ADAPTIVE COVERING CODES IN THE Q-ARY HYPERCUBE

BY

DANIEL TIETZER

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science in Applied Mathematics
in the Graduate College of the
Illinois Institute of Technology

Approved _____
Advisor

Chicago, Illinois
December 2011

ACKNOWLEDGMENT

This thesis constitutes a straightforward improvement on the work of Joshua Cooper and Robert Ellis [Cooper and Ellis, 2010], who are responsible for the "clever tricks" which appear here. Their work derives in turn from clever tricks in work on adaptive codes in the binary hypercube due to Joel Spencer and Peter Winkler [Spencer, 1992,Spencer and Winkler, 1992]; on median bounds for particular random variables [Siegel, 2001] due to Alan Siegel; and on the variability of the "Rotor-Router Model" due to Joshua Cooper, Benjamin Doerr, Joel Spencer, and Gábor Tardos [Cooper et al., 2007]. Thanks are due especially to my M.S. advisor, Robert Ellis. In apparent violation of basic results in coding theory, he has been able to filter all of the errors from my transmissions, even when they were extremely long and more than half-full of errors. I continue to learn from and because of him, even when we are apart. He is also an exceptional human being, as evidenced by my continued good health. This work results from a collaboration with James Williamson, whose forthcoming thesis covers complementary topics [Williamson, 2012]. I must also thank Abe Sklar, a great friend and mentor and the first to warn me about my writing style; Hemanshu Kaul, for his indulgent (and thereby more challenging!) research work with me and for his insights as a member of my thesis committee; and my academic advisor Xiaofan Li, for smoothing my institutional way.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENT iii

LIST OF SYMBOLS vi

ABSTRACT ix

CHAPTER

1. THE PROBLEM 1

1.1. Games to Be Examined 1
1.2. Our Approach 6

2. OUTLINE OF RESULTS 8

2.1. The Main Result 8
2.2. Proof Structure 8

3. THE MODEL 10

3.1. The Linear Machine 10
3.2. Tracking the Excess Chips 11
3.3. Carole's Choices 13
3.4. The Liar Machine 13

4. AUXILIARY RESULTS 16

4.1. Pointwise Lower Bound on the Number of Chips at the End 16
4.2. Some Basic Calculus 23
4.3. Weighted Binomial Coefficients 24

5. DISCREPANCY BOUNDS ON INTERVALS 32

6. THE MAIN RESULT 39

6.1. Preliminaries 39
6.2. Proof of the Main Theorem 42

APPENDIX 45

A. GUIDE TO CITED RESULTS 46
A.1. Correcting Typos 47
A.2. Generalizing Corollary 2.3 48
A.3. Understanding Corollary 2.3 51

B. WRITING CONVENTIONS 52
C. STATEMENT ON COLLABORATION 55
BIBLIOGRAPHY 56

LIST OF SYMBOLS

Symbol	Definition	Defined	Discussed
$\binom{t}{i_1 \ldots i_2}_{q,a}$	sum of weighted binomial coefficients	p11	p11
a	size of an answer	p10	pp3-5
b	parameter for $X_\bullet,\ X_\circ$	p18	App. A
B	length of an interval or parameter for $X_\bullet,\ X_\circ$ [1]	pp18, 27, 30, 32	App. A
c_0	a constant	p27	p24
c_4	a constant	p30	p24
c_5	a constant	pp8, 32	p24
c^+, c', c''	specific numbers	pp27-30	–
D_B^s	difference of weighted binomial coefficients	p25	–
e	number of errors	p8	pp1-6
e_i	unit vector	–	p10
f	error rate or a specific probability density[2]	p10	pp5-6
$H_{q,a}$	entropy function	p23	p23
i, i_1, i_2, j, k, l	integers	variable	–
i', j', k'	real numbers	variable	–
j_{max}	specific number	p27	–

[1]The symbol "B" represents a parameter for $X_\bullet,\ X_\circ$ in Theorem 2, Lemma 3, and Appendix A; in the rest of the thesis, it represents the length of an interval.

[2]The symbol "f" is used as a probability density in appendix A; in the rest of the thesis, it represents a real "error rate."

k_0, k_1, k_2, k_3	specific numbers	pp32-39	–
$\mathcal{L}$	linear machine	p10	–
$\mathcal{L}_\chi$	liar machine (single question)	p14	–
$\mathcal{L}_{\mathbf{X},s},\ \mathcal{L}_{\mathbf{X}}$	liar machine (multiple questions)	p15	–
M	number of messages	pp8, 42	pp1-5, 16
mod q	modulo q	p11	p11
n	number of questions (total)	–	pp1-5
n'	element of $\{n, n_1\}$	p39	–
$n_1,\ n_2$	number of questions (in phases)	pp19, 39, 42	p16
n_3	specific small integer	p40, p41	–
o, O	Landau notation	–	p53
p	arithmetic combination of binomial coefficients	p40, p41	–
Pr	probability measure	variable	App. B
q	size of a question	p10	pp3-5
$r,\ R$	parameters for $X_\bullet,\ X_\circ$	p18	App. A
S	set	variable	–
s, t	numbers of questions	variable	–
$w_\bullet,\ w_\circ$	parameters for $X_\bullet,\ X_\circ$	p18	App. A

$X_\bullet,\ X_\circ$	hypergeometric random variables	p18	AppA
$y,\ y_0$	vectors indexed by $\mathbb{N}$	pp12,14,42-??	–
β	real number	pp24, 27	–
$\bar{\gamma},\ \bar{\bar{\delta}}$	functions satisfying long hypotheses	pp19, 41	–
$\bar{\gamma},\ \bar{\bar{\delta}}$	functions satisfying long hypotheses	pp19, 41	–
δ	specific real number	pp19, 41	–
ϵ	real number	pp37-38	–
$\kappa,\ \lambda$	specific real numbers	pp19, 39, 41	–
μ	mean of hypergeometric random variable	pp18, 19	p16
ν_x	chip-counting function	p11	–
ν_x^-	"inverse" chip-counting function	p11	–
ρ	parameter for $X_\bullet,\ X_\circ$	p18	App. A
σ	element of S	variable	–
Θ	Landau notation	–	p53
χ	single answer	variable	p13
$\mathbf{X}$	answer sequence	variable	p13
ω, Ω	Landau notation	–	p53

ABSTRACT

We investigate a generalized version of the Pathological Liar Game of Ellis and Yan. In our version, there are nonnegative integer parameters M, n, e, a, q with $0 < a < q$. The player (Carole) is equipped with M messages, each with an integral error count. Carole performs the following procedure n times in sequence: she numbers the messages in increasing order of error count, divides them into consecutive contiguous size-q blocks, and selects a size-a subset of each block. The messages she does not select have their error count increased by 1. Carole loses iff, after this process is complete, there is at least one message with error count $\leq e$. We allow n to approach infinity and suppose $e = \lfloor fn \rfloor$ for a fixed $f \in (0, 1)$. We establish an upper bound on the minimum M for which Carole always loses, as a function of n, by extending a technique of Cooper and Ellis; we develop a simple process that approximates the course of the game for any choice by Carole, and bound the difference between this approximating process and any state of the game achievable by Carole. Along the way we generalize several of the results of Cooper and Ellis in ways that suggest future application to similar problems in adaptive coding theory.

CHAPTER 1

THE PROBLEM

1.1 Games to Be Examined

We will be dealing with a set of related Games. All feature two opposing "players," Paul and Carole,[3] although Paul will often be constrained to a single choice and have a purely formal role.

Game 1 ("Liar Game," due to Berlekamp, Rényi and Ulam [Berlekamp, 1964, Rényi, 1961, Ulam, 1976]). *There are nonnegative integer parameters M, n, e. There is a set of M interchangeable messages. Carole picks one of the M messages, and Paul asks Carole n questions to try to guess which message Carole picked. For each question, Paul divides the messages into two disjoint subsets, and Carole tells him which subset contains her message. Just before each question, Paul remembers each of Carole's previous answers. However, Carole can give an incorrect answer up to e times. Paul wins iff, after the nth question, there is at most one message that Carole could have chosen if she answered according to the rules.*[4]

Observe:

- If we remove the requirement that Carole pick a message at the start of the game, nothing changes: if her answers are consistent with at least one of the messages, then she might as well have picked that message, and if her answers are inconsistent with all of the messages then she loses anyway. What matters

---

[3] following the trend set by Spencer and Winkler [Spencer and Winkler, 1992]

[4] In other words, for each $i \in \{1, ..., n\}$, let A_i be the subset selected by Carole in response to the ith question; Paul wins iff, after the nth question, $\left|\bigcup_{S \subseteq \{1, ..., n\},\ |S|=n-e} \bigcap_{i \in S} A_i\right| \leq 1$.

is whether Paul has narrowed down the possible messages to one by the end of the game.

- We can therefore formulate the game equivalently by removing the restriction on the number of incorrect answers Carole is allowed to give. Instead, we can assign an error count to each message, increasing whenever Carole selects a set not containing that message, and eliminate any message with error count $\geq e + 1$.

- If Paul knows the error count of each message just after each previous question, he can easily deduce Carole's answer at each previous round.

- At the beginning of a given question, Paul can win the game iff, by the end of his last question, he can increase the error count of all but one of the messages by at least $e + 1$ minus its current error count. Carole's precise answer history is therefore irrelevant.

We reformulate the liar game in light of these observations:

Game 1 ("Liar Game," reformulated [Rényi, 1961])**.** *There are nonnegative integer parameters M, n, e. There is a set of M interchangeable messages. Each message is given an error count of* 0*. Paul asks Carole n questions. For each question, Paul divides the messages into two disjoint subsets, and Carole picks one of them. Each message not contained in the set selected by Carole has its error count increased by one. Just before each question, Paul knows only the current error count of each message. Paul wins iff, after the nth question, there is at most one message with an error count of at most e.*

Given a strategy by Paul, an answer sequence by Carole can be expressed as

a binary string.[5] A strategy by Paul determines, for each message, a set of answer sequences by Carole placing $\leq e$ errors on that message, so a winning strategy by Paul corresponds to a packing of $\{0,1\}^n$ with these sets. Such packings are called adaptive packings.

We can also change the object of the Game:

Game 2 ("Pathological Liar Game," due to Ellis and Yan [Ellis and Yan, 2004])**.** *There are nonnegative integer parameters M, n, e. There is a set of M interchangeable messages. Each message is given an error count of* 0*. Paul asks Carole n questions. For each question, Paul divides the messages into two disjoint subsets, and Carole picks one of them. Each message not contained in the set selected by Carole has its error count increased by one. Just before each question, Paul knows only the current error count of each question. Paul wins iff, after the nth question, there is at least one message with an error count of at most e.*

In this Game, we can think of Paul as the leader of M hackers, trying to access a computer system controlled by Carole. There are two passwords, and Paul knows them both. Every minute, Paul picks which password each of his hackers should use to try to access the system; Carole immediately knows which hackers are using which passwords. She then picks one password to be the correct one for that minute. Once a hacker has used an incorrect password $e + 1$ times, he is caught. Carole wants to catch all the hackers.

Analogously to the previous observation, always-winning strategies for Paul correspond to covers of $\{0,1\}^n$ by particular types of sets. Such covers are known as adaptive covering codes.

[5]We can do this by indexing the disjoint subsets Paul selects at each question by 0 and 1, and letting the *ith* coordinate of the string be equal to the index of the set Carole selects at the *ith* question, for each positive integer $i \leq n$.

We can let Paul pick q subsets instead of just 1, and we can let Carole pick a subsets instead of just 1.

Game 3 ("q-ary a-pooled Pathological Liar Game," adapted from Du and Hwang [Du and Hwang, 2006] and Ellis and Yan [Ellis and Yan, 2004])**.** *There are nonnegative integer parameters M, n, e, q, a, with $0 < a < q$. There is a set of M interchangeable messages. Each message is given an error count of* 0*. Paul asks Carole n questions. For each question, Paul divides the messages into q disjoint subsets, and Carole picks a of them. Each message not contained in one of the sets selected by Carole has its error count increased by one. Just before each question, Paul knows only the current error count of each message. Paul wins iff, after the nth question, there is at least one message with an error count of at most e.*

Analogously to the previous observation, when $a = 1$, this corresponds to a covering of $\{0, \ ..., \ q-1\}^n$ by certain types of sets. In general, this corresponds to a covering of A^n, where A is the set of size-a subsets of $\{0, \ ..., \ q-1\}$.

