location 1. At least one of $|R_1|, |R_1^c|$ are bigger than $\lceil (n-1)/2 \rceil$, let T_1 be that set. Now we fix σ_{n-2} to be any element in T_1.

Having fixed σ_{n-1} and σ_{n-2}, the bit Alice writes on location σ_{n-2} also only depends on the value of σ_n. Now let R_2 be the subset of indices j in T_1 such that setting $\sigma_n = j$ would cause Alice to write a 1 in location σ_{n-2}. At least one of $|R_2|, |R_2^c|$ are bigger than $\lceil (|T_1|-1)/2 \rceil$, let $T_2 \subseteq T_1$ be that set. This process is iteratively repeated. At step i we set σ_{n-i} to be an arbitrary element of T_{i-1}. With $\sigma_{n-1}, \cdots, \sigma_{n-i}$ now fixed, the value written in location σ_{n-i} depends only on the value of σ_n. The set R_i is defined to be all such values of σ_n that result in Alice writing a 1 in location σ_{n-i} and $T_i \subseteq T_{i-1}$ is defined to be the larger of $|R_i|$ and $|R_i^c|$. We proceed until the set T_i has only one element in it, in this case we assign σ_n to be that element. This process will take at least $\lceil \log_2(n) \rceil$ steps. We then assign the remaining elements to $\sigma_1, \cdots, \sigma_{n-i-1}$ in an arbitrary order.

We now claim that $\Pi_A(\sigma)$ and $\Pi_A(\text{swap}_k(\sigma))$ collide for $k = \sigma_n, \sigma_{n-1}, \cdots, \sigma_{n-\lceil \log_2(n) \rceil}$.

Claim 5.20. *Let $i < \lceil \log_2(n) \rceil$, and let $k = \sigma_{n-i}$. Then $\Pi_A(\sigma)_\ell = \Pi_A(swap_k(\sigma))_\ell$ for all $\ell \neq k, \sigma_n$.*

Proof. Let $\sigma' = \text{swap}_k(\sigma)$. If $\ell <_\sigma k$ then $S_\ell(\sigma) = S_\ell(\sigma')$ and so Alice writes the same bit to location ℓ under both permutations.

Suppose that $\ell >_\sigma k$. Let j be such that $\sigma_{n-j} = \ell$. Note that $\sigma_{n-1} = \sigma'_{n-1}, \cdots, \sigma_{n-j} = \sigma'_{n-j}$. Recall that holding $\sigma_{n-1}, \cdots, \sigma_{n-j}$ fixed, the bit Alice writes at location ℓ depends only on the value of σ_n, and furthermore that bit is the same as for all

settings of $\sigma_n \in T_j$. Since both σ_n and $\sigma'_n = k$ are in the set T_j, it follows that $\Pi_A(\sigma)_\ell = \Pi_A(\sigma')_\ell$. □

By the above claim, σ collides with $\text{swap}_k(\sigma)$ for at least $\lceil \log_2(n) \rceil$ values of k. Furthermore, at least one of the vertices in $\Pi_A(\sigma)$ has degree more than $\lceil \log_2(n)/2 \rceil$. This concludes the proof.

□

Two Stage Protocols

Here we include a lower bound for another class of protocols (which does not appear in the journal version of this work). We include it here because the method of proof differs significantly from the lower bound methods used so far, however the result turns out to be implied by the monotone lower bound.

We say a protocol Π is a *two stage protocol* if it satisfies the following: For an integer $t < n$ and each set $S \in \binom{[n]}{t}$ fix an assignment $\mathbf{x}_S \in \{0,1\}^n$ for which $\mathbf{x}_S$ is 0 at the locations in S. Alice writes 0 in the first $n - t$ locations of σ. Alice then defines S to be the remaining t locations, and as the remaining elements of σ arrive, she fills in the locations so that they agree with the assignment $\mathbf{x}_S$ (she is still forced to label σ_n with the bit b).

Theorem 5.21. *For any two stage protocol* Π, $c(\Pi) = \Omega(\sqrt{n})$.

Remark: Two stage protocols are not order oblivious in general, thus one can apply the reduction from Section 5.6 to make a two stage protocol order oblivious. It turns out that after the reduction, the protocol obtained is monotone, and thus this result is implied by our monotone lower bound. However, we include this bound because the method of proof uses a double counting argument that is very different from the monotone lower bound.

Proof. Recall that an edge $e \in \mathbb{H}_n$ may be written as a vector in $\{0,1,*\}^n$ for which $e_\ell = *$ on exactly one location ℓ. We call this location ℓ to be the *free location* of that edge, and denote it as $\phi(e)$.

For $t < n$ fixed, let Π be a two stage protocol which is defined by a set of assignments $\mathbf{x}_S \in \{0,1\}^n$ indexed by the sets $S \in \binom{[n]}{t}$. Recall by definition each $\mathbf{x}_S$ is 0 at the locations outside of S. For such a protocol it holds that $E(\Pi)$ is the union over all $S \in \binom{[n]}{t}$ and $\ell \in S$ of the edges $(\mathbf{x}_S, \mathbf{x}_S \oplus \mathbf{e}_\ell)$. Our goal is to show a lower bound on the maximum degree of such a graph.

Let r be the max over all edges $e \in E(\Pi)$ of the number of edges $e' \in E(\Pi)$ which collide with e. Then $c(\Pi) \geq r/2$ because one of the vertices in e will have degree at least $r/2$. Let $F_t(\sigma)$ be the set of t elements at the end of σ. For each $S \in \binom{[n]}{t}$ and $e \in E(\Pi)$ we say that the edge e is a *descendant* of S if e is of the form $(\mathbf{x}_S, \mathbf{x}_S \oplus \mathbf{e}_\ell)$ for some $\ell \in S$. Equivalently, e is a descendant of S means that $\Pi_A(\sigma) = e$ for every permutation σ for which $F_t(\sigma) = S$ and $\sigma_n = \phi(e)$ (recall that $\Pi_A(\sigma)$ is the state of the vector $\mathbf{v}$ after Alice has been streamed $\sigma_1, \cdots, \sigma_{n-1}$).

For e a descendant of a set S we refer to the pair (e, S) as a *descendant pair*. The proof proceeds by first finding an exact expression for the number of pairs (e, S). We will then upper bound this quantity in terms of r which will then imply our desired lower bound on r.

Claim 5.22. *There are $t\binom{n}{t}$ descendant pairs (e, S).*

Proof. For each set S there are t pairs (e, S); they are the t edges of the form $(\mathbf{x}_S, \mathbf{x}_S \oplus \mathbf{e}_\ell)$ for ℓ in S. Since there are $\binom{n}{t}$ sets S, the claim follows. □

To give an upper bound on the number of such pairs, we count the number of pairs that contain a particular edge e. For this we need to classify the edges e. Define $J(e)$ to be the set of locations that are nonzero for the edge (which includes $\phi(e)$) and define the *weight* of an edge, $w(e)$, to be the size of $J(e)$.

Claim 5.23. *An edge e of weight w belongs to at most $\binom{r-w}{t-w}$ descendant pairs.*

Proof. Let e be a fixed edge of weight w which is a descendant of a set S. By definition of a two stage protocol, Alice always writes 0 on the first $n - t$ indices of σ. It follows that $J(e) \subseteq S$, because $J(e)$ contains all of the locations labeled a 1 as well as the free location of the edge, which is σ_n. If e' is another descendant of S, then both e and

e' are equal to the assignment $\mathbf{x}_S$ except on their respective free locations; thus they collide. As a result, for each $\ell \in S$ there is an edge e' with $\phi(e') = \ell$ which collides with e. Let U be the union of all sets T for which (e, T) is a descendant pair, then $|U| \leq r$. Since for each such T it holds that $J(e) \subseteq T \subseteq U$, there are at most $\binom{r-w}{t-w}$ sets T for which (e, T) is a descendant pair. □

Let M_k denote the set of edges $e \in E(\Pi)$ for which $w(e) = k$.

Claim 5.24. *The size of M_k is at most $r\binom{n}{k-1}$.*

Proof. Let V_k denote the set of vertices in $\mathbb{H}_n$ which have hamming weight $k-1$. For each $v \in V_k$ let d_v be the number of edges of weight k which appear in $E(\Pi)$ and are incident with v. Note each $d_v \leq r$ by the definition of r. Then

$$|M_k| = \sum_{v \in V_k} d_v \leq r|V| = r\binom{n}{k-1}.$$

□

For each edge $e \in E(\Pi)$ let $p(e)$ denote the number of sets T which are paired with e. Then the number of descendant pairs may be counted as

$$\sum_{k=1}^{t} \sum_{e \in M_k} p(e).$$

By applying Claims 5.23 and 5.24 the above is at most

$$\sum_{k=1}^{t} r\binom{n}{k-1}\binom{r-k}{t-k}.$$

Comparing with the count from Claim 5.22 we have the inequality

$$\sum_{k=1}^{t} r\binom{n}{k-1}\binom{r-k}{t-k} \geq \binom{n}{t}t.$$

Let T_k be the k'th term in the above sum, looking at the ratio of consecutive terms we get

$$\frac{T_{k+1}}{T_k} = \frac{n-k+1}{k}\frac{t-k}{r-k}.$$

We may assume $r \leq \sqrt{n}$ (otherwise the theorem holds trivially), also it is clear that $k \leq t \leq s$. Under these conditions the above ratio is greater than 1, thus the biggest term in the sum is when $k = t$. We can then bound the sum above by

$$tr\binom{n}{t-1}\binom{r-t}{0}.$$

And thus we have

$$tr\binom{n}{t-1} \geq \binom{n}{t}t.$$

This implies

$$r \geq \frac{n-t+1}{t}.$$

To conclude the proof, note that $r \geq t$ because all t descendants of a set S collide. Thus $r = \Omega(n^{1/2})$. □

5.5 A Protocol with Lower Cost than the AND-OR Protocol

In this section we present a construction of a protocol with $c(\Pi) \leq \sqrt{\frac{999}{1000}}\sqrt{n}$ which is the lowest cost protocol we know. The construction is a variant of the AND-OR protocol defined in the introduction.

Assume n and k are integers where $n - k$ is a perfect square. A set of assignments $\{\mathbf{x}_S \in \{0,1\}^n | S \in \binom{[n]}{k}\}$ is an (n,k)-*proper code* if the hamming distance between any $\mathbf{x}_S, \mathbf{x}_{S'}$ is at least $2\sqrt{n}$ and each $\mathbf{x}_S$ is 0 on the locations $i \in S$. Let $\{\mathbf{x}_S | S \in \binom{[n]}{k}\}$ be an (n,k)-proper code. We construct a protocol Π as follows: Alice writes 0 at locations $\sigma_1, \cdots, \sigma_k$. Alice then takes the set $S = \{\sigma_1, \cdots, \sigma_k\}$ and splits $[n] \setminus S$ into $\sqrt{n-k}$ disjoint blocks of size $\sqrt{n-k}$. When Alice continues and receives σ_j (for $k < j < n$) she writes the mod-2 sum of the bit b_j and the bit in location σ_j of $\mathbf{x}_S$, where b_j is 1 if σ_j is the last element in its block, and 0 otherwise.