We can play the same Game with more restrictions on the kinds of questions Paul can ask:

Game 4 ("q-ary a-pooled Alternating Pathological Liar Game," adapted from Du and Hwang [Du and Hwang, 2006] and Cooper and Ellis [Cooper and Ellis, 2010])**.** *There are nonnegative integer parameters M, n, e, q, a, with $0 < a < q$. There is a set of M interchangeable messages. Each message is given an error count of* 0*. Paul asks Carole n questions. For each question, the messages are placed in ascending order of error count (breaking ties arbitrarily). Carole picks a size-a subset $K \subset \{0, \ ..., \ q-1\}$. Each message whose position in the ordering is not congruent to an element of K* mod *q has its error count increased by one. Paul wins iff, after the nth question, there is at least one message with an error count of at most e.*

Paul's questions don't affect the error count except insofar as they restrict Carole's responses. We could instead restrict Carole's responses in some other way:

Game 5 ("q-block a-pooled Alternating Pathological Liar Game"). *There are nonnegative integer parameters M, n, e, q, a, with $0 < a < q$. There is a set of M interchangeable messages. Each message is given an error count of 0. Paul asks Carole n questions. For each question, the messages are placed in ascending order of error count (breaking ties arbitrarily) and broken into blocks, with the first q messages in the first block, the next q messages in the second block, and so on, until the last $M \mod q$ messages are placed in the last block. For each block, Carole picks a size-a subset K of $\{0, ..., q-1\}$, and each message in that block whose position in the ordering is not congruent to an element of $K \mod q$ has its error count increased by one. Paul wins iff, after the nth question, there is at least one message with an error count of at most e.*[6]

In this game, we can again think of Paul as the leader of M hackers, and of Carole as wanting to catch all the hackers. The hackers work out of $\lceil \frac{M}{q} \rceil$ offices, each of which can accommodate q hackers. This time, Carole's security is much better. Every day and for each office, she selects a of the hackers in that office to be detected, and any hacker who is detected $e+1$ times is caught. However, no two offices have equally comfy chairs, so at the end of each day Paul rewards the hackers by assigning some set of the best q hackers to the office with the comfiest chairs, a set of the next q best to the office with the second-comfiest chairs, and so on.[7]

---

[6]Observe that Carole selects K for the last block as though she is pretending that the last block has full size q.

[7]In this game, Paul doesn't do anything except place the hackers in different offices, and the only thing he knows about any hacker is the number of times that hacker has been detected. Also, Carole's handling of the least comfy office is a little different: she assumes that the least comfy office is full, selects a chairs in that office, and any hacker sitting in one of those chairs is detected.

Note that Game 4 is strictly harder for Paul than Game 3 (because Paul has fewer choices of constraint to place on Carole's answer at each round), and Game 5 is strictly harder for Paul than Game 4 (because Carole has more possible answers to choose from).

We will be interested in finding a sufficient condition for Paul to win Game 5, and thus also to win Games 4 and 3. We will then also have a sufficient condition for Paul to win Game 2, which is just Game 3 with $q = 2$ and $a = 1$.

1.2 Our Approach

As mentioned above, the only information that our strategies will need is the error count associated with each message at the beginning of the current question. We will call this information a "state," and express it as a vector indexed by the nonnegative integers,[8] with its ith coordinate equal to the number of messages with error count i.

For intuition's sake, we will think of each message as a "chip" occupying a nonnegative integer position. For each nonnegative integer i we place all chips with error count i in a vertical stack at position i. We think of the positions as arranged horizontally, with higher-numbered positions to the right of lower-numbered positions. When Carole picks a set of messages, the messages she didn't select all move one position to the right. In discussions, we will appeal to the notion of "fractional chips" when thinking about vectors with arbitrary real values.

We can then interpret Game 4 as Paul distributing chips into q piles, "dealer"-

[8]Cooper and Ellis consider functions with domain $\mathbb{Z}$ and finite support when considering Game 2. After t rounds have been played, the number of messages with j errors is equal to the number of chips at position -t+2j in their parametrization. This convention allows the distribution of chips to be approximately described by the pmf of a sum of t i.i.d. mean-zero indicator random variables.

style, starting with a leftmost chip and ending at a rightmost chip. The same interpretation works for Game 5, except that, after every q chips are distributed, the order of the piles can be changed by Carole. This process will tend to make all of the piles close to equal in size, so we would expect that every position i should be sending about a $\frac{q-a}{q}$ fraction of the chips to the $(i+1)th$ position after each question, and keeping the rest. This will not always accurately describe the Game, since the number of chips at each position will not always be divisible by q.

In Chapter 3, we will define the "linear machine," which describes the situation in which every position i sends precisely $\frac{q-a}{q}$ fractional chips to the $(i+1)th$ position and keeps the rest. We will also define the "liar machine," which describes the precise course of Game 5 given Carole's choices. The idea will be to describe the behavior of the linear machine,[9] and then show that it is not too far away from describing the liar machine. Williamson has shown that these bounds are optimal up to a constant factor [Williamson, 2012]. [10] We then derive lower bounds on the number of chips in certain regions of the linear machine to produce lower bounds on the analogous quantities in the liar machine.

Using this framework, fixing a real error rate $f \in (0, \frac{q-a}{q})$, and letting the number of questions n approach infinity, we will find an asymptotic lower bound on M, in terms of n, such that Paul can always win, subject to the condition that $e = \lfloor fn \rfloor$.

---

[9]The terms "liar machine" and "linear machine" are due to Cooper and Ellis [Cooper and Ellis, 2010]

[10]For each q, he finds a constant C_q such that, for a fixed integer N and a sufficiently large integer T, the total difference between the liar machine and the linear machine over the interval $\{\lceil \frac{T}{2} \rceil, \ ..., \ \lceil \frac{T}{2} \rceil + N\}$ is at least C_q times our upper bound. He treats the case where $a = 1$ and $\mathbf{X}_t$ takes a particular value for each nonnegative integer t, but his arguments are easily extensible to the general case. His argument uses results from our Sections 4.2 and 4.1.

CHAPTER 2

OUTLINE OF RESULTS

In Section 2.1, we describe the main result of this paper. In Section 2.2, we concisely outline the dependencies between sections.

2.1 The Main Result

Our goal is to prove the following:

Theorem 0. *Let q, a be positive integers with $a < q$ and f be an element of $(0, \frac{q-a}{q})$. Then there is a real constant c (depending on q and a) s.t., for all sufficiently large integers n, and letting $e := \lfloor fn \rfloor$, Paul can win Game 5, regardless of Carole's responses, as long as*

$$M \geq c \cdot \sqrt{\lfloor \log_q \ln(n) \rfloor} \cdot \frac{q^n}{\sum_{i=0}^{\lfloor fn \rfloor} \binom{n}{i} (q-a)^i a^{n-i}}.$$

This corresponds to a sufficient condition for Paul to win Game 5, and thus Games 4, 3, and 2 as well.[11] The special case of this result for Game 2 (and stated slightly differently) is due to Cooper and Ellis [Cooper and Ellis, 2010, Theorem 4].

2.2 Proof Structure

Definitions We define the Liar Machine, Linear Machine, and associated devices in Chapter 3.

Lemma 4

- In Appendix A, we make minor modifications to a result of Siegel [Siegel, 2001]. We state the modified result as Theorem 2.

[11] As far as we know, such conditions for Games 3, 4 and 5 have not previously been considered explicitly, though in our experience many arguments by other authors used to investigate Game 2 and its variations seem to lend themselves to generalization to Game 3.

- In Section 4.1, we use Theorem 2 to derive Lemma 4.

Lemma 12 and Fact 6

- In Section 4.2, we prove Facts 6 and 8.
- In Section 4.3, we use Facts 6 and 8 to prove Lemma 11.
- In Chapter 5, we use Lemma 11 to prove Lemma 12.

Theorem 16 (the main result)

- In Section 6.1, we prove Fact 13, and use Lemma 4 to prove Fact 15.
- In Section 6.2, we use Lemma 12 and Facts 6, 13, and 15 to prove Theorem 16.

Theorem 16 then immediately implies Theorem 0.

While we draw the ideas for our arguments from many sources, the only outside source a reader will require for our results is Siegel's paper on median bounds [Siegel, 2001], which we discuss in more detail (as it relates to our results) in Appendix A.

In Appendix B, we describe our conventions for thesis organization and for the use of common mathematical symbols.

CHAPTER 3

THE MODEL

We formalize what we discussed in Section 1.2.

Convention 1. *We will now fix positive integers q, a with $a < q$, and $f \in (0, \frac{q-a}{q})$.*

Convention 2. *Given a nonnegative integer i, we will use "e_i" to denote the ith unit vector.*[12]

3.1 The Linear Machine

The linear machine expresses the approximate development of the Game, and is independent of Carole's choices.[13]

Definition 3. *For a vector y indexed by the nonnegative integers and a nonnegative integer i, let*

$$\mathcal{L}(y)(i) := \frac{a}{q}y(i) + \frac{q-a}{q}y(i-1).$$

By the binomial theorem we can see the following:

Fact 1. *Given a vector y indexed by the nonnegative integers and nonnegative integers t and j, $\mathcal{L}^t(y)(j) = \sum_{j'=0}^{\infty} y(j')q^{-t}\binom{t}{j-j'}(q-a)^{j-j'}a^{t-j+j'}$.*

---

[12] i.e., the vector, indexed by the nonnegative integers, whose *ith* coordinate is 1 and whose other coordinates are 0

[13] Given a game state described by a vector y, it happens that $\mathcal{L}(y)$ is equal to the pointwise average of all possible states that Carole can produce from y in a single round. We can also think of $\mathcal{L}$ as describing what happens to "most" of the messages from one round to the next: when there are lots of messages, most of the length-q blocks of messages consist completely of messages with all equal error counts, and each such length-q block sends exactly $q-1$ of its messages to the right by one position, regardless of Carole's selected messages for that block.

If we start with 1 chip at position 0 and no chips elsewhere, and iterate the linear machine t times, the number of fractional chips between positions i_1 and i_2 is equal to $q^{-t} \cdot \sum_{i=i_1}^{i_2} \binom{t}{i}(q-a)^i a^{t-i}$; we want to express this compactly.

Definition 4. *For integers t, i_1, i_2, let*

$$\binom{t}{i_1 \ldots i_2}_{q,a} := \sum_{i=i_1}^{i_2} \binom{t}{i}(q-a)^i a^{t-i}.$$

3.2 Tracking the Excess Chips

Consider a fixed state with its chips ordered and divided into blocks according to the rules of Game 5. Then for any block all of whose chips occupy the same position, Carole's choices for that block have no effect on the Game, since exactly $q-a$ chips from that block will always move to the right. We can therefore restrict our attention to the "excess chips"; we want to keep track of these excess chips at each position.

In Definition 5, for a state y, we will define functions ν_y and ν_y^-. For each i, $\nu_y(i)+1$ counts the total number of chips at positions $\leq i$ contained in any block spanning multiple positions. For each nonnegative integer k, $\nu_y^-(k)$ simply gives the position of the kth such chip.

Definition 5.

- *For an integer k, let $(k \mod q)$ be the unique element of $\{0, \ldots q-1\}$ congruent to $k \mod q$.* [14]
- *For an integer vector y indexed by the nonnegative integers and a nonnegative integer i, let $excess_y(i) :=$*

[14] Note that we are not treating $(k \mod q)$ as an element of $\mathbb{Z}_q$, but of $\mathbb{Z}$.

$$\begin{cases} y(i) & \text{if } y(i) \leq (-\sum_{j=0}^{i-1} y(j) \mod q) \\ (-\sum_{j=0}^{i-1} y(j) \mod q) + (\sum_{j=0}^{i} y(j) \mod q) & \text{if } (-\sum_{j=0}^{i-1} y(j) \mod q) < y(i). \end{cases}$$

- *For an integer vector y indexed by the nonnegative integers and a nonnegative integer i, let $\nu_y(i) := -1 + \sum_{j=0}^{i} excess_y(j)$.*
- *For an integer vector y indexed by the nonnegative integers, let $\nu_y(\infty) := -1 + \sum_{j=0}^{\infty} excess_y(j)$.*
- *For an integer vector y indexed by the nonnegative integers and a nonnegative integer k, let $\nu_y^-(k) := \min\{i \ : \nu_y(i) \geq k\}$.*

Example 1. *Let $q := 7$ and $y := (7, 0, 11, 7, 0, 2, 8, 0, 0, 3, 0, 0, ...)$.*

- *The first block consists entirely of messages with 0 errors, and the second block consists entirely of messages with 2 errors.*
- *The third block consists of 4 messages with 2 errors and 3 messages with 3 errors, so the messages in this block are "counted" by the ν_y function.*
- *The fourth block consists of 4 messages with 3 errors, 2 messages with 5 errors, and 1 message with 6 errors, so the messages in this block are "counted" by the ν_y function.*
- *the fifth block consists of of 7 messages with 6 errors, so the messages in this block are not "counted" by the ν_y function.*
- *The sixth (and last) block consists of 3 message with 9 errors, so the messages in this block are "counted" by the ν_y function (we think of the last block as spanning multiple blocks since it is never "completed").*
- *We thus have the following: $\nu_y(0) = -1,\ \nu_y(1) = -1,\ \nu_y(2) = 3,\ \nu_y(3) = 10,\ \nu_y(5) = 12,\ \nu_y(6) = 13,\ \nu_y(7) = 13,\ \nu_y(8) = 13,\ \nu_y(9) = 16,\ \nu_y(\infty) = 16$.*

- *Thus also:* $\nu_y^-(0) = 2,\ \nu_y^-(3) = 2,\ \nu_y^-(4) = 3,\ \nu_y^-(10) = 3,\ \nu_y^-(11) = 5,\ \nu_y^-(13) = 6, \nu_y^-(14) = 9$.