We claim that upon receiving vector $\mathbf{v}$, Bob knows that the value of σ_n is one of $\sqrt{n-k}$ possible locations. First note that the vector $\mathbf{v}$ is within distance $\sqrt{n-k}$ of the vector $\mathbf{x}_S$, and thus Bob may decode $\mathbf{v}$ to learn the assignment $\mathbf{x}_S$ (and thus the set S as well). Consider the assignment $\mathbf{v} \oplus \mathbf{x}_S$ restricted to the locations outside of S. If the final bit b is 0, then exactly one of the $\sqrt{n-k}$ blocks will be all 0's. Bob can

output J to be that block. If the final bit b is 1, then every block will have exactly a single 1 in it. Bob can output J to be the set of locations that are set to 1. In each case $|J| = \sqrt{n-k}$.

To conclude the construction of this protocol we prove the existence of an $(n, n/1000)$-proper code. Consider the following random code indexed by the sets $S \in \binom{[n]}{k}$: Each $\mathbf{x}_S$ is set to 0 on locations in S, and set to an independently and uniformly chosen random bit on locations outside of S. We claim that with nonzero probability this set is a proper code. The second property holds by definition, it only remains to check the pairwise distances of the code words. Given sets S, S' let $E_{S,S'}$ be the event that $d(\mathbf{x}_S, \mathbf{x}_{S'}) < 2\sqrt{n}$. This may be upper bounded by the probability that $\mathbf{x}_S, \mathbf{x}_{S'}$ differ on less than $2\sqrt{n}$ locations in the set $[n] \setminus (S \cup S')$. This probability is exactly the probability that two random $n - |S \cup S'|$ bit strings are within distance $2\sqrt{n}$. Since $n - |S \cup S'| \geq n/2$ this probability is at most $\exp(-n/32)$ by a standard Chernoff bound. By a union bound the probability of any event $E_{S,S'}$ occurring is at most

$$\binom{n}{n/1000}^2 \exp(-n/32) < 1.$$

Thus with nonzero probability this is a proper code.

Corollary 5.25. *There is an $\epsilon > 0$ and a protocol Π for which $c(\Pi) \leq (1-\epsilon)\sqrt{n}$.*

Chapter 6

The Density of Happy Numbers

Acknowledgment of a Journal Publication:

This chapter contains large overlap with the journal version of this work (see [Gil13]).

6.1 Introduction

Consider the map $H : \mathbb{N} \to \mathbb{N}$ which sends a positive integer to the sums of the squares of its digits. We are interested in the trajectory of integers under this map. For example the trajectory of 7 eventually ends up at 1 which maps to itself

$$7 \to 49 \to 97 \to 130 \to 10 \to 1 \to 1 \cdots$$

If instead we start at 4 we have the following

$$4 \to 16 \to 37 \to 58 \to 89 \to 145 \to 42 \to 20 \to 4 \to \cdots .$$

It is not too hard to show that every number eventually ends up at either 1 or the above cycle starting with 4. Notice that the image of a 4 digit number $H(a_1a_2a_3a_4) \leq 4 \cdot 81 < a_1a_2a_3a_4$, likewise for any number x with at least 4 digits, $H(x) < x$. So every number eventually reaches a number less than 1000, and if we check the trajectories of all numbers less than 1000 we see that there are only two possible cycles.

A number is *happy* if it eventually reaches 1. In this chapter we will show that the asymptotic density of happy numbers does not exist. In particular we show that the upper density is at least .18 and the lower density is at most .12. Our methods also apply to many generalizations of this map.

If we change the map, instead sending an integer to the sum of the cubes of its digits, then there are 9 different possible cycles (see Section 6.5.2). Many generalizations of

these kinds of maps have been studied. For instance, [GT07] considered the map which sends an integer n to the sum of the e'th power of its base-b digits. In this chapter, we study a more general class of functions.

Definition 6.1. *Let $b > 1$ be an integer, and let h be a sequence of b non-negative integers such that $h(0) = 0$ and $h(1) = 1$. Define $H : \mathbb{Z}^+ \to \mathbb{Z}^+$ to be the following function: for $n \in \mathbb{Z}^+$, with base-b representation $n = \sum_{i=0}^{k} a_i b^i$, $H(n) := \sum_{i=0}^{k} h(a_i)$. We say H is the b-happy function with digit sequence h.*

As a special case, the b-happy function with digit sequence $\{0, 1, 2^e, \ldots, (b-1)^e\}$ is called the (e, b)-function.

Definition 6.2. *Let H be any b-happy function and let $C \subseteq \mathbb{N}$. We say $n \in \mathbb{N}$ is type-C if there exists k such that $H^k(n) \in C$.*

For example, for the $(2, 10)$-function, happy numbers are type-$\{1\}$. Numbers which are not happy are type-$\{4, 16, 37, 58, 89, 145, 42, 20\}$.

Fix a b-happy function H and let $\alpha := \max_{i=0,\ldots,b-1} \big(H(i)\big)$. If n is a d-digit integer in base-b, then $H(n) \leq \alpha d$. If d^* is the smallest $d \in \mathbb{N}$ such that $\alpha d < b^{d-1}$, then for all n with $d \geq d^*$ digits, $H(n) \leq \alpha d < b^{d-1} \leq n$. This implies the following

Fact 6.3. *For all $n \in \mathbb{N}$, there exists an integer i such that $H^i(n) < b^{d^*} - 1$.*

Moreover, to find all possible cycles for a b-happy function, it suffices to perform a computer search on the trajectories of the integers in the interval $[0, b^{d^*} - 1]$.

Richard Guy asks a number of questions regarding $(2, 10)$-happy numbers and their generalizations, including the existence (or not) of arbitrarily long sequences of consecutive happy numbers and whether or not the asymptotic density exists [Guy04]. To date, there have been a number of papers in the literature addressing the former question ([ESS+00],[GT07],[Pan08]). This work addresses the latter question. Informally, our main result says that if the asymptotic density exists, then the density function must quickly approach this limit.

Theorem 6.4. *Fix a b-happy function H. Let I be a sufficiently large interval and let $S \subseteq I$ be a set of type-C integers. If $\frac{|S|}{|I|} = d$, then the upper density of type-C integers is at least $d\,(1 - o(1))$.*

Note as a corollary we can get an upper bound on the lower density by taking C to be the union of all cycles except the one in which we are interested. In Sections 3 and 4 we will define explicitly what constitutes a sufficiently large interval and provide an expression for the $o(1)$ term. Using Theorem 6.4, one can prove the asymptotic density of (e, b)-happy numbers (or more generally type-$\{C\}$ numbers) does not exist by finding two large intervals I_1, I_2 for which the density in I_1 is large and in I_2 is small. In the case of $(2, 10)$-happy numbers, taking $I_1 = [10^{403}, 10^{404} - 1]$ and $I_2 = [10^{2367}, 10^{2368} - 1]$, we show that the upper density $\bar{d} \geq .185773(1 - 10^{-49})$ and lower density $\underline{d} \leq .11379(1 + 10^{-100})$ respectively.

We also show that the asymptotic density does not exist for 8 of the cycles for the $(3, 10)$-function (see Section 5). It should be noted that our methods only give one sided bounds. In an earlier version of this manuscript, we asked if $\bar{d} < 1$ for $(2, 10)$-happy numbers. Recently, David Moews has announced a proof of this (see his homepage at http://djm.cc/dmoews.html). Specifically, he proves that $.1962 < \bar{d} < .38$, and $0.002937 < \underline{d} < .1217$.

6.2 Preliminaries

Throughout the chapter we regard an interval $I = [a, b]$ as a set of integers $\{n \in \mathbb{Z}^+ : a \leq n \leq b\}$ where, in general, $a, b \in \mathbb{R}$. We denote $|I|$ to be the cardinality of this set. We also denote the set $\{0, 1, \ldots, n\}$ by $[n]$. Throughout this section let H denote an arbitrary b-happy function with digit sequence h.

Definition 6.5. *Let I be an interval and Y the random variable uniformly distributed amongst the set of integers in I. Then we say the random variable Y is induced by the interval I.*

Definition 6.6. *The type-C density of an integer interval I is defined to be the quantity*

$$d_C(I) := \frac{|\{n \in I : n \text{ is type-}C\}|}{|I|}.$$

Observation 6.7. *If Y is the random variable induced by an interval I, then*

$$d_C(I) = \mathbb{P}\big(H(Y) \text{ is type-}C\big).$$

Usually, we take C to be one of the cycles arising from a b-happy function H. However, if we wish to upper bound the lower density of type-C integers, then we study the density of type-C' integers, where C' is the union of all cycles except C.

6.2.1 The Random Variable $H(Y_m)$

Consider the random variable Y_m induced by the interval $[0, b^m - 1]$, i.e., Y_m is a random m-digit number. If X_i is the random variable corresponding to the coefficient of b^i in the base-b expansion of Y_m, then

$$H(Y_m) = \sum_{i=0}^{m-1} h(X_i). \tag{6.1}$$

We will be interested in the mean and variance of the $h(X_i)$ (i.e., the image of a random digit) which we refer to as the *digit mean* (μ) and *digit variance* (σ^2) of H. The random variables $h(X_i)$ in (6.1) are all independent and identically distributed (i.i.d.), thus,

$$\mathbf{E}[H(Y_m)] = \mu m \qquad \mathbf{Var}[H(Y_m)] = \sigma^2 m. \tag{6.2}$$

The random variable $H(Y_m)$ is equivalent to rolling m times a b-sided die with faces $0, 1, h(2), \ldots, h(b-1)$ and taking the sum. Since it is a sum of m i.i.d. random variables, it approaches a normal distribution as m gets large. Also, the distribution of $H(Y_m)$ is concentrated near the mean. This observation leads to the following key insight which underlies the proofs in this work: **For a sufficiently large integer m, the density of type-C integers among m digit integers is approximately determined by the set of type-C integers near μm.**

6.2.2 Computing Densities

In order to apply Theorem 6.4 it is necessary to compute the number of m-digit integers which are type-C for m large. In this section we discuss how this can be done efficiently (even for $m \geq 1000$).

Let $P_{m,i} := \mathbb{P}\big(H(Y_m) = i\big)$. Then

$$P_{m,i} = \frac{\Big|\{(a_1, a_2, \ldots, a_m) : a_k \in h \text{ and } \sum_{k=1}^{m} a_k = i\}\Big|}{b^m}.$$

For fixed m, the sequence $\{P_{m,i}\}_{i=1}^{\infty}$ has generating function

$$f_m(x) = \sum_{i=0}^{\infty} P_{m,i} x^i = \left(\frac{1 + x + x^{h(2)} + \cdots + x^{h(b-1)}}{b}\right)^m. \tag{6.3}$$

This implies the following recurrence relation with initial conditions $P_{0,0} = 1$, and $P_{0,i} = 0$ for $i \in \mathbb{Z} - \{0\}$.

$$P_{m,i} = \frac{P_{m-1,i} + P_{m-1,i-1} + P_{m-1,i-h(2)} + \cdots + P_{m-1,i-h(b-1)}}{b}. \tag{6.4}$$

To see this, write $f_m(x) = \left(\frac{1+x+x^{h(2)}+\cdots+x^{h(b-1)}}{b}\right)^{m-1}\left(\frac{1+x+x^{h(2)}+\cdots+x^{h(b-1)}}{b}\right)$ and consider the coefficient of x^i.