3.3 Carole's Choices

We want to describe Carole's choices in an efficient way. We will do this using two observations, given a state y:

- We can think about Carole's answers in terms of the deviation of the resulting chip movements from those prescribed by the linear machine. When Carole moves a chip at position i to the right, she is transferring $\frac{a}{q}$ fractional chips from the ith position to the $(i+1)th$. When she fixes a chip at position i, she is transferring $\frac{a-q}{q}$ fractional chips from the ith position to the $(i+1)th$.
- Since we don't need to consider blocks all of whose chips occupy the same position, we can express Carole's answer as a selection of a residue classes mod q for each block of q consecutive integers in $\{0,\ ...,\ \nu_y(\infty)\}$.

Now we can describe Carole's responses:

Definition 6. *Given a vector y indexed by the nonnegative integers, a function χ is a legal response to y if χ is the restriction to $\{0,\ ...,\ \nu_y(\infty)\}$ of a function $\psi : \{0,\ ...,\ q(\lfloor\frac{\nu_y(\infty)}{q}\rfloor + 1) - 1\} \to \{\frac{a-q}{q}, \frac{a}{q}\}$ s.t. the following holds: for each $k \in \{0,\ ...,\ \lfloor\frac{\nu_y(\infty)}{q}\rfloor + 1\}$, the set $\{kq,\ ...,\ (k+1)q - 1\}$ contains exactly a elements of $\psi^{-1}(\frac{a-q}{q})$.*

Convention 7. *We will use the symbol χ to represent choices for Carole, and the symbol $\mathbf{X}$ to represent a finite-length sequence of such choices.*

3.4 The Liar Machine

Now we can define the liar machine, which expresses the exact development of the Game given legal choices by Carole.

Definition 8. *For an integer vector y indexed by the nonnegative integers and a real-valued function χ defined on $\{0,\ ...,\ \nu_y(\infty)\}$, let*

$$\mathcal{L}_\chi(y) := \mathcal{L}(y) + \sum_{k=0}^{\nu_y(\infty)} \chi(k) \cdot (e_{\nu_y^-(k)+1} - e_{\nu_y^-(k)}).$$

For each k and each chip with position $\nu_y^-(k)$, Carole transfers $\chi(k)$ fractional chips to position $\nu_y^-(k) + 1$. [15]

Example 2. *Let $q := 7$ and $y := (7, 0, 11, 7, 0, 2, 8, 0, 0, 3, 0, 0, ...)$ as in Example 1, and $a = 5$. The "blocks" are tracked in Example 1. Suppose that Carole wants to pick a legal response χ to y; then she has the following options:*

- *$\chi(k) = \frac{a-q}{q}$ for exactly 5 elements k of $\{0,\ ...,\ 6\}$ and exactly 5 elements of $\{7,\ ...,\ 13\}$, and $\chi(k) = \frac{a}{q}$ for exactly 2 elements k of each of those two sets;*
- *$\chi(k) = \frac{a-q}{q}$ for between 1 and 3 elements k of $\{14, 15, 16\}$, and $\chi(k) = \frac{a-q}{q}$ for the remaining elements k of $\{14, 15, 16\}$.*

Suppose that she selects χ so that $\chi^{-1}(\frac{a}{q}) = \{5, 6, 12, 13\}$; then

- $\mathcal{L}(y) = (\frac{7a}{q}, \frac{7(q-a)}{q}, \frac{11a}{q}, \frac{-4a+11q}{q}, \frac{7(q-a)}{q}, \frac{7q-5a}{q}, \frac{2q+6a}{q}, \frac{8(q-a)}{q}, 0, \frac{3a}{q}, \frac{3(q-a)}{q}, 0, ...)$
- $\mathcal{L}_\chi(y) = \mathcal{L}(y) + (0, 0, -4\frac{a-q}{q}, 4\frac{a-q}{q} - \frac{a-q}{q} - 2\frac{a}{q} - 4\frac{a-q}{q}, \frac{a-q}{q} + 2\frac{a}{q} + 4\frac{a-q}{q}, -\frac{a-q}{q} - \frac{a}{q}, \frac{a-q}{q} + \frac{a}{q} - \frac{a}{q}, \frac{a}{q}, 0, -3\frac{a-q}{q}, 3\frac{a-q}{q}, 0, ...)$, *so*
- $\mathcal{L}_\chi(y) = (\frac{7a}{q}, \frac{7(q-a)}{q}, \frac{7a+4q}{q}, \frac{7q}{q}, \frac{2q}{q}, \frac{8q-7a}{q}, \frac{q+7a}{q}, \frac{8q-7a}{q}, 0, \frac{3q}{q}, 0, 0, ...) = (5, 2, 9, 7, 2, 3, 6, 2, 0, 3, 0, 0)$;

[15] Note that, for any function $\chi : \{0,\ ...,\ \nu_y(\infty)\} \to \{\frac{a-q}{q}, \frac{a}{q}\}$ and any integer vector y, $\mathcal{L}_\chi(y)$ is an integer vector.

this last vector corresponds to "fixing" the first 5 *chips, "shifting" the next* 2*, "fixing" the next* 5*, "shifting" the next* 2*, and so on, until the last* 3 *are "fixed."*

Finally, we describe the evolution of the Game over multiple rounds:

Definition 9. *For an integer vector* y *indexed by the nonnegative integers, a positive integer* T*, and a sequence of real-valued functions* $\{\mathbf{X}_t : t \in \{1,\ ...,\ T\}$ *defined on* $\{0,\ ...,\ \nu_y(\infty)\}$*,*

- *let* $\mathcal{L}_{\mathbf{X},0}(y) := y$*;*
- *for each* $t \in \{1,\ ...,\ T\}$*, let* $\mathcal{L}_{\mathbf{X},\ t}(y) := \mathcal{L}_{\mathbf{X}_t}(\mathcal{L}_{\mathbf{X},\ t-1}(y))$*;*
- *if* $\mathbf{X}_t$ *is a legal response to* $\mathcal{L}_{\mathbf{X},\ t-1}(y)$ *for each* $t \in \{1,\ ...,\ T\}$*, say that* $\mathbf{X}$ *is a legal response sequence to* y*;*
- *let* $\mathcal{L}_{\mathbf{X}}(y) := \mathcal{L}_{\mathbf{X},\ T}(y)$*.*

CHAPTER 4

AUXILIARY RESULTS

This is the Section to which a reader familiar with the arguments of Cooper and Ellis [Cooper and Ellis, 2010] should devote the most attention; we generalize their Lemmas 6 (without lower bounds) and 14. In particular, the arguments of Section 4.1 are probably the most complicated in this thesis.

Ultimately, we want two things: an interval-wise lower bound on the number of chips in the linear machine, and an interval-wise upper bound on the difference between the liar machine and linear machine. Section 4.1 will give us a pointwise lower bound on the number of chips in the linear machine, extended to an interval-wise bound in Section 6.1. Section 4.2 will give us some necessary calculus results. Section 4.3 will give us a pointwise upper bound on the difference between the liar machine and the linear machine, extended to an interval-wise bound in Chapter 5.

Sections 4.1 and 4.2 are self-contained, while Section 4.3 invokes a result from 4.2. Two results from 4.2 appear in the proof of the main result (16).

4.1 Pointwise Lower Bound on the Number of Chips at the End

In Chapter 6, we will analyze the linear machine in two successive phases: a long phase of length n_1 and a short phase of length n_2, with $n_1 + n_2 = n$. We start with $M = m \cdot \dfrac{q^n}{\binom{n}{\lfloor fn_1 \rfloor \ldots \lfloor fn \rfloor}_{q,a}}$ chips at position 0 (with error count 0) and no chips elsewhere. At the end of the Game, the following expression describes the contribution to the number of fractional chips at position i by chips which, at time n_1, occupied positions between $\lfloor fn_1 \rfloor$ and $\lfloor fn \rfloor$, inclusive:

$$m \cdot \frac{q^n}{\binom{n}{\lfloor fn_1 \rfloor \ldots \lfloor fn \rfloor}_{q,a}} \cdot \sum_{j=\lfloor fn_1 \rfloor}^{\min(\lfloor fn \rfloor, i)} \frac{\binom{n_1}{j} \cdot (q-a)^j a^{n_1-j}}{q^{n_1}} \frac{\binom{n_2}{i-j} \cdot (q-a)^{i-j} a^{n_2-(i-j)}}{q^{n_2}}.$$

Ignoring a factor of m, cancelling powers of q, and pulling out a factor of

$(q-a)^i a^{n-i}$, we obtain the expression $\frac{(q-a)^i a^{n-i}}{\binom{n}{\lfloor fn_1 \rfloor \ldots \lfloor fn \rfloor}_{q,a}} \cdot \sum_{j=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \binom{n_1}{j}\binom{n_2}{i-j}$. We wish bound this number above a constant times $\frac{(q-a)^i a^{n-i}}{\binom{n}{\lfloor fn_1 \rfloor \ldots \lfloor fn \rfloor}_{q,a}} \cdot \binom{n}{i}$ as $n \to \infty$; we want lots of fractional chips to remain at the end of the Game.

A result implying this appears in the paper on the binary case by Cooper and Ellis [Cooper and Ellis, 2010, Proposition 14], but we give a slightly more explicit proof of a slightly more general result here (Lemma 4). We could actually do a bit better if we accepted additional tedium.[16]

We will prove this result in (approximately) the following steps:

(i) Show that the sum $\sum_{j=\lceil \frac{n_1}{n} \cdot i \rceil}^{i-1} \frac{\binom{n_1}{j}\binom{n_2}{i-j}}{\binom{n}{i}}$ is small.

(ii) Note that $\frac{n_1}{n} \cdot i$ is the mean of a hypergeometric random variable (counting the number of red balls selected lying in a pool when a total of $i \leq \lfloor fn \rfloor$ balls are selected from a pool of n balls of which n_1 are red), so that $\sum_{j=\lceil \frac{n_1}{n} \cdot i \rceil}^{\lfloor fn \rfloor} \frac{\binom{n_1}{j}\binom{n_2}{i-j}}{\binom{n}{i}} \geq \frac{1}{2}$.

(iii) Note that our sum can be interpreted as $(q-a)^i$ times a tail of the corresponding hypergeometric distribution.

(iv) Use the fact that the mean and median of a hypergeometric random variable differ by at most 1 to show that *tail* $\geq 1/2 -$ (*terms above the median*), where (*terms above the median*) is given by the small quantity in (i).

Steps (ii) and (iii) are pretty easy, but step (iv) follows from a clever result appearing in a paper by Siegel, which we invoke as Lemma 2 below.

[16] We have reason to believe that this generalization may be of use in analyzing more general versions of Game 1.

The use of the hypergeometric distribution seems natural if we interpret each position i as the result of i "rightward moves," and interpret the number of fractional chips as the sum of the probabilities of all possible allocations of i rightward moves between the first and second phases (appropriately scaled). Indeed, subject to the strategy we want to use, we can say that the denominator $\binom{n}{\lfloor fn_1 \rfloor \ldots \lfloor fn \rfloor}_{q,a}$ in our starting chip quantity is in a sense the biggest we can get away with, since it "matches up" with our $(q-a)^i \cdot \sum_{j=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \binom{n_1}{j}\binom{n_2}{i-j}$ terms.