If $\alpha = \max\limits_{i=0,\ldots,b-1}\big(h(i)\big)$, then $H(Y_m) \subseteq [0, m\alpha]$. In particular, $P_{m,i} = 0$ if $i > m\alpha$. Using this fact combined with (6.4), we can implement the following simple algorithm for quickly calculating the type-C density of the interval $[0, b^m - 1]$.

1. First, using the recurrence (6.4), calculate $P_{m,i}$ for $i = 0, \ldots, m\alpha$.

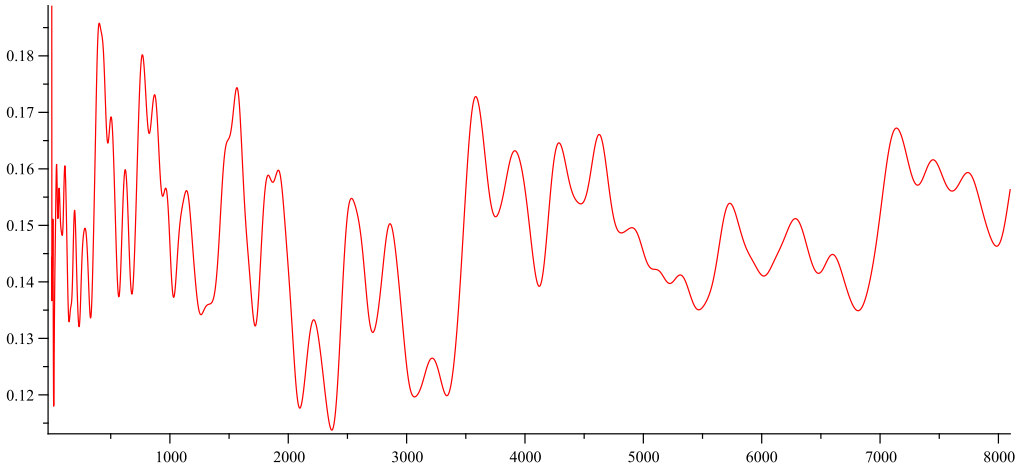
2. Using brute force, find the type-C integers in the interval $[0, m\alpha]$.

3. Output $\sum\limits_{\substack{i \in [0, m\alpha] \\ i \text{ type-C}}} P_{m,i}$.

Using this algorithm, calculating the density for large m becomes computationally feasible. Figure 1 graphs the density of $(2, 10)$-happy numbers $< 10^m$ for m up to 8000.

The peak near 10^{400} and valley near 10^{2350} will be used to imply the bounds obtained in this work.

Figure 6.1: Relative Density of $(2, 10)$-Happy Numbers $< 10^m$



6.2.3 A Local Limit Law

The random variable $H(Y_m)$ approaches a normal distribution as m becomes large. The following theorem[1], presented in [FS09], gives a bound.

Theorem 6.8. *(Local limit law for sums). Let $X_1, \ldots, X_n$ be i.i.d. integer-valued variables with probability generating function (PGF) $B(z)$, mean μ, and variance σ^2, where it is assumed that the X_i are supported on $\mathbb{Z}^+$. Assume that $B(z)$ is analytic in some disc that contains the unit disc in its interior and that $B(z)$ is aperiodic with $B(0) \neq 0$. Then the sum,*

$$S_n := X_1 + X_2 + \cdots + X_n$$

satisfies a local limit law of the Gaussian type: for t in any finite interval, one has

$$\mathbb{P}(S_n = \lfloor \mu n + t\sigma\sqrt{n} \rfloor) = \frac{e^{-t^2/2}}{\sqrt{2\pi n}\sigma} \left(1 + O(n^{-1/2})\right).$$

.

Here aperiodic means that the $\gcd\{j : b_j > 0, j > 0\} = 1$, where $B(z) = \sum_{j=0}^{\infty} b_j z^j$ (or more informally, the digit sequence for H cannot all be divisible by some integer larger than 1). In our case, the PGF of the $H(X_i)$ is the polynomial

$$p(x) = \frac{x^{H(0)} + x^{H(1)} + \cdots + x^{H(b-1)}}{b}.$$

[1]We quote a simpler version, with a minor typo corrected

It is important in our definition of b-happy functions that we assume that $H(0) = 0$, and $H(1) = 1$. This guarantees that $p(x)$ is aperiodic and in particular that the above theorem applies for the sum $H(Y_m)$. As a consequence, for a fixed interval $[-T, T]$, if $i = \mu m + t\sigma\sqrt{m}$ for some $t \in [-T, T]$, then

$$P_{m,i} = \frac{e^{-t^2/2}}{\sqrt{2\pi m}\sigma}\left(1 + O\left(m^{-1/2}\right)\right).$$

The above error term, $O\left(m^{-1/2}\right)$, will prove to be a technical difficulty which will be discussed later.

6.2.4 Overview of the Proof

The following heuristic will provide the general motivation for the proofs. Recall that the random variable Y_m is concentrated near its mean μm.

Primary motivation for the proofs: Suppose I is a large interval with type-C density d. Consider the choices of m such that the mean of $H(Y_m)$ is in the interval I, then for some choices of m we likely have

$$\mathbb{P}\big(H(Y_m) \text{ is type-}C\big) \geq d.$$

Thus if there is an interval I with high type-C density and midpoint a, we should be able to find a much larger interval $I' \approx [0, b^a/\mu]$ (where b is the base we are working in) also with high type-C density. We can apply this reasoning iteratively to conclude the upper density of type-C numbers is large.

The key idea to turn this heuristic into a proof is to average over all reasonable choices of m in order to imply there is an m with the desired property.

We will use Theorem 6.8 to show that, for small k, $H(Y_m)$ and $H(Y_{m+k})$ have essentially the same distribution only shifted by a factor of μk. Thus, as k varies, the distributions $H(Y_{m+k})$ should uniformly cover the interval I. It is crucial here to use the fact that $H(Y_m)$ is approximately locally normal, otherwise the proof will fail. For example, suppose all the happy numbers in I are odd. In this case, if $H(Y_m)$ is not locally normal and instead is supported on the even numbers for all m, then every shift $H(Y_{m+k})$ will miss all of the happy numbers in I.

Unfortunately, the fuzzy term in the local limit law prevents us from obtaining explicit bounds on the error (and any explicit bounds seem unsatisfactory for our purposes). Section 3 adds a necessary step, which is to construct an interval within $[b^{n-1}, b^n]$ with high type-C density for n arbitrarily large. The trick is to consider intervals of the form $I_k := [1^k 0^{n-k}, 1^k 0^m (b-1)^{n-k-m}]$ (here 1^k denotes k consecutive 1's). This solves the issue where $H(Y_m)$ and $H(Y_{m+k})$ are not exact shifts of each other, as the distributions induced by the I_k *are* exact shifts under the image of H. These distributions will uniformly cover the base interval I with much better provable bounds. The main result is presented in Section 4, the proof uses the local limit law with the result from Section 3.

6.3 Constructing Intervals

Throughout this section, if Y is a random variable and k is an integer, let $\tau_k(Y)$ denote the random variable $Y + k$.

Definition 6.9. *We say an integer interval I is n-strict if $I \subseteq [b^{n-1}, b^n - 1]$ and $|I| = b^{3n/4}$.*

The primary goal of this section is to construct n-strict intervals of high type-C density for arbitrarily large n.

Our choice of the definition of n-strict is only for the purpose of simplifying calculations, there is nothing special about the value $\frac{3}{4}$. In fact, any ratio $> \frac{1}{2}$ would work. Note if 4 does not divide n, then no n-strict intervals exist.

For the entirety of this section we will make the following assumptions:

- H is a b-happy function with digit mean μ and digit variance σ^2.
- We wish to lower bound the upper density of type-C integers for some $C \subset \mathbb{N}$.
- We have found, by computer search, an appropriate starting interval I_1, which is n_1-strict and has suitably large type-C density $d_C(I_1)$.

The results in this section apply only if this n_1 is sufficiently large, so we state here exactly how large n_1 must be so one knows where to look for the interval I_1. In particular, we say an integer n satisfies the bounds (**B**) if

B1: $4\left(1+3\mu+\sqrt{2}\sigma b^{5n/8}\right) \leq b^{n-1}$,

B2: $\sqrt{3\mu b}\sigma \leq b^{3n/8}$,

B3: $4\mu\left(3\mu+1+b^{3n/4}+2\sigma\mu^{-1/2}b^{5n/8}\right) \leq b^{n-1}$.

Generally, n need not be too large to satisfy these bounds. For example, if H is a $(2,10)$-happy function, assuming $n > 13$ is enough to guarantee that it satisfies bounds (**B**). This is well within the scope of the average computer as it is possible to compute the density of type-C integers in $[0, b^n-1]$ for n up to (and beyond) 1000 using the algorithm in Section 2. These bounds are necessary in the proof of Theorem 6.13.

Our first goal is to use an arbitrary n-strict interval I to construct a second interval, I_2, which is n_2-strict for some n_2 much larger than n and contains a similar density of type-C integers as I. The next lemma will be a helpful tool.

Lemma 6.10. *Let $I := [i_1, i_2]$, $J := [j_1, j_2]$ be integer intervals. Let $S \subseteq I$ and Y be an integer-valued random variable whose support is in J. For $k \in \mathbb{Z}$, denote the random variable $Y + k$ as $\tau_k(Y)$. Then there exists an integer $k \in [i_1 - j_2, i_2 - j_1]$ such that $\mathbb{P}\big(\tau_k(Y) \in S\big) \geq \frac{|S|}{|I|+|J|-1}$.*

Proof. The idea of the proof is that by averaging over all appropriate k, the distributions of $\tau_k(Y)$ should uniformly cover I. More formally, let $k_1 := i_1 - j_2$, $k_2 := i_2 - j_1$, and let K be the set of integers in the interval $[k_1, k_2]$. Note that $|K| = |I| + |J| - 1$. Pick k uniformly at random from K and consider the random variable $Z := \mathbb{P}\big(\tau_k(Y) \in S\big)$. Then

$$\mathbf{E}[Z] = \frac{1}{|K|}\sum_{k=k_1}^{k_2} \mathbb{P}\big(\tau_k(Y) \in S\big)$$

$$= \frac{1}{|I|+|J|-1}\sum_{k=k_1}^{k_2}\sum_{i\in S} \mathbb{P}\big(\tau_k(Y) = i\big)$$

$$= \frac{1}{|I| + |J| - 1} \sum_{i \in S} \sum_{k=k_1}^{k_2} \mathbb{P}(\tau_k(Y) = i). \tag{6.5}$$

Note that $\mathbb{P}\big(\tau_k(Y) = i\big) = \mathbb{P}\big(Y = i - k\big)$ and for $i \in S \subseteq I$ we have,

$$J \subseteq [i - k_2, i - k_1].$$

Thus, for all $i \in S$,

$$\sum_{k=k_1}^{k_2} \mathbb{P}\big(\tau_k(Y) = i\big) = \mathbb{P}(Y \in [i - k_2, i - k_1]) = 1.$$

Therefore,

$$(6.5) = \frac{|S|}{|I| + |J| - 1}.$$

So there exists k such that $\mathbb{P}\big(\tau_k(Y) \in S\big) \geq \mathbf{E}[Z] = \frac{|S|}{|I|+|J|-1}$.

□

Using Lemma 6.10, we will not lose much density assuming that $|I|$ is much larger than $|J|$. However, if Y_m is induced by the interval $[0, b^m - 1]$, then the random variable $H(Y_m)$ will be supported on a set J that is much too large. As a result, it will be more useful to consider a smaller interval where the bulk of the distribution lies.

Lemma 6.11. *Let Y be an integer-valued random variable with mean μ_Y and variance σ_Y^2, and let $\lambda > 0$. Let $S \subseteq [i_1, i_2] = I$ be a set of integers where $|S|/|I| = d$. Then there exists an integer $k \in [i_1 - (\mu_Y + \sigma_Y\lambda), i_2 - (\mu_Y - \sigma_Y\lambda)]$ such that*

$$\mathbb{P}\big(\tau_k(Y) \in S\big) \geq \left(1 - \frac{1}{\lambda^2}\right)\left(\frac{d}{1 + \frac{2\sigma_Y\lambda}{|I|}}\right).$$

Proof. By Chebyshev's Inequality[2] we have $\mathbb{P}(|Y - \mu_Y| < \sigma_Y\lambda) > 1 - \frac{1}{\lambda^2}$. Let Y' be Y conditioned on being in the interval $J := [\mu_Y - \sigma_Y\lambda, \mu_Y + \sigma_Y\lambda]$. Note that

$$|J| \leq 2\sigma_Y\lambda + 1.$$

Then, by Lemma 6.10, there exists $k \in [i_1 - (\mu_Y + \sigma_Y\lambda), i_2 - (\mu_Y - \sigma_Y\lambda)]$ such that $\mathbb{P}\big(\tau_k(Y') \in S\big) \geq \frac{|S|}{|I|+|J|-1} = \frac{d}{1+\frac{2\sigma_Y\lambda}{|I|}}$. Therefore, we have

$$\mathbb{P}\big(\tau_k(Y) \in S\big) \geq \mathbb{P}(Y \in J)\mathbb{P}\big(\tau_k(Y') \in S\big) \geq \left(1 - \frac{1}{\lambda^2}\right)\left(\frac{d}{1 + \frac{2\sigma_Y\lambda}{|I|}}\right).$$

[2]We certainly could do better than Chebyshev's Inequality here. However, the bounds it gives will suit our purposes fine.

□

It is possible to construct sets of intervals which, under the image of H, act as shifts of each other. For example, in base-10 (recall $H(0) = 0, H(1) = 1$) if the random variable X_1 is induced by $[1100, 1199]$ and X_2 is induced by $[0, 99]$, then $H(X_1) = H(X_2) + 2$.

We will now further expand on the example above. Let $n \in \mathbb{N}$ be divisible by 4. Let $B_0 := [0, b^{3n/4} - 1]$ and, for $k = 1, \ldots, \frac{n}{4}$, consider the interval

$$B_k := [b^{n-1} + b^{n-2} + \cdots + b^{n-k}, b^{n-1} + b^{n-2} + \cdots + b^{n-k} + b^{3n/4} - 1].$$

Then the intervals B_k will all be n-strict (with exception of B_0), and a random integer $x \in B_k$ will have the following base-b expansion:

$$x = \underbrace{11\ldots1}_{\text{k digits}}\ \underbrace{00\ldots0}_{\frac{n}{4}-k \text{ digits}}\ \underbrace{X_i X_{i-1} \ldots X_1}_{\frac{3n}{4} \text{ digits}}.$$

That is, x will have its first k digits equal to 1, the next $\frac{n}{4} - k$ digits equal to 0, and the remaining $\frac{3n}{4}$ digits will be i.i.d. random variables X_i taking values uniformly in the set $\{0, 1, \ldots, b-1\}$.

For $k = 0, \cdots, \frac{n}{4}$ let Y_k be the random variable which is uniform in B_k. Note that

$$H(Y_k) = H(Y_0) + k = \tau_k\left(H(Y_0)\right).$$

Recall $H(Y_0)$ has mean $\frac{3n}{4}\mu$, and variance $\frac{3n}{4}\sigma^2$. Consider an interval $I = [i_1, i_2]$ containing a set of type-C integers S, and let $\lambda > 0$. By Lemma 6.11, there exists $k' \in \left[i_1 - \left(\frac{3n}{4}\mu + \sqrt{\frac{3n}{4}}\lambda\sigma\right), i_2 - \left(\frac{3n}{4}\mu - \sqrt{\frac{3n}{4}}\lambda\sigma\right)\right]$ such that

$$\mathbb{P}\Big(\tau_{k'}\big(H(Y_0)\big) \in S\Big) \geq \Big(1 - \frac{1}{\lambda^2}\Big)\Big(\frac{d_C(I)}{1 + \frac{\sqrt{3n}\lambda\sigma}{|I|}}\Big). \tag{6.6}$$

Thus, if $I \subseteq \left[1 + \frac{3}{4}n\mu + \lambda\sigma\sqrt{\frac{3}{4}n}, \frac{1}{4}n + \frac{3}{4}n\mu - \lambda\sigma\sqrt{\frac{3}{4}n}\right]$, then $1 \leq k' \leq \frac{n}{4}$. Setting $k := k'$ produces the interval B_k, which will be n-strict with

$$d_C(B_k) \geq \Big(1 - \frac{1}{\lambda^2}\Big)\Big(\frac{d}{1 + \frac{\sqrt{3n}\lambda\sigma}{|I|}}\Big).$$

In fact, we have proven the following

Theorem 6.12. *Let $n \in \mathbb{N}$ be divisible by 4 and let $C \subset \mathbb{N}$. For $\lambda > 0$, define $J_{n,\lambda} := \left[1 + \frac{3}{4}n\mu + \lambda\sigma\sqrt{\frac{3}{4}n}, \frac{1}{4}n + \frac{3}{4}n\mu - \lambda\sigma\sqrt{\frac{3}{4}n}\right]$. Fix an interval $I \subseteq J_{n,\lambda}$. Then there exists an n-strict interval, I_2, such that $d_C(I_2) \geq d_C(I)\left(1 - \frac{1}{\lambda^2}\right)\left(\frac{1}{1+\frac{\sqrt{3n}\sigma\lambda}{|I|}}\right)$.*

The goal for the rest of the section is to use the previous theorem iteratively to construct a sequence of intervals $\{I_i\}_{i=1}^{\infty}$, each with high type-C density, such that each I_i is n_i-strict and the sequence $\{n_i\}_{i=1}^{\infty}$ grows quickly. One technical issue to worry about is that $d_C(I_{i+1}) < d_C(I_i)$ for all i. How much smaller $d_C(I_{i+1})$ is depends on how large we choose λ_i to be in each step. We wish to choose λ_i as large as possible, but choose λ_i too large and two bad things can happen: First, I_i will not be contained in J_{n_{i+1},λ_i} for any choice of n_{i+1}. Second, $\frac{\sqrt{3n}\sigma\lambda_i}{|I|}$ will not be small. We are helped by the fact that the sequence $\{n_i\}_{i=1}^{\infty}$ will grow super exponentially (in fact, $n_{i+1} = \Omega(b^{n_i})$). Choosing $\lambda_i = b^{n_i/8}$ in each step will work well; however, we will need the initial n_1 to be sufficiently large. The next theorem gives precise calculations. The proof follows from a number of routine calculations and estimations, some of which we have left for the appendix.

Theorem 6.13. *Suppose I is n-strict, where n satisfies bounds (**B**). Then there exists $n_2 \geq \frac{b^{n-1}}{\mu}$, and an n_2-strict interval I_2 such that*

$$d_C(I_2) \geq d_C(I)\left(1 - b^{-n/4}\right)\left(1 - \frac{2\sigma}{\sqrt{\mu}}b^{-n/8}\right).$$

Proof. As before, let $J_{m,\lambda} := \left[1 + \frac{3}{4}m\mu + \lambda\sigma\sqrt{\frac{3}{4}m}, \frac{1}{4}m + \frac{3}{4}m\mu - \lambda\sigma\sqrt{\frac{3}{4}m}\right]$. We assumed that I is n-strict, so $|I| = b^{3n/4}$. Write I as $[a, a + b^{3n/4} - 1]$. Setting $\lambda := b^{n/8}$, we attempt to find an n_2 divisible by 4 such that $I \subseteq J_{n_2,\lambda}$. It would be prudent to consider $f(m) := 1 + \frac{3}{4}m\mu + \lambda\sigma\sqrt{\frac{3}{4}m}$, which is the left endpoint of $J_{m,\lambda}$. We first find an integer n_2 such that $f(n_2) \leq a$ and $a - f(n_2)$ is small. By Lemma 6.22 in the Appendix, assuming n satisfies bounds (**B**), it follows that there exists n_2 such that:

- $4 \mid n_2$,
- $\frac{b^{n-1}}{\mu} \leq n_2 \leq \frac{4}{3\mu}b^n$,
- $0 \leq a - f(n_2) \leq 3\mu + 1$.

We now check that $I \subseteq J_{n_2,\lambda}$ in order to invoke Theorem 6.12. We already have that the left endpoint $f(n_2) \leq a$. It remains to check the right endpoints of I and J_{n_2}. We need to show that

$$a - 1 + b^{3n/4} \leq \frac{n_2}{4} + \frac{3n_2}{4}\mu - \lambda\sigma\sqrt{\frac{3n_2}{4}}. \tag{6.7}$$

The above is equivalent to

$$a - \left(\frac{3n_2}{4}\mu + \lambda\sigma\sqrt{\frac{3n_2}{4}} + 1\right) + b^{3n/4} \leq \frac{n_2}{4} - 2\lambda\sigma\sqrt{\frac{3n_2}{4}}.$$

Simplifying, the above follows from showing that

$$a - f(n_2) + b^{3n/4} + 2\lambda\sigma\sqrt{\frac{3n_2}{4}} \leq \frac{n_2}{4}.$$

Now let

$$\text{LHS} := a - f(n_2) + b^{3n/4} + 2\lambda\sigma\sqrt{\frac{3n_2}{4}}.$$

Then

$$\text{LHS} \leq 3\mu + 1 + b^{3n/4} + \lambda\sigma\sqrt{3n_2}.$$

Using the facts that $\lambda = b^{n/8}$ and $n_2 \leq \frac{4b^n}{3\mu}$, we get that

$$\text{LHS} \leq 3\mu + 1 + b^{3n/4} + 2\frac{\sigma}{\sqrt{\mu}}b^{5n/8}.$$

Now consider $\text{RHS} := \frac{n_2}{4}$. By the assumptions on n_2, we have

$$\text{RHS} \geq \frac{b^n}{4b\mu}.$$

So (6.7) follows from showing that

$$3\mu + 1 + b^{3n/4} + 2\frac{\sigma}{\sqrt{\mu}}b^{5n/8} \leq \frac{b^n}{4b\mu}.$$