The following Lemma is effectively proven by Siegel [Siegel, 2001]:[17]

Lemma 2. *Let an urn contain R red balls and B black balls. Suppose each red ball has weight $w_\circ$, and each black ball has weight $w_\bullet$. Suppose that the balls are selected one-by-one without replacement where each as yet unselected ball is given a probability of being selected at the next round that equals its current fraction of the total weight of all unselected balls. Suppose r and b are nonnegative real numbers with integer sum satisfying $r = R(1 - e^{-w_\circ \rho})$ and $b = B(1 - e^{-w_\bullet \rho})$, for some fixed $\rho > 0$. Let $r + b$ balls be drawn from the urn as prescribed. Let $X_\circ$ be the number of red balls selected by this random process, and let $X_\bullet$ be the number of black, so that $X_\circ + X_\bullet = r + b$. Then $\Pr\{X_\circ \geq \lfloor r \rfloor\} > \frac{1}{2}$ and $\Pr\{X_\bullet \geq \lfloor b \rfloor\} > \frac{1}{2}$.*

Following Cooper and Ellis [Cooper and Ellis, 2010, Corollary 13], we note an immediate consequence of the above Lemma:

Lemma 3. *If μ is the mean of a hypergeometric random variable X, then $\Pr\{X \leq \lceil \mu \rceil\} \geq \frac{1}{2}$.*

Proof. Consider such a random variable X, and suppose that $\Pr\{X = j\} = \frac{\binom{R}{j}\cdot\binom{B}{i-j}}{\binom{R+B}{i}}$ for some nonnegative integers R, B, i. Let r and b be nonnegative real numbers

[17] For a discussion of this, see Appendix A.2.

s.t. $\frac{r}{R} = \frac{b}{B}$ and $r + b = i$, and let $w := -\ln(1 - \frac{r}{R})$. Let $\rho := 1$. Now define $X_\circ$ as in the hypothesis of Lemma 2, with $w_\bullet = w = w_\circ$ and R, r, B, b, ρ as just defined. Then clearly $\Pr\{X_\circ \geq \lfloor r \rfloor\} > \frac{1}{2}$, so it suffices to show that X and $X_\circ$ have the same probability mass function; this will now be done.

Let S be the set of injective mappings from $\{1, ..., r+b\}$ to the set of $R+B$ balls. For each $\sigma \in S$, let $\Pr_\sigma$ be the probability that, for each $k \in \{1, ..., r+b\}$, the *kth*-selected ball is equal to $s(k)$. For each $j \in \mathbb{Z}^+$, let S_j be the set of mappings in S whose ranges contain precisely j red balls; then clearly $|S_j| = (r+b)! \cdot \binom{R}{j} \cdot \binom{B}{r+b-j}$. Thus it suffices to show that, for each $\sigma \in S$, $\Pr_\sigma = \frac{1}{(r+b)! \cdot \binom{R+B}{r+b}} = \frac{(R+B-r-b)!}{(R+B)!}$. But for any $\sigma \in S$, $\Pr_\sigma$ is proportional to $\prod_{k=1}^{r+b} \frac{w}{(R+B)w-(k-1)w} = \frac{(R+B-r-b)!}{(R+B)!}$. Since $|S| = \frac{(R+B)!}{(R+B-r-b)!}$, $\Pr_\sigma$ is thus equal to $\frac{(R+B-r-b)!}{(R+B)!}$.

□

We are now ready to state our result for this Section (beware that the "r" in our proof is not a direct analogue of the "r" above):

Lemma 4. *Let κ be an element of $(f, 1)$, and let $\lambda := \frac{12}{f\kappa(\kappa - f)}$. Suppose that, for each positive integer n, nonnegative integers n_1 and n_2 are defined s.t. $n_1 + n_2 = n$, $\frac{n_1}{n} \geq \kappa$ for n sufficiently large, and $n_2 \geq 2\lambda$ for n sufficiently large. Let $\bar{\gamma} : \mathbb{Z}^+ \to \mathbb{R}^+$ be s.t. $\bar{\gamma}(n) < \frac{1}{\lambda}$ for n sufficiently large. Let $\bar{\delta}(n) := \sqrt{\min(\bar{\gamma}(n) \cdot n_2, \log_{1-\lambda \cdot \bar{\gamma}(n)} \frac{1}{2})}$ for each positive integer n s.t. $\bar{\gamma}(n) < \frac{1}{\lambda}$.*

Then for n sufficiently large, and for any nonnegative integer r s.t. $fn_2 - \bar{\delta}(n) + \frac{n}{n_1} \leq r \leq n - \lfloor fn_1 \rfloor$,

$$\sum_{j=\lfloor fn_1 \rfloor}^{\lfloor fn_1 \rfloor + r} \frac{\binom{n_1}{j}\binom{n_2}{\lfloor fn_1 \rfloor + r - j}}{\binom{n}{\lfloor fn_1 \rfloor + r}} \geq \frac{1}{2} - \max\left(\frac{8}{\bar{\delta}(n)}, \frac{8}{\bar{\delta}(n)^2}\right). \quad [18]$$

[18] Think of $\bar{\gamma}(n) = \frac{1}{\sqrt{n_2}}$.

Proof. Suppose that

(i) n is large enough that $n-1 \geq n_2 \geq 2\lambda$, $\frac{n_1}{n} \geq \kappa$, $n \geq \frac{6}{f\kappa}$, and $\bar{\gamma}(n) \leq \frac{1}{\lambda}$. [19]

Clearly,

$$\sum_{j=\lfloor fn_1 \rfloor}^{\lfloor fn_1 \rfloor + r} \frac{\binom{n_1}{j}\binom{n_2}{\lfloor fn_1 \rfloor + r - j}}{\binom{n}{\lfloor fn_1 \rfloor + r}} = \sum_{j=0}^{r} \frac{\binom{n_1}{\lfloor fn_1 \rfloor + r - j}\binom{n_2}{j}}{\binom{n}{\lfloor fn_1 \rfloor + r}}.$$

Let $p(j) := \frac{\binom{n_1}{\lfloor fn_1 \rfloor + r - j}\binom{n_2}{j}}{\binom{n}{\lfloor fn_1 \rfloor + r}}$ for each nonnegative integer j. Then p, as a function of j, is the probability mass function of a hypergeometric random variable with mean $\frac{n_2}{n}(\lfloor fn_1 \rfloor + r)$; denote this quantity by μ. Then by Lemma 3 we have that

$$\sum_{j=0}^{\lceil \mu \rceil} \frac{\binom{n_1}{\lfloor fn_1 \rfloor + r - j}\binom{n_2}{j}}{\binom{n}{\lfloor fn_1 \rfloor + r}} \geq \frac{1}{2}. \tag{4.1}$$

Note next that

$$\mu - r = \frac{n_1}{n}(fn_2 - r) - \frac{n_2}{n}(fn_1 - \lfloor fn_1 \rfloor). \tag{4.2}$$

Thus in particular, if $r \geq fn_2$ then $r \geq \mu$, so since r is an integer $r \geq \lceil \mu \rceil$ and thus the result follows immediately by (4.1). Thus suppose that

(ii) $r \leq fn_2$;

then $-\frac{n_2}{n}(fn_1 - \lfloor fn_1 \rfloor) \leq \lceil \mu \rceil - r \leq (\lceil \mu \rceil - \mu) + \frac{n_1}{n}(fn_2 - r) \leq (\lceil \mu \rceil - \mu) - 1 + \frac{n_1}{n}\bar{\delta}(n) \leq \bar{\delta}(n)$;[20] in particular,

$$\lceil \mu \rceil - r \leq \bar{\delta}(n). \tag{4.3}$$

---

[19] Note that this hypothesis is independent of r.

[20] This is where the addend $\frac{n}{n_1}$ in the hypothesis comes in.

Now for any positive integer $k = \mu + \delta$ *s.t.* $1 \geq \delta \geq -\bar{\gamma}(n) \cdot n_2$, we have that

$$\begin{aligned}\frac{p(k)}{p(k-1)} =& \frac{\binom{n_1}{\lfloor fn_1 \rfloor + r - k}\binom{n_2}{k}}{\binom{n_1}{\lfloor fn_1 \rfloor + r - (k-1)}\binom{n_2}{k-1}} = \\ & \frac{(n_2 - k + 1)((\lfloor fn_1 \rfloor + r) - k + 1)}{k(n - n_2 - (\lfloor fn_1 \rfloor + r) + k)} = \\ & 1 + \frac{[n_2(\lfloor fn_1 \rfloor + r) - (k-1)(n_2 + \lfloor fn_1 \rfloor + r) + (k-1)^2]}{k(n - n_2 - (\lfloor fn_1 \rfloor + r) + k)} - \\ & \frac{[k^2 + kn - k(n_2 + \lfloor fn_1 \rfloor + r)]}{k(n - n_2 - (\lfloor fn_1 \rfloor + r) + k)} = \\ & 1 + \frac{n_2(\lfloor fn_1 \rfloor + r) + (n_2 + \lfloor fn_1 \rfloor + r) - 2k + 1 - kn}{k(n - n_2 - (\lfloor fn_1 \rfloor + r) + k)} = \\ & 1 + \frac{(n_2 + \lfloor fn_1 \rfloor + r) - 2k + 1 - n\delta}{k(n - n_2 - (\lfloor fn_1 \rfloor + r) + k)}.\end{aligned}$$

Now since

- $n - 1 \geq n_2 \geq 1$, $\frac{n_1}{n} \geq \kappa$, $n \geq \frac{6}{f\kappa}$, and $\bar{\gamma}(n) \leq \frac{f\kappa}{2}$ (by (i));
- thus also $k \leq \mu + 1 \leq n_2 + 1 \leq n$; and
- $r \leq fn_2$ (by (ii)),

we have that

- $|(n_2 + \lfloor fn_1 \rfloor + r) - 2k + 1 - n\delta| \leq 4 \cdot \max(1, |\delta|) \cdot n$;
- $n - n_2 - (\lfloor fn_1 \rfloor + r) + k \geq n_1 - fn \geq (\kappa - f)n$;
- $k = \mu + \delta \geq \frac{\kappa}{2} fn_2 - \frac{n_2}{n} \geq \frac{\kappa}{3} fn_2$,

so

$$\left| \frac{(n_2 + \lfloor fn_1 \rfloor + r) - 2k + 1 - n\delta}{k(n - n_2 - (\lfloor fn_1 \rfloor + r) + k)} \right| \leq \lambda \cdot \frac{\max(1, |\delta|) \cdot n}{n \cdot n_2} \leq \lambda \cdot \max(\frac{1}{n_2}, \bar{\gamma}(n));$$

thus in particular,

$$\left| \frac{p(k)}{p(k-1)} - 1 \right| \leq \lambda \cdot \max(\frac{1}{n_2}, \bar{\gamma}(n)). \tag{4.4}$$

Now by (4.4),

$$1 \geq \sum_{k=\lceil \mu - \bar{\delta}(n)^2 \rceil}^{\lfloor \mu+1 \rfloor} p(k) \geq$$

$$p(\lfloor \mu + 1 \rfloor) \cdot \bar{\delta}(n)^2 \cdot \left(1 - \lambda \cdot \max(\frac{1}{n_2}, \bar{\gamma}(n))\right)^{\bar{\delta}(n)^2} \geq$$

$$\bar{\delta}(n)^2 \cdot \frac{1}{2} \cdot p(\lfloor \mu + 1 \rfloor).^{21}$$

Thus in particular

$$p(\lfloor \mu + 1 \rfloor) \leq \frac{2}{\bar{\delta}(n)^2}. \tag{4.5}$$

Now also, since $(1 + \lambda \cdot \bar{\gamma}(n))(1 - \lambda \cdot \bar{\gamma}(n)) \leq 1$ and $1 - \lambda \cdot \bar{\gamma}(n) > 0$ (by (i)),

$$p(k) \leq p(\lfloor \mu + 1 \rfloor) \cdot (1 + \lambda \cdot \bar{\gamma}(n))^{\bar{\delta}(n)^2} \leq p(\lfloor \mu + 1 \rfloor) \cdot (1 - \lambda \cdot \bar{\gamma}(n))^{-\bar{\delta}(n)^2} \leq 2p(\lfloor \mu + 1 \rfloor)$$

whenever $\lceil \mu - \bar{\delta}(n)^2 \rceil \leq k \leq \lfloor \mu + 1 \rfloor$; in particular we have that

$$p(k) \leq 2p(\lfloor \mu + 1 \rfloor) \tag{4.6}$$

whenever $\lceil \mu - \bar{\delta}(n)^2 \rceil \leq k \leq \lfloor \mu + 1 \rfloor$.