The above is exactly the bound (**B3**). Therefore, $I \subseteq J_{n_2,\lambda}$. Thus, by applying Theorem 6.12 with $\lambda = b^{n/8}$, there exists an n_2-strict interval I_2 such that

$$d_C(I_2) \geq d_C(I)\Big(1 - \frac{1}{b^{n/4}}\Big)\Big(\frac{1}{1 + \frac{\sqrt{3n_2}\sigma b^{n/8}}{b^{3n/4}}}\Big).$$

Since $n_2 \leq \frac{4}{3\mu}b^n$, it follows that

$$\frac{1}{1 + \frac{\sqrt{3n_2}\sigma b^{n/8}}{b^{3n/4}}} \geq \frac{1}{1 + \frac{2\sigma}{\sqrt{\mu}}b^{-n/8}} \geq 1 - \frac{2\sigma}{\sqrt{\mu}}b^{-n/8}.$$

Thus, we conclude that $d_C(I_2) \geq d_C(I)\left(1 - b^{-n/4}\right)\left(1 - \frac{2\sigma}{\sqrt{\mu}}b^{-n/8}\right)$. □

Apply the previous theorem to our starting n_1-strict interval I_1 to get an n_2-strict interval I_2. Since $n_2 > n_1$, we can apply Theorem 6.13 again on I_2. Continuing in this manner produces a sequence of integers $\{n_i\}_{i=1}^{\infty}$ and n_i-strict intervals $\{I_i\}_{i=1}^{\infty}$ such that, for all i:

- $n_{i+1} \geq \frac{b^{n_i-1}}{\mu}$,
- $d_C(I_{i+1}) \geq d_C(I_i)\left(1 - b^{-n_i/4}\right)\left(1 - \frac{2\sigma}{\sqrt{\mu}} b^{-n_i/8}\right)$.

The second condition implies that, for all i,

$$d_C(I_i) \geq d_C(I_1) \prod_{i=1}^{\infty} \left(\left(1 - b^{-n_i/4}\right)\left(1 - \frac{2\sigma}{\sqrt{\mu}} b^{-n_i/8}\right)\right). \tag{6.8}$$

The following fact will help simplify the above expression. For positive real numbers x and α, if $x \geq 2\alpha > 0$, then

$$1 - \alpha x^{-1} \geq \frac{1}{1 + 2\alpha x^{-1}} \geq e^{-2\alpha x^{-1}}.$$

Therefore, (6.8) implies that

$$d_C(I_i) \geq d_C(I_1) \cdot \exp\left(\sum_{i=1}^{\infty} -2b^{-n_i/4} - \frac{4\sigma}{\sqrt{\mu}} b^{-n_i/8}\right).$$

For all $i \in \mathbb{N}$, it holds that $n_i \geq i n_1$ (it may happen that $n_2 < 2n_1$ if μ is very large, but assuming the bounds (**B**) this will not be the case). The sum in the previous inequality is the sum of two geometric series, one with ratio $r = b^{-n_1/4}$ and first term $a = -2b^{-n_1/4}$. The second has $r = b^{-n_1/8}$ and $a = \frac{-4\sigma}{\sqrt{\mu}} b^{-n_1/8}$. Recall that an infinite geometric series with $|r| < 1$, and first term a sums to

$$\frac{a}{1-r}.$$

Therefore, the first series sums to $\frac{-2b^{-n_1/4}}{1-b^{-n_1/4}}$, the second sums to $\frac{-4\sigma b^{-n_1/8}}{\sqrt{\mu}(1-b^{-n_1/8})}$. After simplifying we conclude that, for all i,

$$d_C(I_i) \geq d_C(I_1) \cdot \exp\left(\frac{-2}{b^{n_1/4} - 1} + \frac{-4\sigma}{\sqrt{\mu}(b^{n_1/8} - 1)}\right).$$

Thus, we have proven the following

Theorem 6.14. *Assume there exists n_1 satisfying the bounds (**B**) and an n_1-strict interval I_1. Then, for all $N \in \mathbb{N}$, there exists $n > N$ and an n-strict interval I such that*

$$d_C(I) \geq d_C(I_1) \cdot exp\left(\frac{2}{1-b^{n_1/4}} + \frac{4\sigma}{\sqrt{\mu}(1-b^{n_1/8})}\right).$$

6.4 Main Result

As in the previous section we continue to assume that H is a b-happy function with digit mean μ and digit variance σ^2. Also, we assume that we have experimentally found a suitable starting n_1-strict interval, I_1, with large type-C density for some $C \subset \mathbb{N}$. As in Section 2, for positive integers m, let Y_m denote the random variable induced by the interval $[0, b^m - 1]$.

In this section we give a proof of the following

Theorem 6.15. *Suppose I_1 is n_1-strict, where n_1 satisfies bounds (**B**). Let $\bar{d}$ denote the upper density of the set of type-C integers. Then*

$$\bar{d} \geq d_C(I_1) \cdot exp\left(\frac{2}{1-b^{n_1/4}} + \frac{4\sigma}{\sqrt{\mu}(1-b^{n_1/8})}\right).$$

The digit mean and digit variance for the case $(e, b) = (2, 10)$ are 28.5 and 721.05 respectively. In this case, if $n > 13$, then it satisfies bounds (**B**). After performing a computer search we find that the density of happy numbers in the interval $[10^{403}, 10^{404} - 1]$ is at least .185773; thus, there exists a 404-strict interval containing at least this density of happy numbers as a subset. Consider

$$\delta(n) := \Big(\frac{2}{1-b^{n/4}} + \frac{4\sigma}{\sqrt{\mu}(1-b^{n/8})}\Big).$$

Plugging in the value for n, we find that $e^{\delta(404)} > 1 - 10^{-49}$. Thus, by Theorem 6.15, the upper density of type-$\{1\}$ integers is at least .1857729. For the lower density, the type-$\{1\}$ density of $[10^{2367}, 10^{2368} - 1]$ is at most .11379. This implies that there is a 2368-strict interval with type-$\{4, 16, 37, 58, 89, 145, 42, 20\}$ density at least $1 - .11379$ (recall that there are only two cycles for the $(2, 10)$-function). We can then apply the main result to conclude that the upper density of type-$\{4, 16, 37, 58, 89, 145, 42, 20\}$ integers is at least $1 - .1138$. This gives the following

Corollary 6.16. *Let $\underline{d}$ and $\bar{d}$ be the lower and upper density of $(2, 10)$-happy numbers respectively. Then $\underline{d} < .1138$ and $\bar{d} > .18577$.*

The proof of Theorem 6.15 is somewhat technical despite having a rather intuitive motivation. For the sake of clarity we first give a sketch of how to use Theorems 6.14 and 6.8 in order to prove a lower bound on the upper density of type-C numbers.

Given our starting interval I_1, apply Theorem 6.14 to construct an n-strict interval I, where

$$d_C(I) \geq (1 - o(1))d_C(I_1).$$

Do this with n large enough as to make all the following error estimations arbitrarily small. Pick m_1 such that μm_1 (i.e., the mean of $H(Y_{m_1})$) lands in the interval I. Since $I \subseteq [b^{n-1}, b^n]$, we have that $m_1 = \Theta(b^n)$. This implies that the standard deviation of $H(Y_{m_1})$ is roughly $b^{n/2}$. This will be much less than $|I| = b^{3n/4}$ and thus the bulk of the distribution of $H(Y_{m_1})$ will lie in the interval I.

Next, use Lemma 6.11 with a large λ to find an integer k for which

$$\mathbb{P}\left(\tau_k\big(H(Y_{m_1})\big) \text{ is type-C}\right) \geq (1 - o(1))d_C(I_1).$$

Note that this k will be smaller than $|I| = b^{3n/4}$ and that the mean of $\tau_k\big(H(Y_{m_1})\big)$ is equal to $\mu m_1 + k$. Clearly, there exists an integer m_2 such that

$$|\mu m_2 - (\mu m_1 + k)| \leq \mu.$$

Consider the random variable $H(Y_{m_2})$. The means of $H(Y_{m_2})$ and $\tau_k\left(H(Y_{m_1})\right)$ are almost equal. Since k is much smaller relative to m_1 and m_2, the variance of these two distributions will be close. Furthermore, the distributions of $H(Y_{m_2})$ and $\tau_k\big(H(Y_{m_1})\big)$ are asymptotically locally normal, so we may apply the local limit law to conclude that the distributions are point-wise close near the means. Thus,

$$\mathbb{P}\big(H(Y_{m_2}) \text{ is type-C}\big) \geq (1 - o(1))d_C(I_1).$$

This implies that the interval $[0, b^{m_2} - 1]$ has type-C density at least $d_C(I_1)\,(1 - o(1))$. Note that in the above analysis, we may take n (and therefore m_2) to be arbitrarily large. This lower bounds the upper density of type-C integers by $d_C(I_1)(1 - o(1))$. In

fact, the only contribution to the error term is from the application of Theorem 6.14 (the rest of the error tends to 0 as n tends to infinity).

6.4.1 Some Lemmas

We will now begin to prove the main result. We have broken some of the pieces down for 3 lemmas. The proofs primarily consist of calculations and we leave them for after the proof of the main result. Note that Lemma 6.19 (part 1) is the only place where the local limit law is used.

Lemma 6.17. *There exists a sufficiently large N such that, if $n > N$ and I is an n-strict interval, then there exists $m \in \mathbb{N}$ with the property that*

$$[\mu m - \sigma m^{5/8}, \mu m + \sigma m^{5/8}] \subseteq I.$$

Lemma 6.18. *Let $\epsilon > 0$ be given (assume as well that $\epsilon \leq 1$). Let $\lambda := \sqrt{\frac{6}{\epsilon}}$. Then there exists a sufficiently large N such that, if n, m_1, and I all satisfy:*

- $n > N$,
- $m_1 \in [\frac{b^{n-1}}{\mu}, \frac{b^n}{\mu}]$,
- *I is n-strict,*

then the following hold:

1. $\lambda \leq {m_1}^{1/8}$,
2. $\left| 1 - \frac{1}{1+\frac{2\lambda\sigma\sqrt{m_1}}{b^{3n/4}}} \right| \leq \frac{\epsilon}{6}$.

Lemma 6.19. *Let $\epsilon > 0$ be given (assume as well that $\epsilon \leq 1$). Let $T := \frac{2\sqrt{6}}{\sqrt{\epsilon}}$, and $\lambda := \frac{\sqrt{6}}{\sqrt{\epsilon}}$. Then there exists a sufficiently large N such that, if n, m_1, m_2, k, and I all satisfy:*

- $n > N$,
- $m_1 \in [\frac{b^{n-1}}{\mu}, \frac{b^n}{\mu}], m_2 \in [\frac{b^{n-2}}{\mu}, \frac{b^{n+1}}{\mu}]$,

- $|k| \leq b^{3n/4}$,
- $|\mu m_1 + k - \mu m_2| \leq \mu$,
- *I is n-strict,*

then the following hold:

1. *For $i \in \{1,2\}$,* $\max_{|t|\leq T} \left| 1 - \frac{\mathbb{P}(H(Y_{m_i}) = \lfloor \mu m_i + t\sigma\sqrt{m_i} \rfloor)}{\frac{e^{t^2/2}}{\sqrt{2\pi m_i}\sigma}} \right| \leq \frac{\epsilon}{6}$.
2. $\left| 1 - \sqrt{\frac{m_1}{m_2}} \right| \leq \frac{\epsilon}{6}$.
3. *For any real numbers t_1 and t_2, where $t_1 \in [\frac{-T}{2}, \frac{T}{2}]$ and*

$$\mu m_1 + k + t_1\sigma\sqrt{m_1} = \mu m_2 + t_2\sigma\sqrt{m_2},$$

it holds that $t_2 \in [-T, T]$ and $|1 - e^{{t_1}^2 - {t_2}^2/2}| \leq \frac{\epsilon}{6}$.

6.4.2 Proof of Theorem 4.1

Proof. In order to lower bound the upper density of type-C integers, it suffices to show that, for all $\epsilon > 0$ and $N_1 \in \mathbb{N}$, there exists $m > N_1$ such that

$$d_C([0, b^m - 1]) \geq d_C(I_1) \cdot \exp\left(\frac{2}{1 - b^{n/4}} + \frac{4\sigma}{\sqrt{\mu}(1 - b^{n/8})} \right) (1 - \epsilon).$$

Let ϵ and N_1 be arbitrary (with $\epsilon \leq 1$). Set $T := \frac{2\sqrt{6}}{\sqrt{\epsilon}}$. Also, in anticipation of applying Lemma 6.11, set $\lambda := \frac{\sqrt{6}}{\sqrt{\epsilon}}$.

First, pick $N > N_1$ large enough to apply Lemmas 6.17, 6.18, and 6.19. By Theorem 6.14, there exists an n-strict interval I, where $n > N$ and

$$d_C(I) \geq d_C(I_1) \cdot \exp\left(\frac{2}{1 - b^{n_1/4}} + \frac{4\sigma}{\sqrt{\mu}(1 - b^{n_1/8})} \right). \tag{6.9}$$

For $m \in \mathbb{N}$, let

$$J_m := [\mu m - \sigma m^{5/8}, \mu m + \sigma m^{5/8}].$$

Recall that $\mathbf{E}[H(Y_m)] = \mu m$ and $\mathbf{Var}[H(Y_m)] = \sigma^2 m$. Hence, J_m is where the bulk of the distribution of $H(Y_m)$ lands. Pick m_1 such that $J_{m_1} \subseteq I$ (the existence of such m_1 is guaranteed by Lemma 6.17). Note that $m_1 \in [\frac{b^{n-1}}{\mu}, \frac{b^n}{\mu}]$ since I is n-strict. Let S be

the set of type-C integers in I. Apply Lemma 6.11 on the random variable $H(Y_{m_1})$ to find an integer k such that

$$\mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) \in S\Big) \geq d_C(I)\left(1-\frac{1}{\lambda^2}\right)\Big(\frac{1}{1+\frac{2\lambda\sigma\sqrt{m_1}}{|I|}}\Big). \tag{6.10}$$

Since $J_{m_1} \subseteq I$ and $|I| = b^{3n/4}$, it follows that $k \leq b^{3n/4}$.

Let $\tau_k(J_{m_1}) := [a+k, b+k]$, where $J_{m_1} = [a, b]$. Let S' be the set of type-C integers in interval $\tau_k(J_{m_1})$. Recall the proof of Lemma 6.11. In particular, we applied Lemma 6.10 after ignoring the tails of the distribution of $H(Y_{m_1})$ outside of $\lambda\sigma\sqrt{m_1}$ from the mean. Since $\lambda \leq {m_1}^{1/8}$ (by Lemma 6.18, part 1), we may replace (6.10) by the stronger conclusion that

$$\sum_{i\in S'}\mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big) \geq d_C(I)\Big(1-\frac{1}{\lambda^2}\Big)\Big(\frac{1}{1+\frac{2\lambda\sigma\sqrt{m_1}}{|I|}}\Big).$$

Using the assumption that $\lambda = \sqrt{\frac{6}{\epsilon}}$ and part 2 of Lemma 6.18, we simplify the above as

$$\sum_{i\in S'}\mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big) \geq d_C(I)\left(1-\frac{\epsilon}{6}\right)^2. \tag{6.11}$$

Now pick $m_2 \in \mathbb{N}$ such that $|m_1\mu + k - m_2\mu| \leq \mu$. Since $|k| \leq b^{3n/4}$, it follows that $m_2 \in [\frac{b^{n-2}}{\mu}, \frac{b^{n+1}}{\mu}]$. In particular m_1, m_2, n, k, and I now all satisfy the conditions of Lemma 6.19. It remains to show that near the mean of $\tau_k\big(H(Y_{m_1})\big)$, the distributions of $\tau_k\big(H(Y_{m_1})\big)$ and $H(Y_{m_2})$ are similar. This will imply that the interval $[0, b^{m_2}-1]$ contains a large density of type-C integers. Making this precise, we prove the following

Claim 6.20. *For integers $i \in \tau_k\big(J_{m_1}\big)$,*

$$\frac{\mathbb{P}\big(H(Y_{m_2}) = i\big)}{\mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big)} \geq \left(1-\frac{\epsilon}{6}\right)^4.$$

Proof. Let $i \in \tau_k(J_{m_1})$ be fixed and pick t_1, t_2 such that

$$i = \mu m_1 + k + t_1\sigma\sqrt{m_1} = \mu m_2 + t_2\sigma\sqrt{m_2}.$$

It is important now that we had chosen $\lambda = \frac{T}{2}$, this implies that $|t_2| \leq T$ (see Lemma 6.19 part 3). We can use the local limit law to estimate the distributions of $\tau_k\big(H(Y_{m_1})\big)$

and $H(Y_{m_2})$. By Lemma 6.19 part 1,

$$\mathbb{P}\big(H(Y_{m_2}) = i\big) = \mathbb{P}\big(H(Y_{m_2}) = \mu m_2 + t_2\sigma\sqrt{m_2}\big) \geq \frac{e^{-{t_2}^2/2}}{2\pi\sigma\sqrt{m_2}}\left(1 - \frac{\epsilon}{6}\right)$$

and

$$\mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big) = \mathbb{P}\big(H(Y_{m_1}) = \mu m_1 + t_1\sigma\sqrt{m_1}\big) \leq \frac{e^{-{t_1}^2/2}}{2\pi\sigma\sqrt{m_1}}\left(1 + \frac{\epsilon}{6}\right).$$

Hence,

$$\frac{\mathbb{P}\big(H(Y_{m_2}) = i\big)}{\mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big)} \geq \exp\left(({t_1}^2 - {t_2}^2)/2\right)\frac{\sqrt{m_1}}{\sqrt{m_2}}\frac{(1-\frac{\epsilon}{6})}{(1+\frac{\epsilon}{6})}.$$

The above, by Lemma 6.19 parts 2 and 3, is at least $\left(1 - \frac{\epsilon}{6}\right)^4$. □

Putting it all together, we have shown that

$$\begin{aligned}
d_C([0, b^{m_2} - 1]) &\geq \sum_{i \in S'} \mathbb{P}\big(H(Y_{m_2}) = i\big) \\
&= \sum_{i \in S'} \frac{\mathbb{P}\big(H(Y_{m_2}) = i\big)}{\mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big)} \mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big). \\
&\geq \left(1 - \frac{\epsilon}{6}\right)^4 \sum_{i \in S'} \mathbb{P}\Big(\tau_k\big(H(Y_{m_1})\big) = i\Big) && \text{(Claim 1)} \\
&\geq d_C(I)\left(1 - \frac{\epsilon}{6}\right)^6 && \text{(Equation 6.11)} \\
&\geq d_C(I)(1 - \epsilon).
\end{aligned}$$

To conclude the proof, equation (6.9) implies that

$$d_C([0, b^{m_2} - 1]) \geq d_C(I_1) \cdot \exp\left(\frac{2}{1 - b^{n_1/4}} + \frac{4\sigma}{\sqrt{\mu}(1 - b^{n_1/8})}\right)(1 - \epsilon).$$

□

We conclude this section with the proofs of the lemmas used in the previous theorem.

Proof of Lemma 6.17

Proof. For $m \in \mathbb{N}$, let $J_m := [\mu m - \sigma m^{5/8}, \mu m + \sigma m^{5/8}]$. If I is an n-strict interval, then $I \subseteq [b^{n-1}, b^n - 1]$. Note that $\mu m \in I$ implies that $m = O(b^n)$. This in turn shows that

$$|J_m| = O(b^{5n/8}) \ll |I| = b^{3n/4}.$$

Comparing the growth rates of $|J_m|$ and $|I|$ it is clear that we can pick N_1 large enough such that $n > N_1$ implies that there exists m with $J_m \subseteq I$. □

Proof of Lemma 6.18

Proof. We find N_1, N_2 for the two parts respectively and then choose $N = \max(N_1, N_2)$.

1. λ is a fixed constant here and it is assumed that $m_1 \geq \frac{b^{n-1}}{\mu}$, so the result is trivial (this gives N_1).

2. For $x > 0$, to show that $\left|1 - \frac{1}{1+x}\right| \leq \frac{\epsilon}{6}$, it is equivalent to prove that

$$\left(1 - \frac{\epsilon}{6}\right)(1+x) \leq 1 \leq \left(1 + \frac{\epsilon}{6}\right)(1+x).$$

The above follows if $x \leq \frac{\epsilon}{6}$. Thus, the result will follow by finding N large enough such that $\frac{2\sigma\lambda\sqrt{m_1}}{b^{3n/4}} \leq \frac{\epsilon}{6}$. Using the assumption that $m_1 \leq \frac{b^n}{\mu}$, we get

$$\frac{2\sigma\lambda\sqrt{m_1}}{b^{3n/4}} \leq \frac{2\sigma\lambda}{\sqrt{\mu}b^{n/4}}.$$

This is equivalent to

$$\frac{12\sigma\lambda}{\sqrt{\mu\epsilon}} \leq b^{n/4}.$$

Hence, picking $N_2 \geq 4\log_b(\frac{12\sigma\lambda}{\sqrt{\mu\epsilon}})$ suffices.

□

Proof of Lemma 6.19

Proof. We first find N_1, N_2, and N_3 for the 3 parts respectively, and then define $N := \max(N_1, N_2, N_3)$.

1. For each m, we have

$$H(Y_m) = \sum_{i=1}^{m} H(X_i),$$

where each X_i is uniform in the set $\{0, 1, \cdots, b-1\}$. Recall that it is assumed that $H(0) = 0$ and $H(1) = 1$. In particular, the random variables $H(X_i)$ satisfy the aperiodic condition required by Theorem 6.8. Thus, the result follows from applying Theorem 6.8 to the sum $\sum_{i=1}^{m} H(X_i)$ with finite interval $[-T, T]$. Fix M large enough such that $m > M$ implies that the $O(m^{-1/2})$ term in Theorem 6.8 is less than $\frac{\epsilon}{6}$. By assumption,

we have that both m_1 and m_2 are larger than $\frac{b^{n-2}}{\mu}$. Hence, setting $N_1 = \log_b(\mu M) + 2$ suffices.