By (4.6) and (4.5), we have that

$$\sum_{k=\lceil \mu - \bar{\delta}(n) \rceil}^{\lfloor \mu+1 \rfloor} p(k) \leq (\bar{\delta}(n) + 1) \cdot 2 \cdot \frac{2}{\bar{\delta}(n)^2} \leq \max\left(\frac{8}{\bar{\delta}(n)}, \frac{8}{\bar{\delta}(n)^2}\right). \tag{4.7}$$

Now

$$\sum_{j=0}^{r} \frac{\binom{n_1}{\lfloor f n_1 \rfloor + r - j}\binom{n_2}{j}}{\binom{n}{\lfloor f n_1 \rfloor + r}} = \sum_{j=0}^{\lfloor \mu+1 \rfloor} p(j) - \sum_{j=r+1}^{\lfloor \mu+1 \rfloor} p(j). \tag{4.8}$$

---

[21] This is true since, by (i),

- $\bar{\delta}(n)^2 \leq \log_{1-\lambda\cdot\bar{\gamma}(n)} \frac{1}{2} \in \mathbb{R}^+$ and,
- since $n_2 \geq 2\lambda$, we have that, if $\log_{1-\frac{\lambda}{n_2}} \frac{1}{2} \leq \log_{1-\lambda\cdot\bar{\gamma}(n)} \frac{1}{2}$, then $\frac{1}{n_2} \geq \bar{\gamma}(n)$ and thus $\bar{\gamma}(n) \cdot n_2 \leq 1 \leq \log_{1-\frac{\lambda}{n_2}} \frac{1}{2}$.

Now $r+1 \geq \lceil \mu - \bar{\delta}(n) \rceil$ by (4.3), so (4.7) and (4.1) dictate that the right-hand side of (4.8) is bounded below by

$$\sum_{j=0}^{\lfloor \mu+1 \rfloor} p(j) - \frac{8}{\bar{\delta}(n)} \geq \frac{1}{2} - \max\left(\frac{8}{\bar{\delta}(n)}, \frac{8}{\bar{\delta}(n)^2}\right),$$

as desired. □

4.2 Some Basic Calculus

Suppose we iterate the linear machine n times starting with 1 chip at position 0 and no chips elsewhere. Then given $g \in (0,1)$ *s.t.* $gn \in \mathbb{Z}$, the number of fractional chips at position gn is given by

$$q^{-n}\binom{n}{gn}(q-a)^{gn}a^{(1-g)n},$$

which by the Stirling approximation is equal to

$$q^{-n}\frac{\Theta(1)}{\sqrt{n}\sqrt{g}\sqrt{1-g}} \cdot \left(\frac{1}{g^g(1-g)^{(1-g)}}\right)^n \left((q-a)^g\ a^{1-g}\right)^n,$$

where the $\Theta(1)$ function is uniform in g. We are interested in the exponential "growth rate" of this expression for a fixed g, and wish to express it as -1 plus a function of g.

Definition 10. *Given $g \in (0,1)$, let*

$$H_{q,a}(g) := g\log_q(\frac{1}{g}) + (1-g)\log_q(\frac{1}{1-g}) + g\log_q(q-a) + (1-g)\log_q(a).$$

This generalizes the common entropy function, so that

$$q^{-n}\binom{n}{gn}(q-a)^{gn}a^{(1-g)n}$$

is equal to

$$q^{-n}\frac{\Theta(1)}{\sqrt{n}\sqrt{g}\sqrt{1-g}} \cdot \left(\frac{1}{g^g(1-g)^{(1-g)}}\right)^n \left((q-a)^g\ a^{1-g}\right)^n,$$

where again the $\Theta(1)$ function is uniform in g.

We will want a few simple properties of this function:

Fact 5. *For* $g \in (0,1)$, $H'_{q,a}(g) = \log_q(\frac{1-g}{g}) + \log_q(q-a) - \log_q(a)$.

Fact 6. $H_{q,a}$ *is strictly increasing on* $(0, \frac{q-a}{q})$ *and strictly decreasing on* $(\frac{q-a}{q}, 1)$.

Fact 7. $H_{q,a}(\frac{q-a}{q}) = 1$.

Fact 8. *Given* $\beta \in \mathbb{R}$, $\lim_{n\to\infty} n \cdot (H_{q,a}(\frac{q-a}{q} + \beta n^{-1/2}) - 1) \in \mathbb{R}$.

Proof. By Facts 5 and 7 and L'Hôpital's rule,

$$\lim_{n\to\infty} n \cdot (H_{q,a}(\frac{q-a}{q} + \beta n^{-1/2}) - 1) =$$
$$\lim_{n\to\infty} n^2 \cdot \frac{\beta n^{-3/2}}{2} \cdot [\log_q(\frac{a - q\beta n^{-1/2}}{q-a+q\beta n^{-1/2}}) + \log_q(\frac{q-a}{a})] =$$
$$\lim_{n\to\infty} \frac{\beta}{2} n^{1/2} \cdot [\log_q(\frac{a - q\beta n^{-1/2}}{q-a+q\beta n^{-1/2}}) - \log_q(\frac{a}{q-a}];$$

again by L'Hôpital's rule, this expression is equal to

$$\lim_{n\to\infty} \beta^2 \cdot (-\log_q(e)) \cdot \frac{q-a+q\beta n^{-1/2}}{a-q\beta n^{-1/2}} \cdot \left(\frac{q}{q-a+q\beta n^{-1/2}}\right)^2 =$$
$$\lim_{n\to\infty} \beta^2 \cdot (-\log_q(e)) \cdot \frac{q^2}{(a-q\beta n^{-1/2})(q-a+q\beta n^{-1/2})} =$$
$$\beta^2 \cdot (-\log_q(e)) \cdot \frac{q^2}{a(q-a)}.$$

□

4.3 Weighted Binomial Coefficients

In Chapter 5, we will want to show that the gap between the liar machine and linear machine[22] does not grow too large. Consider a chip-distribution function

[22] of Chapter 3

y, a position i, and one of the "excess chips" at position i. After one step of the liar machine, this chip gives rise to an appropriate (positive or negative) number of "fractional discrepancy chips" at positions i and $i+1$. As we iterate the liar machine, these fractional chips propagate via the linear machine. After iterating the linear machine a specified number of times, the total number of these discrepancy chips hitting a particular position j can be expressed as a difference of weighted binomial coefficients. We will show that each of these differences is small, and also that the differences originating from distinct excess chips tend largely to cancel each other out when added together.

Remember that we consider q and a to be fixed, so that the bounds we will derive may depend on q and a. [23]

Definition 11. *Given $s \in \mathbb{Z}^+,\ B, j \in \mathbb{Z}$, let*

$$D_B^s(j) := \binom{s}{j-B}(q-a)^{j-B}a^{s-j+B} - \binom{s}{j}a^{s-j}(q-a)^j.$$

4.3.1 Bimodality of Distance between Two Coefficients. The first step is to show that $D_B^s(j)$ is bimodal in j, so that a sum of alternating-sign coefficients times D_B^s terms is bounded above in absolute value by a constant times the maximum value of D_B^s.

Lemma 9. *Given $s \in \mathbb{Z}^+, B \in \mathbb{Z}$, D_B^s is bimodal in j, i.e. either*

(a) there exist integers j_1, j_2 s.t., for any integer j, $D_B^s(j) - D_B^s(j-1) < 0$ if and only if $j_1 < j < j_2$, or

(b) there exist integers j_1, j_2 s.t., for any integer j, $D_B^s(j) - D_B^s(j-1) > 0$ if and only if $j_1 < j < j_2$.

[23] This dependence would not turn out to be too bad if we considered it. Our bounds are independent of f, though our later results will not depend on this.

Proof. Since $D_B^s(j) = -D_{(-B)}^s(j-B)$, we may suppose without loss of generality that $B \geq 0$. We will show that (a) holds in this case.

Given $j \in \mathbb{Z}$,

$$D_B^s(j-1) - D_B^s(j) =$$

$$\binom{s}{j-B-1}(q-a)^{j-B-1}a^{s-j+B+1} - \binom{s}{j-B}(q-a)^{j-B}a^{s-j+B} -$$
$$\binom{s}{j-1}(q-a)^{j-1}a^{s-j+1} + \binom{s}{j}(q-a)^j a^{s-j} =$$

$$\binom{s+1}{j-B}(q-a)^{j-B-1}a^{s-j+B+1}\left[\frac{j-B}{s+1} - \frac{q-a}{a}\frac{s+1-j+B}{s+1}\right] -$$
$$\binom{s+1}{j}(q-a)^{j-1}a^{s-j+1}\left[\frac{j}{s+1} - \frac{q-a}{a}\frac{s+1-j}{s+1}\right];$$

this quantity has the same sign as

$$\binom{s+1}{j-B}[q(j-B) - (q-a)(s+1)] - \binom{s+1}{j}\left(\frac{q-a}{a}\right)^B [qj - (q-a)(s+1)]. \quad (4.9)$$

Now

- whenever $\frac{q-a}{q}(s+1) \leq j \leq \frac{q-a}{q}(s+1) + B$, (4.9) is negative;
- whenever $j < \frac{q-a}{q}(s+1)$ and $j < B$, (4.9) is nonnegative;
- whenever $j > s+1$ and $j > \frac{q-a}{q}(s+1) + B$, (4.9) is nonnegative;

thus it suffices to show that the sign of (4.9), as a function of an integer j, is nonincreasing on $\{B,\ ...,\ \lceil\frac{q-a}{q}(s+1)\rceil - 1\}$ and nondecreasing on $\{\lfloor\frac{q-a}{q}(s+1) + B\rfloor + 1,\ ...,\ s+1\}$. This will now be done.

Now for $j \in \{B,\ ...,\ s+1\}$, the sign of (4.9) is equal to the sign of

$$[q(j-B) - (q-a)(s+1)] - \left[\prod_{i=0}^{B-1}\frac{s+2-j+i}{j-i}\frac{q-a}{a}\right][qj - (q-a)(s+1)];$$

also,

(i) $[q(j-B)-(q-a)(s+1)]$ and $[qj-(q-a)(s+1)]$ are negative for $j \in \{B,\ ...,\ \lceil \frac{q-a}{q}(s+1)\rceil - 1\}$;

(ii) $[q(j-B)-(q-a)(s+1)]$ and $[qj-(q-a)(s+1)]$ are positive for $j \in \{\lfloor \frac{q-a}{q}(s+1)+B\rfloor + 1,\ ...,\ s+1\}$;

(iii) as a function of j, $\prod_{i=0}^{B-1} \frac{s+2-j+i}{j-i}\frac{q-a}{a}$ is nonnegative and nonincreasing on $\{B,\ ...,\ s+1\}$;

(iv) as a function of j, $\frac{q(j-B)-(q-a)(s+1)}{qj-(q-a)(s+1)}$ is nondecreasing on $\{B,\ ...,\ \lceil \frac{q-a}{q}(s+1)\rceil - 1\}$ and $\{\lfloor \frac{q-a}{q}(s+1)+B\rfloor + 1,\ ...,\ s+1\}$.

Now by (i), (ii), and (iii), it suffices to show that the sign of $\frac{q(j-B)-(q-a)(s+1)}{qj-(q-a)(s+1)} - \left[\prod_{i=0}^{B-1} \frac{s+2-j+i}{j-i}\frac{q-a}{a}\right]$, as a function of j, is nondecreasing on $\{B,\ ...,\ \lceil \frac{q-a}{q}(s+1)\rceil - 1\}$ and $\{\lfloor \frac{q-a}{q}(s+1)+B\rfloor+1,\ ...,\ s\}$. But by (iii) and (iv), this quantity is equal to a sum of two expressions which, as functions of j, are nondecreasing on $\{B,\ ...,\ \lceil \frac{q-a}{q}(s+1)\rceil - 1\}$ and $\{\lfloor \frac{q-a}{q}(s+1)+B\rfloor + 1,\ ...,\ s+1\}$. □

4.3.2 Distance between Two Close Coefficients.

Now we bound D_1^s.

Lemma 10. *There is a real constant c_0 s.t., given $s \in \mathbb{Z}^+, \max\{|D_1^s(j)| : j \in \mathbb{Z}\} \leq q^s \cdot \frac{c_0}{s}$.*

Proof. For $s \in \mathbb{Z}^+,\ j \in \mathbb{Z}$,

$$D_1^s(j) = (q-a)^{j-1}a^{s-j+1}\left[\binom{s}{j-1} - \binom{s}{j}\frac{q-a}{a}\right] =$$
$$(q-a)^{j-1}a^{s-j+1}\binom{s+1}{j}\left[\frac{j}{s+1} - \frac{q-a}{a}\frac{s+1-j}{s+1}\right] =$$
$$\frac{1}{a(s+1)}(q-a)^{j-1}a^{s-j+1}\binom{s+1}{j}[qj-(q-a)(s+1)], \tag{4.10}$$

so

$$D_1^s(j+1) - D_1^s(j) =$$

$$\frac{1}{a(s+1)}(q-a)^j a^{s-j}\binom{s+1}{j+1}[qj-(q-a)(s+1)+q]- \\ \frac{1}{a(s+1)}(q-a)^{j-1}a^{s-j+1}\binom{s+1}{j}[qj-(q-a)(s+1)] =$$

$$\frac{1}{a(s+1)}(q-a)^{j-1}a^{s-j}((q-a)\binom{s+1}{j+1}[qj-(q-a)(s+1)+q]- \\ a\binom{s+1}{j}[qj-(q-a)(s+1)]); \quad (4.11)$$

also, if $j \in \{0,\ ...,\ s+1\}$, then the right-hand side of equation 4.11 has the same sign as

$$(q-a)(s+1-j)[qj-(q-a)(s+1)+q]-a(j+1)[qj-(q-a)(s+1)]. \quad (4.12)$$

Now for $s \in \mathbb{Z}^+$, the right-hand side of equation (4.10) is equal to 0 for any integer j *s.t.* $j \leq -1$ or $j \geq s+2$, so $\max\{|D_1^s(j)| : j \in \mathbb{Z}\}$ is well-defined; letting j_{max} be an integer *s.t.* $|D_1^s(j_{max})| = \max\{|D_1^s(j)| : j \in \mathbb{Z}\}$, we have the following:

- If $j_{max} \leq 0$ or $j_{max} \geq s+1$, then $|D_1^s(j_{max})| \leq \frac{1}{a(s+1)}(q-a)^s(s+1)[q(s+1)+(q-a)(s+1)]$, which for s large enough is $\leq \frac{1}{s}q^s$.

- If $1 \leq j_{max} \leq s$ and neither $D_1^s(j_{max}+1) = D_1^s(j_{max})$ nor $D_1^s(j_{max}) - D_1^s(j_{max}-1)$, then the sign of $D_1^s(j_{max}+1) - D_1^s(j_{max})$ is equal to -1 times the sign of $D_1^s(j_{max}) - D_1^s(j_{max}-1)$, and thus the sign of expression (4.12) with $j = j_{max}$ is equal to -1 times the sign of expression (4.12) with $j = j_{max}-1$. Since (4.12), as a function of j, is continuous on $[0, s+1]$, it follows that $j_{max} = \lceil j' \rceil$ or $j_{max} = \lfloor j' \rfloor$ for some $j' \in [0, s+1]$ *s.t.* (4.12) $= 0$ for $j = j'$.

- If $1 \leq j_{max} \leq s$ and either $D_1^s(j_{max}+1) = D_1^s(j_{max})$ or $D_1^s(j_{max}) - D_1^s(j_{max}-1)$, then $D_1^s(j_{max}) = D_1^s(j')$ for some $j' \in \{0,\ ...,\ s\}$ *s.t.* $D_1^s(j'+1) - D_1^s(j') = 0$ and thus *s.t.* (4.12) $= 0$ for $j = j'$.

Thus it suffices to show that there is a real constant c_0 *s.t.* the following holds: for any $s \in \mathbb{Z}^+$ and any $j' \in [0, s+1]$ *s.t.* $(q-a)(s+1-j')[qj' - (q-a)(s+1)+q] - a[qj' - (q-a)(s+1)] = 0$, we have that $|D_1^s(\lfloor j' \rfloor)| \leq q^s \cdot \frac{c_0}{s} \geq |D_1^s(\lceil j' \rceil)|$. This will now be done.

Consider such a j'; then expression (4.12) is equal to

$$\begin{aligned}(q-a)(s+1-j)[qj-(q-a)(s+1)+q] - a(j+1)[qj-(q-a)(s+1)] = \\ -qj[qj-(q-a)(s+1)] - a[qj-(q-a)(s+1)] + \\ (q-a)(s+1)[qj-(q-a)(s+1)] + q(q-a)(s+1-j) = \\ -q^2j^2 + 2q(q-a)(s+1)j - q^2j + (q+a)(q-a)(s+1) - (q-a)^2(s+1)^2,\end{aligned} \tag{4.13}$$

so

$$j' = \frac{2(q-a)(s+1)-q}{2q} \pm \frac{1}{2}\sqrt{q^2 - 4q(q-a)(s+1) + 4(q+a)(q-a)(s+1)} = \frac{(q-a)(s+1)}{q} - \frac{1 \pm \sqrt{q^2 + 4a(q-a)(s+1)}}{2};$$

thus by (4.10), for $j \in \{\lceil j' \rceil, \lfloor j' \rfloor\}$,

$$|D_1^s(j)| \leq \frac{1}{a(s+1)}(q-a)^{j-1}a^{s-j+1}\binom{s+1}{j} q \cdot \frac{3+\sqrt{q^2+4a(q-a)(s+1)}}{2}.$$

Let $\beta \in \mathbb{R}^+$ be *s.t.*, for s large enough, $\frac{3+\sqrt{q^2+4a(q-a)(s+1)}}{2} \leq \beta(s+1)^{1/2}$.

Let c^+ be *s.t.*, for s large enough and $j \in \{\lfloor \frac{q-a-1/2}{q}s \rfloor, \ ..., \lceil \frac{q-a+1/2}{q}s \rceil\}$, $(q-a)^j a^{s-j+1}\binom{s+1}{j} \leq \frac{c^+}{\sqrt{s+1}} q^{(s+1)H_{q,a}(\frac{j}{s+1})}$; this is possible by the Stirling approximation and Definition 10.

Then for s large enough and $j' \in [0, s+1]$ *s.t.*

$$(q-a)(s+1-j')[qj' - (q-a)(s+1)+q] - a[qj' - (q-a)(s+1)] = 0,$$

both $|D_1^s(\lceil j' \rceil)|$ and $|D_1^s(\lfloor j' \rfloor)|$ are less than or equal to

$$\frac{qc^+\beta}{a(s+1)} \cdot q^{(s+1)H_{q,a}(\frac{j'}{s+1})};$$

now by Fact 6, for s large enough this is less than or equal to

$$\frac{qc^+\beta}{a(s+1)} \cdot q^{(s+1)H_{q,a}(\frac{q-a}{q} \pm \beta(s+1)^{-1/2})}$$

and, by Fact 8, there is a real constant c' *s.t.*, for s large enough, this is less than or equal to $\frac{c'}{s} q^{s+1}$. Thus, for s large enough, we have that $|D_1^s(j_{max})| \leq \frac{q^2 c' c^+ \beta}{a} \cdot \frac{q^s}{s}$.

□

Since each term D_B^s is just a sum of D_1^s terms, we can now bound D_B^s.

Lemma 11. [24] *There is a real constant c_4 s.t., given $s \in \mathbb{Z}^+,\ B \in \mathbb{Z}$,*

(a) $\max\{|D_B^s(j)| \ : \ j \in \mathbb{Z}\} \leq q^s \cdot \frac{c_4 B}{s}$;

(b) $\max\{|D_B^s(j)| \ : \ j \in \mathbb{Z}\} \leq q^s \cdot \frac{c_4}{\sqrt{s}}$.

Proof. For $s \in \mathbb{Z}^+, B, j' \in \mathbb{Z}$,

- $|D_B^s(j')| = \left|\sum_{k=0}^{B-1} D_1^s(j'-k)\right|$; by Lemma 10, this is less than or equal to $\leq q^s \cdot \frac{c_0 B}{s}$, proving (a);
- also, $|D_B^s(j')| \leq \max\{\binom{s}{j}(q-a)^j a^{s-j} \ : \ j \in \mathbb{Z}\}$.

Now for $j \in \mathbb{Z}$ *s.t.* $j \leq \frac{1}{2}\frac{q-a}{q}$ or $j \geq 2\frac{q-a}{q}$, we have by Fact 6 that $\binom{s}{j}(q-a)^j a^{s-j} \leq q^s \frac{1}{\sqrt{s}}$ for $s \in \mathbb{Z}^+$ large enough. By the Stirling approximation and Facts 6 and 8, there is a real constant c'' *s.t.*, for $s \in \mathbb{Z}^+$ large enough and $j \in \mathbb{Z}$ *s.t.* $\frac{1}{2}\frac{q-a}{q} \leq$

[24] We name our constant "c_4" to harmonize notation with Cooper and Ellis [Cooper and Ellis, 2010, Lemma 6] .

$j \leq 2\frac{q-a}{q}$, $\binom{s}{j}(q-a)^j a^{s-j} \leq q^s \frac{c''}{\sqrt{s}}$. Thus $\max\{\binom{s}{j}(q-a)^j a^{s-j} \ : \ j \in \mathbb{Z}\} \leq q^s \frac{\max(1,c'')}{\sqrt{s}}$, proving (b). $\square$

CHAPTER 5

DISCREPANCY BOUNDS ON INTERVALS

We are now ready to bound the difference between the liar machine and the linear machine[25] in Lemma 12 by applying the results of Section 4.3. This generalizes Theorems 2 and 3 of Cooper and Ellis [Cooper and Ellis, 2010]. We conclude the chapter by discussing a further generalization.

Lemma 12. *There is a real constant c_5 s.t. the following holds: Let y be a vector indexed by the nonnegative integers, let t be an integer greater than 1, let i, B be nonnegative integers with $B \leq i$, and let $\mathbf{X}$ be a legal response sequence to y with length t. Then*

1. *if $B + 1 \geq \frac{\sqrt{t}}{2}$, then $\left|\sum_{j=i-B}^{i}[\mathcal{L}_{\mathbf{X}}(y)(j) - \mathcal{L}^t(y)(j)]\right| \leq c_5 \cdot \frac{\sqrt{t}}{2}$, and*
2. *if $B+1 \leq \frac{\sqrt{t}}{2}$, then $\left|\sum_{j=i-B}^{i}[\mathcal{L}_{\mathbf{X}}(y)(j) - \mathcal{L}^t(y)(j)]\right| \leq c_5 \cdot (B+1) \cdot \ln(t/(B+1)^2)$.*

Proof. For each $s \in \{0, \ ..., \ t-1\}$, let

$$\nu_s(\infty) := \nu_{\mathcal{L}_{\mathbf{X}},\ t-1-s(y)}(\infty),$$

and let

$$\nu_s^- := \nu^-_{\mathcal{L}_{\mathbf{X}},\ t-1-s(y)}.$$

Let c_4 be as in Lemma 11.

---

[25] of Chapter 3

By Definitions 8 and 9, we have that

$$\mathcal{L}_{\mathbf{X}}(y) =$$
$$\mathcal{L}^t(y) + \sum_{s=0}^{t-1} \mathcal{L}^s \left(\sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot (e_{\nu_s^-(k)+1} - e_{\nu_s^-(k)}) \right) =$$
$$\mathcal{L}^t(y) + \sum_{s=0}^{t-1} \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot \mathcal{L}^s(e_{\nu_s^-(k)+1} - e_{\nu_s^-(k)}),$$

so by Fact 1 and Definition 11,

$$\sum_{j=i-B}^{i} [\mathcal{L}_{\mathbf{X}}(y)(j) - \mathcal{L}^t(y)(j)] =$$
$$\sum_{j=i-B}^{i} \sum_{s=0}^{t-1} \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot \mathcal{L}^s(e_{\nu_s^-(k)+1} - e_{\nu_s^-(k)})(j) =$$
$$\sum_{s=0}^{t-1} \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot \sum_{j=i-B}^{i} \mathcal{L}^s(e_{\nu_s^-(k)+1} - e_{\nu_s^-(k)})(j) =$$

$$\sum_{s=0}^{t-1} \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot \sum_{j=i-B}^{i} \sum_{j'=0}^{\infty} q^{-s} \binom{s}{j-j'} (q-a)^{j-j'} a^{s-j+j'} \cdot$$
$$[e_{\nu_s^-(k)+1}(j') - e_{\nu_s^-(k)}(j')] =$$

$$\sum_{s=0}^{t-1} \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot q^{-s} \cdot \sum_{j=i-B}^{i} [\binom{s}{j-\nu_s^-(k)-1} (q-a)^{j-\nu_s^-(k)-1} a^{s-j+\nu_s^-(k)+1} -$$
$$\binom{s}{j-\nu_s^-(k)} (q-a)^{j-\nu_s^-(k)} a^{s-j+\nu_s^-(k)}] =$$

$$\sum_{s=0}^{t-1} \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot q^{-s} \cdot [\binom{s}{i-B-\nu_s^-(k)-1} (q-a)^{i-B-\nu_s^-(k)-1} a^{s-i+B+\nu_s^-(k)+1} -$$
$$\binom{s}{i-\nu_s^-(k)} (q-a)^{i-\nu_s^-(k)} a^{s-i+\nu_s^-(k)}] =$$

$$\sum_{s=0}^{t-1} q^{-s} \cdot \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot [D_{B+1}^s(i-\nu_s^-(k))];$$

thus in particular

$$\left|\sum_{j=i-B}^{i}[\mathcal{L}_{\mathbf{X}}(j)-\mathcal{L}^t(j)]\right| \quad\leq\quad \sum_{s=0}^{t-1}q^{-s}\ \cdot\ \left|\sum_{k=0}^{\nu_s(\infty)}\mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right|. \quad (5.1)$$