2. Ignoring the square root, it suffices to show that

$$\left|1 - \frac{m_1}{m_2}\right| \leq \frac{\epsilon}{6}. \tag{6.12}$$

By assumption

$$|\mu m_1 + k - \mu m_2| \leq \mu.$$

Dividing through by μm_2, it follows that

$$\left|1 - \frac{m_1}{m_2} - \frac{k}{\mu m_2}\right| \leq \frac{1}{m_2}.$$

This implies that

$$\frac{k}{\mu m_2} - \frac{1}{m_2} \leq 1 - \frac{m_1}{m_2} \leq \frac{1}{m_2} + \frac{k}{\mu m_2}.$$

Thus, (6.12) follows from showing that $\frac{1}{m_2} + \frac{k}{\mu m_2} \leq \frac{\epsilon}{6}$. Using the assumption that $m_2 \geq \frac{b^{n-2}}{\mu}$ and $|k| \leq b^{3n/4}$, it follows that

$$\frac{1}{m_2} + \frac{k}{\mu m_2} \leq \mu b^{-(n-2)} + b^{2-(n/4)}.$$

Therefore, picking $N_2 := \max(\log_b(\frac{12\mu}{\epsilon}) + 2, 4\log_b(\frac{12}{\epsilon}) + 2)$ suffices.

3. We first find N' such that $n > N'$ implies that $t_2 \in [-T, T]$. We start with the assumption that

$$\mu m_1 + k + t_1\sigma\sqrt{m_1} = \mu m_2 + t_2\sigma\sqrt{m_2}.$$

Using the facts that $|\mu m_1 + k - \mu m_2| \leq \mu$ and $|t_1| \leq \frac{T}{2}$, this implies that

$$|t_2| \leq \frac{\mu}{\sigma\sqrt{m_2}} + \frac{T\sqrt{m_1}}{2\sqrt{m_2}}.$$

We assumed that $m_2 \geq \frac{b^{n-2}}{\mu}$. Also, in part (2) we showed that there exists N_2 such that $n > N_2$ implies that $\frac{\sqrt{m_1}}{\sqrt{m_2}} \leq \left(1 + \frac{\epsilon}{6}\right) \leq \frac{7}{6}$. Hence, if we take $N' > N_2$, it follows that

$$|t_2| \leq \frac{\mu^2}{\sigma b^{n-2/2}} + \frac{7T}{12}.$$

Pick $N' > N_2$ large enough such that $n > N'$ implies that $\frac{\mu^2}{\sigma b^{n-2/2}} \leq \frac{5T}{12}$. This will take care of the size of t_2.

Now we must show that there exists N'' large enough such that $n > N''$ implies that

$$\left|1 - e^{({t_1}^2 - {t_2}^2)/2}\right| \leq \frac{\epsilon}{6}.$$

For a real number x, if we wish to show that $|1 - e^x| \leq \frac{\epsilon}{6}$, it is equivalent to prove that

$$\ln\left(1 - \frac{\epsilon}{6}\right) \leq x \leq \ln\left(1 + \frac{\epsilon}{6}\right).$$

Set $\epsilon^* := \min\left(\ln\left(1 + \frac{\epsilon}{6}\right), \left|\ln\left(1 - \frac{\epsilon}{6}\right)\right|\right)$. We find N'' such that $n > N''$ implies that

$$\left|\frac{{t_2}^2 - {t_1}^2}{2}\right| \leq \epsilon^*.$$

It was assumed that

$$\mu m_1 + k + t_1\sigma\sqrt{m_1} = \mu m_2 + t_2\sigma\sqrt{m_2}.$$

Equivalently

$$\mu m_1 + k - \mu m_2 = t_2\sigma\sqrt{m_2} - t_1\sigma\sqrt{m_1}.$$

Applying the assumption that the left hand side is at most μ and dividing both sides by $\sigma\sqrt{m_2}$, we get

$$\left|t_2 - t_1\sqrt{\frac{m_1}{m_2}}\right| \leq \frac{\mu}{\sqrt{m_2}\sigma}.$$

Rearranging, this gives

$$\left|t_2 - t_1 + t_1(1 - \sqrt{\frac{m_1}{m_2}})\right| \leq \frac{\mu}{\sqrt{m_2}\sigma}.$$

This implies that

$$|t_2 - t_1| \leq \frac{\mu}{\sqrt{m_2}\sigma} + \left|t_1(1 - \sqrt{\frac{m_1}{m_2}})\right|.$$

We assumed that $m_2 \geq \frac{b^{n-2}}{\mu}$ and $|t_1| \leq T$. By part (2), if we chose $N'' > N_2$, then

$$\left|1 - \sqrt{\frac{m_1}{m_2}}\right| \leq \mu b^{-(n-2)} + b^{-(n-2)/4}.$$

Putting this together, it follows that

$$\left|\frac{{t_2}^2 - {t_1}^2}{2}\right| = \left|\left(\frac{t_2 + t_1}{2}\right)(t_2 - t_1)\right| \leq T\Big(\frac{\mu^{3/2}b^{-(n-2)/2}}{\sigma} + T(\mu b^{-(n-2)} + b^{-(n-2)/4})\Big).$$

Now, since T, μ, σ, and b are all constants, it follows that the right hand side tends to 0 as n goes to infinity. Therefore, there exists N'' such that $n > N''$ implies that the right hand side is at most ϵ^*. Finally, set $N_3 := \max(N', N'')$. □

6.5 Experimental Data

The data[3] presented in this section is the result of short computer searches, so the bounds surely can be improved with more computing time. Floating point approximation with conservative rounding was used.

6.5.1 Finding an Appropriate n-strict Interval

If n is divisible by 4 and the interval $[b^{n-1}, b^n - 1]$ has type-C density d, then there exists an n-strict interval with type-C density at least d which we may apply Theorem 6.15 to. The type-C density of $[b^{n-1}, b^n - 1]$ for various n can be quickly calculated by first computing the densities of intervals of the form $[0, b^n - 1]$; the algorithm which does this was discussed in Section 2. After an appropriate n-strict interval is found, we check to see that n satisfies bounds (**B**), compute the error term, and find the desired bound. Our results show that in almost all cases, the asymptotic density of type-C numbers does not exist.

6.5.2 Explanation of Results

The following information is given in tables (in the order of column in which they appear):

1. The cycle C in which type-C densities are being computed.
2. The lower bound on the upper density (UD) implied by Theorem 6.15.
3. The upper bound on the lower density (LD) implied by Theorem 6.15.
4. The n such that the interval $[b^{n-1}, b^n - 1]$ is used to find the bound (denoted as UD n or LD n).
5. The $\delta(n) = \left(\frac{2}{1-b^{n/4}} + \frac{4\sigma}{\sqrt{\mu}(1-b^{n/8})}\right)$ part of the error term for Theorem 6.15 (we only present an upper bound on $|\delta(n)|$, the true number is always negative). In all

[3]Data generated by fellow graduate student, Patrick Devlin.

cases the error is small enough not to affect the bounds as we only give precision of about 5 or 6 decimal places.

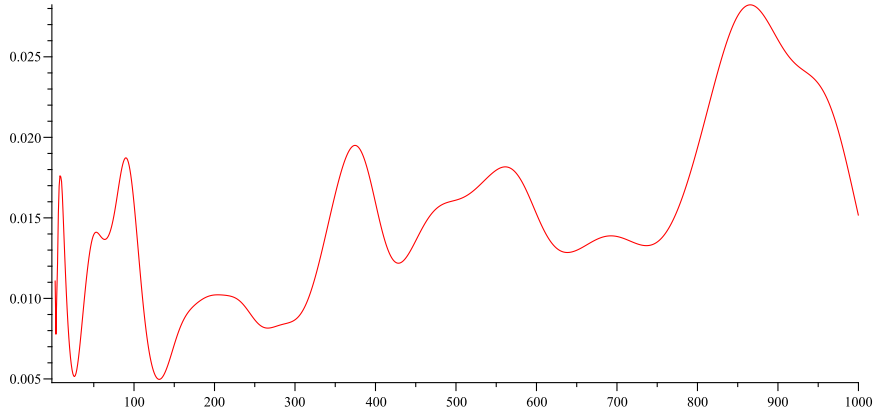
Cubing the Digits in Base-10

In this case, if $n > 16$, then it satisfies bounds (**B**). Table 1 shows the results for the cycles when studying the $(3, 10)$-happy function. There are 9 possible cycles. Figure 2 graphs the density of type-$\{1\}$ integers less than 10^n. It is easy to prove, in this case, that $3 \mid n$ if an only if n is type-$\{153\}$.

Table 6.1: Bounds for the cycles appearing for the $(3, 10)$-function

Cycle	UD	LD	UD n	LD n	UD $\delta(n)$	LD $\delta(n)$
{1}	$> .028219$	$< .0049761$	10^{864}	10^{132}	$< 10^{-106}$	$< 10^{-14}$
{55,250,133}	$> .06029$	$< .0447701$	10^{208}	10^{964}	$< 10^{-24}$	$< 10^{-118}$
{136,244}	$> .024909$	$< .006398$	10^{204}	10^{420}	$< 10^{-23}$	$< 10^{-51}$
{153}	$= \frac{1}{3}$	$= \frac{1}{3}$	N/A	N/A	N/A	N/A
{160,217,352}	$> .050917$	$< .03184$	10^{160}	10^{456}	$< 10^{-18}$	$< 10^{-56}$
{370}	$> .19905$	$< .16065$	10^{276}	10^{560}	$< 10^{-32}$	$< 10^{-68}$
{371}	$> .30189$	$< .288001$	10^{836}	10^{420}	$< 10^{-102}$	$< 10^{-50}$
{407}	$> .04532$	$< .0314401$	10^{420}	10^{836}	$< 10^{-50}$	$< 10^{-103}$
{919,1459}	$> .04425$	$< .01843$	10^{916}	10^{120}	$< 10^{-112}$	$< 10^{-13}$

Figure 6.2: Density of type-$\{1\}$ integers in the interval $[0, 10^n - 1]$ for the $(3, 10)$-function



A More General Function

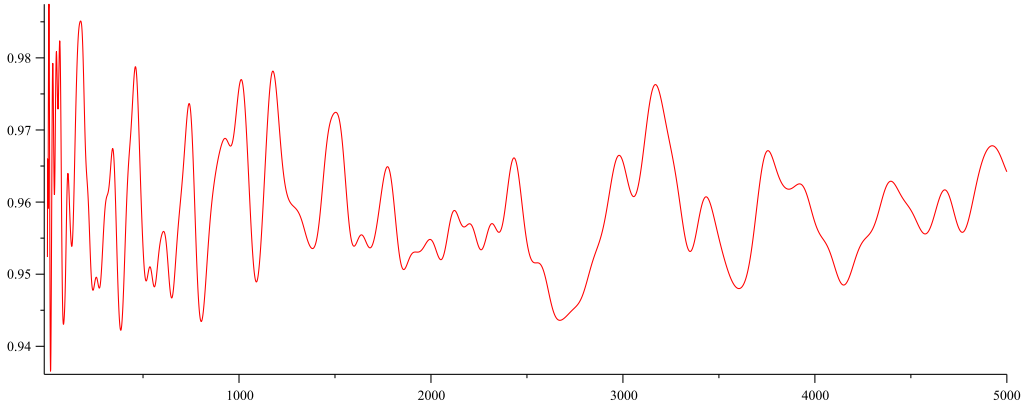
In order to emphasize the generality of Theorem 6.15, we consider the function in base-7 with digit sequence $[0, 1, 7, 4, 17, 9, 13]$. There are only two cycles for this function,

both are fixed points. Written in base-10 the cycles are $\{1\}$ and $\{20\}$. Figure 3 graphs the relative density of type-$\{1\}$ numbers. Table 2 shows the bounds derived. As there are only two cycles, we focus on the cycle $\{1\}$. In this case, if $n > 12$, then it satisfies bounds (**B**).