Note also that

- for any $k\in\{0,\ ...,\ \nu_0(\infty)\},\ D^0_{B+1}(i-\nu_0^-(k))=$
$$\begin{cases}\pm 1 & \text{if } i=\nu_0^-(k) \text{ or } i=\nu_0^-(k)+B+1\\ 0 & \text{otherwise;}\end{cases}$$
- for any integer j, there are at most $2q-2$ integers k *s.t.* $\nu_s^-(k)=j$;

and thus

$$q^{-0}\cdot\left|\sum_{k=0}^{\nu_0(\infty)}\mathbf{X}_{t-1-0}(k)\cdot D^0_{B+1}(i-\nu_0^-(k))\right|\leq$$

$$\max(q-a,a)\cdot(4q-4)\leq 4q^2.\ ^{26} \quad (5.2)$$

Now if $B+1\leq\frac{\sqrt{t}}{2}$, then $4q^2\leq 4q^2\cdot(B+1)\ln(\frac{t}{(B+1)^2})$, so by (5.1) and (5.2) it suffices to show that

$$\sum_{s=1}^{t-1}q^{-s}\cdot\left|\sum_{k=0}^{\nu_s(\infty)}\mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right|\leq$$

$$4q^2\cdot c_4\cdot\min((B+1)\ln(\frac{t}{(B+1)^2}),2\sqrt{t}).$$

To do this, it suffices to show that, for $s\in\{1,\ ...,\ t-1\}$,

$$q^{-s}\ \cdot\ \left|\sum_{k=0}^{\nu_s(\infty)}\mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right|\ \leq\ 4q^2\ \cdot\ c_4\ \cdot\ \min\left(\frac{B+1}{s},\frac{1}{\sqrt{s}}\right), \quad (5.3)$$

[26] We could actually do better here by being more careful.

since

$$\sum_{s=1}^{([B+1]/2)^2-1} \frac{1}{\sqrt{s}} + \sum_{s=([B+1]/2)^2}^{t-1} \frac{B+1}{s} \leq (B+1)\left[1+\ln(\frac{4t}{(B+1)^2})\right] \text{ if } B+1 \leq \frac{\sqrt{t}}{2}$$

and

$$\sum_{s=1}^{t-1} \frac{1}{\sqrt{s}} \leq 2\sqrt{t}.$$

To prove (5.3), it suffices to show that, for $s \in \{1, \ ..., \ t-1\}$,

$$q^{-s} \cdot \left| \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot D^s_{B+1}(i-\nu_s^-(k)) \right| \leq 4q^2 \cdot \max\{\left|D^s_{B+1}(j)\right| : j \in \mathbb{Z}\}, \qquad (5.4)$$

by Lemma 11. This will now be done.

Consider a fixed $s \in \mathbb{Z}^+$. Then

- $i-\nu_s^-(k)$ is monotone on $\{0, \ ..., \ \nu_s(\infty)\}$ as a function of k by Definition 5;
- $D^s_{B+1}(j)$ is bimodal on $\mathbb{Z}$ as a function of j by Lemma 9;
- therefore $D^s_{B+1}(i-\nu_s^-(k))$ is bimodal on $\{0, \ ..., \ \nu_s(\infty)\}$ as a function of k.[27]

Let k_1, k_2 be *s.t.*, as a function of k, $D^s_{B+1}(i-\nu_s^-(k))$ is monotone on $\{0, \ ..., \ k_1q-1\}$, $\{(k_1+1)q, \ ..., \ k_2q-1\}$, and $\{(k_2+1)q, \ ..., \ \nu_s(\infty)\}$. Let $k_0 := -1$, and let $k_3 := \lfloor \frac{\nu_s(\infty)}{q} \rfloor + 1$; then

$$\left| \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot D^s_{B+1}(i-\nu_s^-(k)) \right| \leq$$

$$\left| \sum_{k=0}^{k_3q-1} \mathbf{X}_{t-1-s}(k) \cdot D^s_{B+1}(i-\nu_s^-(k)) \right| +$$

$$(q-1)\max(q-a,a) \cdot \max\{\left|D^s_{B+1}(j)\right| : j \in \mathbb{Z}\} \leq$$

[27] We use the term "bimodal" here in the sense of Lemma 9.

$$\left|\sum_{k=0}^{k_1q-1} \mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right| +$$

$$q\max(q-a,a)\cdot\max\{\left|D^s_{B+1}(j)\right| : j\in\mathbb{Z}\}+$$

$$\left|\sum_{k=(k_1+1)q}^{k_2q-1} \mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right| +$$

$$q\max(q-a,a)\cdot\max\{\left|D^s_{B+1}(j)\right| : j\in\mathbb{Z}\}+$$

$$\left|\sum_{k=(k_2+1)q}^{k_3q-1} \mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right| +$$

$$q\max(q-a,a)\cdot\max\{\left|D^s_{B+1}(j)\right| : j\in\mathbb{Z}\} \leq$$

$$\left|\sum_{k=(k_0+1)q}^{k_1q-1} \mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right| +$$

$$\left|\sum_{k=(k_1+1)q}^{k_2q-1} \mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right| +$$

$$\left|\sum_{k=(k_2+1)q}^{k_3q-1} \mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))\right| +$$

$$3q^2\cdot\max\{\left|D^s_{B+1}(j)\right| : j\in\mathbb{Z}\}.\ ^{28} \quad (5.5)$$

Now since

- $D^s_{B+1}(i-\nu_s^-(k))$ is monotone on $\{(k_0+1)q,\ ...,\ k_1q-1\}$, $\{(k_1+1)q,\ ...,\ k_2q-1\}$, and $\{(k_2+1)q,\ ...,\ k_3q-1\}$ as a function of k, and
- for each $b\in\{0,1,2\}$, $\sum_{k=(k_b+1)q}^{k_{b+1}q}\mathbf{X}_{t-1-s}(k)\cdot D^s_{B+1}(i-\nu_s^-(k))$ is between

$$\sum_{k=k_b+1}^{k_{b+1}}[\frac{q-a}{q}\sum_{l=0}^{a-1} D^s_{B+1}(i-\nu_s^-(kq+l))+$$

$$\frac{a}{q}\sum_{l=a}^{q-1} D^s_{B+1}(i-\nu_s^-(kq+l))]$$

[28]Note that k_3 could be infinite, and the argument still works.

and

$$\sum_{k=k_b+1}^{k_{b+1}} [\frac{q-a}{q} \sum_{l=q-a}^{q-1} D^s_{B+1}(i-\nu_s^-(kq+l)) + \frac{a}{q} \sum_{l=0}^{q-a-1} D^s_{B+1}(i-\nu_s^-(kq+l))]$$

(since $\mathbf{X}$ is a legal response sequence to y),

it follows that expression (5.5) is less than or equal to

$$\leq (3\max(q-a,a)+3q^2)\max\{\left|D^s_{B+1}(j)\right| : j \in \mathbb{Z}\} \leq$$

$$4q^2 \cdot \max\{\left|D^s_{B+1}(j)\right| : j \in \mathbb{Z}\},$$

proving (5.4).

□

The above argument is based on the idea that, for nonnegative integers i, s with $s < t$, the contribution of the $(t-s)th$ state to $\sum_{j=i-B}^{i}[\mathcal{L}_{\mathbf{X}}(j) - \mathcal{L}^t(j)]$ is equal to $q^{-s} \cdot \sum_{k=0}^{\nu_s(\infty)} \mathbf{X}_{t-1-s}(k) \cdot D^s_{B+1}(i-\nu_s^-(k))$ (as in (5.1)). In the above proof, we use the bimodality of D_{B+1} and the "alternating" structure of the $\mathbf{X}_{t-1-s}$ values to bound $\left|\sum_{j=i-B}^{i}[\mathcal{L}_{\mathbf{X}}(j) - \mathcal{L}^t(j)]\right|$ below a constant times $q^{-s} \cdot \max\{\left|D^s_{B+1}(j)\right| : j \in \mathbb{Z}\}$, but since $D_{B+1}(i-\nu_s^-(k)) = 0$ whenever $i - \nu_s^-(k) \leq -1$, we could just as easily use the bound

$$q^{-s}\max\{\left|D^s_{B+1}(j)\right| : j \in \mathbb{Z},\ 0 \leq j \leq i\}. \tag{5.6}$$

We briefly sketch what would happen for small i.

If $j \leq i \leq (\frac{q-a}{q} - \epsilon)s$ for some constant ϵ, then $|D^s_{B+1}(j)|$ is bounded above by a constant times $\frac{1}{\sqrt{s}} q^{H_{q,a}(\frac{q-a}{q}-\epsilon)\cdot s}$, by the Stirling approximation and Fact 6 (this constant is actually uniform in B because of our constraint $B \geq 0$). By Fact 7, $H_{q,a}(\frac{q-a}{q} - \epsilon) < 1$, so that $q^{-s} \cdot D^s_{B+1}(j)$ shrinks exponentially quickly as $s \to \infty$. Thus, for a given i,

- the total contribution from the $(t-s)th$ states with $s \geq i/\left(\frac{q-a}{q} - \epsilon\right)$ are bounded above in absolute value by a constant (which can be taken to be decreasing in i), and
- the total contribution from the $(t-s)th$ states with $s < i/\left(\frac{q-a}{q} - \epsilon\right)$ are bounded above in absolute value by a constant times

$$\begin{cases} c_5 \cdot \frac{\sqrt{i}}{2} & \text{if } B+1 \geq \frac{\sqrt{i}}{2} \\ \leq c_5 \cdot (B+1) \cdot \ln(i/(B+1)^2) & \text{if } B+1 \leq \frac{\sqrt{i}}{2} \end{cases} \tag{5.7}$$

  as in the above lemma,

so the total contribution is bounded above by (5.3).

Such an observation may be of use in versions of Game 1, since in this case it is the messages with low error count that worry us the most.

CHAPTER 6

THE MAIN RESULT

In Section 6.1, we will have two objectives. Fact 13 says that the pointwise number of chips in the linear machine is much greater than the difference between the liar machine and the linear machine. Fact 15 extends Lemma 4 to give an interval-wise lower bound on the number of chips in the linear machine. We will then prove our main result in Section 6.2.

Section 6.1 generalizes Lemmas 9,10,11, and 15 of Cooper and Ellis; Section 6.2 generalizes their Theorem 26 [Cooper and Ellis, 2010].

Recall throughout that $f < \frac{q-a}{q}$, and that we make use of the entropy function from Definition 10.

6.1 Preliminaries

Fact 13. *Let κ be an element of $(0,1)$. Suppose that, for each positive integer n, nonnegative integers n_1 and n_2 are defined s.t. $n_1+n_2 = n$ and $\frac{n_1}{n} \geq \kappa$ for n sufficiently large. Then*

$$\frac{q^n}{\binom{n}{\lfloor fn \rfloor}(q-a)^{\lfloor fn \rfloor}a^{n-\lfloor fn \rfloor}} \frac{\binom{n}{\lfloor fn_1 \rfloor}(q-a)^{\lfloor fn_1 \rfloor}a^{n_1-\lfloor fn_1 \rfloor}}{q^{n_1}} = \Theta(q^{[1-H_{q,a}(f)]\cdot n_2}).$$

Proof. By the Stirling approximation, for each $n' \in \{n, n_1\}$, $\frac{q^{n'}}{\binom{n'}{\lfloor fn' \rfloor}(q-a)^{\lfloor fn' \rfloor}a^{n'-\lfloor fn' \rfloor}}$ is between

$$\Theta\left(q^{n'}\frac{1}{\sqrt{n'}}\frac{(fn'-1)^{fn'-1}(n'-fn'+1)^{n'-fn'+1}}{n'^{n'}(q-a)^{fn'-1}a^{n'-fn'+1}}\right) \text{ and}$$
$$\Theta\left(q^{n'}\frac{1}{\sqrt{n'}}\frac{(fn'+1)^{fn'+1}(n'-fn'-1)^{n'-fn'-1}}{n'^{n'}(q-a)^{fn'+1}a^{n'-fn'-1}}\right);$$

now each of the above two expressions is equal to

$$\Theta\left(q^{n'}\frac{1}{\sqrt{n'}}\frac{(fn')^{fn'}(n'-fn')^{n'-fn'}}{n'^{n'}(q-a)^{fn'}a^{n'-fn'}}\right) = \Theta\left(\frac{1}{\sqrt{n'}}q^{[1-H_{q,a}(f)]\cdot n'}\right),$$

so

$$\frac{q^n}{\binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor}}\frac{\binom{n}{\lfloor fn_1\rfloor}(q-a)^{\lfloor fn_1\rfloor}a^{n_1-\lfloor fn_1\rfloor}}{q^{n_1}} =$$

$$\frac{\sqrt{n_1}}{\sqrt{n}}\cdot\frac{\Theta(q^{[1-H_{q,a}(f)]\cdot n})}{\Theta(q^{[1-H_{q,a}(f)]\cdot n_1})} = \Theta(q^{[1-H_{q,a}(f)]\cdot n_2}).$$

□

Fact 14.