Table 6.2: Bounds for the cycles appearing for the function with digit sequence $[0, 1, 7, 4, 17, 9, 13]$

Cycle	UD	LD	UD n	LD n	UD $\delta(n)$	LD $\delta(n)$
$\{1\}$	$> .9858$	$< .94222$	7^{176}	7^{384}	$< 10^{-17}$	$< 10^{-40}$

Figure 6.3: Density of type-$\{1\}$ integers in the interval $[0, 7^n - 1]$ for Digit Sequence $[0, 1, 7, 4, 17, 9, 13]$



6.6 Chapter Appendix

Lemma 6.21. *Fix $a > 0$. Assume that $f : \mathbb{R}^+ \to \mathbb{R}^+$ has continuous first and second derivatives such that, for all $x \in \mathbb{R}^+, f'(x) > 0$ and $f''(x) < 0$. Also, assume that $\lim_{x\to\infty} f(x) = \infty$. Furthermore, suppose we have $x^* \in \mathbb{R}^+$ such that $f(x^* + 1) \leq a$. Then there exists $n \in \mathbb{N}$ such that $n \geq x^*$ and $0 \leq a - f(n) \leq f'(x^*)$.*

Proof. This follows from a first order Taylor approximation of the function f. Let x^* such that $f(x^* + 1) \leq a$ be given. Set $n := \sup\{m \in \mathbb{N} | f(m) \leq a\}$. Since f is strictly increasing and unbounded this n exists. Note that $f(n) \leq a$ and $f(n + 1) > a$. It also follows that $n \geq x^*$ as otherwise $\lceil x^* \rceil$ would be the supremum. By the concavity of f,

we have

$$f(n+1) - f(n) \leq f'(n) \leq f'(x^*).$$

However, $f(n+1) > a$, so we conclude that $0 \leq a - f(n) \leq f'(x^*)$. □

Lemma 6.22. *Let n be a positive integer, $\lambda = b^{n/8}$, and $a \in [b^{n-1}, b^n]$. Let μ and σ be the digit mean and variance of some b-happy function H. Also, assume that n satisfies bounds (**B**). Let $f(n) := 1 + \frac{3}{4}\mu n + \lambda\sigma\sqrt{\frac{3}{4}n}$. Then there exists an integer n_2 such that:*

- $\frac{b^{n-1}}{\mu} \leq n_2 \leq \frac{4}{3\mu}b^n$,
- $4 \mid n_2$,
- $0 \leq a - f(n_2) \leq 3\mu + 1$.

Proof. Since we require that $4 \mid n_2$, we apply Lemma 6.21 on the function

$$g(m) = f(4m) = 1 + 3\mu m + \lambda\sigma\sqrt{3m}.$$

Let $x^* := \frac{b^{n-1}}{4\mu}$. We first check that $g(x^*+1) \leq a$. By assumption, $a \geq b^{n-1}$. Therefore, we need to show that

$$1 + 3\mu\Big(\frac{b^{n-1}}{4\mu} + 1\Big) + b^{n/8}\sigma\sqrt{3\Big(\frac{b^{n-1}}{4\mu} + 1\Big)} \leq b^{n-1}.$$

Simplifying the above, it suffices to show that

$$1 + 3\mu + b^{5n/8}\sigma\sqrt{\frac{3}{4b\mu} + 3b^{-n}} \leq \frac{b^{n-1}}{4}. \tag{6.13}$$

To keep the results of this work as general as possible, we only assumed that $\mu \geq \frac{1}{b}$ (this would correspond to the quite uninteresting b-happy function H which maps all digits to 0 except for the digit 1). Also it is clear that $b^n \geq 3$, and therefore

$$\frac{3}{4b\mu} + 3b^{-n} \leq \frac{3}{4} + 1 \leq 2.$$

Plugging this in and rearranging, we see that equation (6.13) follows if

$$4(1 + 3\mu + \sqrt{2}\sigma b^{5n/8}) \leq b^{n-1}.$$

This is exactly the bound (**B1**) and is true by assumption. Therefore, by Lemma 6.21, there exists $m \in \mathbb{N}$ such that

$$0 \leq a - g(m) \leq g'(x^*).$$

Also,

$$\begin{aligned} g'(x^*) &= 3\mu + \frac{\sqrt{3}\sigma b^{n/8}}{2\sqrt{\frac{b^{n-1}}{4\mu}}} \\ &= 3\mu + \sqrt{3\mu b}\sigma b^{-3n/8}. \end{aligned}$$

Again, by the assumption (**B2**) on n, the previous statement is bounded above by $3\mu+1$. Set $n_2 := 4m$. Then $4 \mid n_2, n_2 \geq \frac{b^{n-1}}{\mu}$, and $0 \leq a - f(n_2) \leq 3\mu + 1$. Finally, we note that $f(\frac{4}{3\mu}b^n) > a$ and, since f is strictly increasing, we conclude that $n_2 \leq \frac{4}{3\mu}b^n$. $\square$

References

[Aar08] Scott Aaronson. Quantum certificate complexity. *J. Comput. Syst. Sci.*, 74(3):313–322, 2008.

[Bar82] AD Barbour. Poisson convergence and random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 92, pages 349–359. Cambridge Univ Press, 1982.

[BBC+01] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald De Wolf. Quantum lower bounds by polynomials. *Journal of the ACM (JACM)*, 48(4):778–797, 2001.

[BdW02] Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theor. Comput. Sci.*, 288(1):21–43, 2002.

[BKR89] Andrew D Barbour, Michal Karoński, and Andrzej Ruciński. A central limit theorem for decomposable random variables with applications to random graphs. *Journal of Combinatorial Theory, Series B*, 47(2):125–145, 1989.

[CFGS88] Fan RK Chung, Zoltán Füredi, Ronald L Graham, and Paul Seymour. On induced subgraphs of the cube. *Journal of Combinatorial Theory, Series A*, 49(1):180–187, 1988.

[Che75] Louis HY Chen. Poisson approximation for dependent trials. *The Annals of Probability*, pages 534–545, 1975.

[DKR14] Bobby DeMarco, Jeff Kahn, and Amanda Redlich. Modular statistics for subgraph counts in sparse random graphs. *arXiv preprint arXiv:1402.2264*, 2014.

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge university press, 2010.

[Erd60] Paul Erdos. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[ESS+00] Esam El-Sedy, Samir Siksek, et al. On happy numbers. *Rocky Mountain Journal of Mathematics*, 30(2):565–570, 2000.

[FS09] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. cambridge University press, 2009.

[Für88] Zoltán Füredi. Matchings and covers in hypergraphs. *Graphs and Combinatorics*, 4(1):115–206, 1988.

[Gil13] Justin Gilmer. On the density of happy numbers. *INTEGERS*, 13:2, 2013.

[GK] Justin Gilmer and Swastik Kopparty. A local central limit theorem for the number of triangles in a random graph. *Random Structures and Algorithms, (submitted).*

[GKS] Justin Gilmer, Michal Koucky, and Michael Saks. A new approach to the sensitivity conjecture. *The 6th Innovations in Theoretical Computer Science (ITCS) conference, (to appear).*

[GL92] Craig Gotsman and Nathan Linial. The equivalence of two problems on the cube. *Journal of Combinatorial Theory, Series A*, 61(1):142–146, 1992.

[GSS] Justin Gilmer, Shrikanth Srinivasan, and Michael Saks. Composition limits and separating examples for some boolean function complexity measures. *Combinatorica, (to appear).*

[GT07] HG Grundman and EA Teeple. Sequences of generalized happy numbers with small bases. *J. Integer Seq*, 10, 2007.

[Guy04] Richard Guy. *Unsolved problems in number theory*, volume 1. Springer, 2004.

[HKP11] Pooya Hatami, Raghav Kulkarni, and Denis Pankratov. *Variations on the Sensitivity Conjecture.* Number 4 in Graduate Surveys. Theory of Computing Library, 2011.

[Jan92] Svante Janson. *Orthogonal decompositions and functional limit theorems for random graph statistics*, volume 534. American Mathematical Soc., 1992.

[Kar84] Michal Karonski. *Balanced subgraphs of large random graphs.* UAM, 1984.

[KK04] Claire Kenyon and Samuel Kutin. Sensitivity, block sensitivity, and l-block sensitivity of boolean functions. *Information and Computation*, 189(1):43–53, 2004.

[KK13] Phokion G Kolaitis and Swastik Kopparty. Random graphs and the parity quantifier. *Journal of the ACM (JACM)*, 60(5):37, 2013.

[KR83] Michał Karoński and Andrezj Ruciński. On the number of strictly balanced subgraphs of a random graph. In *Graph theory*, pages 79–83. Springer, 1983.

[LMP04] Martin Loebl, Jiří Matoušek, and Ondřej Pangrác. Triangles in random graphs. *Discrete mathematics*, 289(1):181–185, 2004.

[LY02] Laszlo Lovasz and Neal E Young. Lecture notes on evasiveness of graph properties. *arXiv preprint cs/0205031*, 2002.

[Mid04] G. Midrijanis. Exact quantum query complexity for total boolean functions. *arXiv preprint quant-ph/0403168*, 2004.

[Nis91] Noam Nisan. CREW PRAMs and Decision Trees. *SIAM J. Comput.*, 20(6):999–1007, 1991.

[NS94] Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994.

[NW88] Krzysztof Nowicki and John C Wierman. Subgraph counts in random graphs using incomplete u-statistics methods. *Annals of Discrete Mathematics*, 38:299–310, 1988.

[Pan08] Hao Pan. On consecutive happy numbers. *Journal of Number Theory*, 128(6):1646–1654, 2008.

[RR10] Adrian Röllin and Nathan Ross. Local limit theorems via landau-kolmogorov inequalities. *arXiv preprint arXiv:1011.3100*, 2010.

[Ruc88] Andrzej Ruciński. When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields*, 78(1):1–10, 1988.

[San95] Miklos Santha. On the monte carlo boolean decision tree complexity of read-once formulae. *Random Structures & Algorithms*, 6(1):75–87, 1995.

[Ste71] Charles Stein. Dependent random variables. 1971.