(a) For each positive integer n, $\binom{n}{0...\lfloor fn\rfloor}_{q,a} \leq \frac{(1-f)(q-a)}{(1-f)(q-a)-fa}\cdot\binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor}$.

(b) Suppose that, for each positive integer n, a nonnegative integer n_3 is defined s.t. $n_3 = \omega(1)$. Then $\binom{n}{0...\lfloor fn\rfloor}_{q,a} = [1-o(1)]\cdot\binom{n}{\lfloor fn\rfloor-n_3...\lfloor fn\rfloor}_{q,a}$.

Proof. For $k \in \{0,\ ...,\ \lfloor fn\rfloor\}$,

$$\frac{\binom{n}{\lfloor fn\rfloor-k}}{\binom{n}{\lfloor fn\rfloor}} = \frac{\lfloor fn\rfloor!(n-\lfloor fn\rfloor)!)}{(\lfloor fn\rfloor-k)!(n-\lfloor fn\rfloor+k)!} \leq \frac{\lfloor fn\rfloor^k}{(n-\lfloor fn\rfloor+1)^k} \leq \frac{(fn)^k}{(n-fn)^k},$$

so since $f < \frac{q-a}{q}$,

$$\frac{\binom{n}{\lfloor fn\rfloor-k}(q-a)^{\lfloor fn\rfloor-k}a^{n-\lfloor fn\rfloor+k}}{\binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor}} \leq \frac{[fa]^k}{[(1-f)(q-a)]^k} < 1; \tag{6.1}$$

thus in particular

$$\binom{n}{0...\lfloor fn\rfloor}_{q,a} = \sum_{j=0}^{\lfloor fn\rfloor}\binom{n}{j}(q-a)^j a^{n-j} \leq$$

$$\left[\binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor}\right]\cdot\sum_{k=0}^{\lfloor fn\rfloor}\left(\frac{fa}{(1-f)(q-a)}\right)^k \leq$$

$$\frac{(1-f)(q-a)}{(1-f)(q-a)-fa}\cdot\binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor},$$

proving (a). Letting n_3 be as in (b), we have also by 6.1 (and the fact that since

$f < \frac{q-a}{q}$) that

$$\binom{n}{0...\lfloor fn\rfloor - n_3}_{q,a} = \sum_{j=0}^{\lfloor fn\rfloor - n_3} \binom{n}{j}(q-a)^j a^{n-j} \leq$$

$$\left[\binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor}\right] \cdot \sum_{k=n_3}^{\lfloor fn\rfloor}\left(\frac{fa}{(1-f)(q-a)}\right)^k \leq$$

$$\left(\frac{fa}{(1-f)(q-a)}\right)^{n_3} \cdot \frac{(1-f)(q-a)}{(1-f)(q-a)-fa} \cdot \binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor};$$

thus the right-hand side of the above expression is equal to

$o(1) \cdot \binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor}$, so

$$\binom{n}{0...\lfloor fn\rfloor}_{q,a} = \binom{n}{0...\lfloor fn\rfloor - n_3}_{q,a} + \binom{n}{\lfloor fn\rfloor - n_3...\lfloor fn\rfloor}_{q,a} =$$

$$o(1) \cdot \binom{n}{\lfloor fn\rfloor}(q-a)^{\lfloor fn\rfloor}a^{n-\lfloor fn\rfloor} + \binom{n}{\lfloor fn\rfloor - n_3...\lfloor fn\rfloor}_{q,a} =$$

$$o(1) \cdot \binom{n}{0...\lfloor fn\rfloor}_{q,a} + \binom{n}{\lfloor fn\rfloor - n_3...\lfloor fn\rfloor}_{q,a},$$

proving (b). □

Lemma 15. *Let κ be an element of $(f, 1)$. Suppose that, for each positive integer n, nonnegative integers n_1 and n_2 are defined s.t. $n_1 + n_2 = n$, $\frac{n_1}{n} \geq \kappa$ for n sufficiently large, and $n_2 = \omega(1)$. Then*

$$\sum_{i=\lfloor fn_1\rfloor}^{\lfloor fn\rfloor}\sum_{j=\lfloor fn_1\rfloor}^{i}\binom{n_1}{j}\binom{n_2}{i-j}(q-a)^i a^{n-i} \geq [\frac{1}{2} - o(1)] \cdot \sum_{i=0}^{\lfloor fn\rfloor}\binom{n}{i}(q-a)^i a^{n-i}.$$

Proof. For $n \in \mathbb{Z}^+$,

- let $\lambda := \frac{12}{f\kappa(\kappa-f)}$;
- let $\bar{\gamma}(n) := \frac{1}{\sqrt{n_2}}$; then $\bar{\gamma}(n) = o(1)$ and $\bar{\gamma}(n) \cdot n_2 = \omega(1)$;
- let $\bar{\delta}(n) := \sqrt{\min(\bar{\gamma}(n) \cdot n_2, \log_{1-\lambda\cdot\bar{\gamma}(n)} \frac{1}{2}}$;
- let $n_3 := \lceil \bar{\delta}(n) + \frac{1}{\kappa}\rceil$; then $n_3 = \omega(1)$;

then by Lemma 4, for n sufficiently large and any integer $i := \lfloor fn_1 \rfloor + r \leq n$ *s.t.* $0 \leq r$ and $fn_2 - n_3 \leq r$,[29]

$$\sum_{j=\lfloor fn_1 \rfloor}^{i} \frac{\binom{n_1}{j}\binom{n_2}{i-j}}{\binom{n}{i}} \geq \frac{1}{2} - \frac{8}{\bar{\delta}(n)};$$

thus

$$\sum_{i=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \sum_{j=\lfloor fn_1 \rfloor}^{i} \binom{n_1}{j}\binom{n_2}{i-j}(q-a)^i a^{n-i} \geq$$
$$\sum_{i=\lfloor fn \rfloor - n_3}^{\lfloor fn \rfloor} \sum_{j=\lfloor fn_1 \rfloor}^{i} \binom{n_1}{j}\binom{n_2}{i-j}(q-a)^i a^{n-i} \geq$$
$$\sum_{i=\lfloor fn \rfloor - n_3}^{\lfloor fn \rfloor} \binom{n}{i} \cdot [\frac{1}{2} - \frac{8}{\bar{\delta}(n)}] \cdot (q-a)^i a^{n-i} =$$
$$[\frac{1}{2} - o(1)] \cdot \sum_{i=\lfloor fn \rfloor - n_3}^{\lfloor fn \rfloor} \binom{n}{i} \cdot (q-a)^i a^{n-i};$$

by 14 (b) and since $n_3 = \omega(1)$, the above expression is equal to

$$[\frac{1}{2} - o(1)] \cdot \sum_{i=0}^{\lfloor fn \rfloor} \binom{n}{i}(q-a)^i a^{n-i}.$$

□

6.2 Proof of the Main Theorem

We are now ready to state a stronger version of our main result Theorem 0:

Theorem 16. *Let q, a be positive integers with $a < q$ and f be an element of $(0, \frac{q-a}{q})$. Let c_5 be as in Lemma 12. For each positive integer n, let $n_2 := \lfloor \frac{1}{1-H_{q,a}(f)} \log_q(\ln(n)) \rfloor$ and let $n_1 := n - n_2$. Then the following holds for every sufficiently large integer n and every integer M s.t. $M \geq 3c_5 \cdot \sqrt{n_2} \cdot \frac{q^n}{\binom{n}{0 \ldots \lfloor fn \rfloor}_{q,a}}$:*

---

[29] and thus $\lfloor fn \rfloor \leq \lfloor fn_1 \rfloor + \lceil fn_2 \rceil \leq \lfloor fn_1 \rfloor + r + n_3 = i + n_3$

Let y_0 be the vector, indexed by the nonnegative integers, s.t. $y_0(0) = M$ and $y_0(i) = 0$ for each $i \in \mathbb{Z}^+$. Let $\mathbf{X}$ be a legal response sequence to y_0 with length n. Then $\sum_{i=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \mathcal{L}_{\mathbf{X}}(y_0)(i) \geq 1$. [30]

Proof. We will use Fact 1 several times without reference.

For $j \in \{\lfloor fn_1 \rfloor, \ ..., \ \lfloor fn \rfloor\}$,

$$\mathcal{L}^{n_1}(y_0)(j) = M \frac{\binom{n_1}{j}(q-a)^j a^{n_1-j}}{q^{n_1}};$$

now if n is sufficiently large that $\frac{q-a}{q}\frac{n_1}{n} \geq f$, then we have by the Stirling approximation and Fact 6 that

$$\begin{aligned} \frac{\binom{n_1}{j}(q-a)^j a^{n_1-j}}{q^{n_1}} \geq \\ \Theta(1) \cdot \frac{\sqrt{n_1}}{\sqrt{n}} \frac{\binom{n_1}{\lfloor fn_1 \rfloor}(q-a)^{\lfloor fn_1 \rfloor} a^{n_1-\lfloor fn_1 \rfloor}}{q^{n_1}} \geq \\ \Theta(1) \cdot \frac{\binom{n_1}{\lfloor fn_1 \rfloor}(q-a)^{\lfloor fn_1 \rfloor} a^{n_1-\lfloor fn_1 \rfloor}}{q^{n_1}}, \end{aligned}$$

so by Facts 13 and 14(a),

$$\mathcal{L}^{n_1}(y_0)(j) \geq \sqrt{n_2} \cdot \Theta\left(q^{[1-H_{q,a}(f)]\cdot n_2}\right) = \omega(1) \cdot \ln(n),$$

so by Lemma 12 (with $B = 0$),

$$\begin{aligned} \mathcal{L}_{\mathbf{X},\ n_1}(y_0)(j) \geq \\ [1 - \frac{c_5 \ln(n)}{\omega(1)\ln(n)}] \cdot M \cdot \frac{\binom{n_1}{j}(q-a)^j a^{n_1-j}}{q^{n_1}} = \\ [1 - o(1)] \cdot M \cdot \frac{\binom{n_1}{j}(q-a)^j a^{n_1-j}}{q^{n_1}} \end{aligned} \tag{6.2}$$

where the $\omega(1)$ and $o(1)$ functions are uniform in j.

---

[30] We could replace the factor of $3c_5$ in our hypothesis with $(2 + o(1))c_5$, where the $o(1)$ function is the same as appears in Lemma 15 .

Let $y := \mathcal{L}_{\mathbf{X},\, n_1}(y_0)$. Then for $i \in \{\lfloor fn_1 \rfloor,\ ...,\ \lfloor fn \rfloor\}$,

$$\mathcal{L}^{n_2}(y)(i) \geq \sum_{j=\lfloor fn_1 \rfloor}^{i} y(j) \frac{\binom{n_2}{i-j}(q-a)^{i-j}a^{n_2-i+j}}{q^{n_2}};$$

by (6.2), this is greater than or equal to

$$[1-o(1)] \cdot \frac{M}{q^n} \cdot \sum_{j=\lfloor fn_1 \rfloor}^{i} \binom{n_1}{j}\binom{n_2}{i-j}(q-a)^i a^{n-i},$$

where the $o(1)$ function is again uniform in j; thus

$$\sum_{i=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \mathcal{L}^{n_2}(y)(i) \geq$$

$$[1-o(1)] \cdot \frac{M}{q^n} \cdot \sum_{i=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \sum_{j=\lfloor fn_1 \rfloor}^{i} \binom{n_1}{j}\binom{n_2}{i-j}(q-a)^i a^{n-i};$$

by Lemma 15, the above expression is greater than or equal to

$$[1-o(1)] \cdot \frac{M}{q^n} \cdot [\frac{1}{2} - o(1)] \cdot \sum_{i=0}^{\lfloor fn \rfloor} \binom{n}{i}(q-a)^i a^{n-i},$$

and by the definition of M this expression is greater than or equal to

$$c_5 \cdot [\frac{3}{2} - o(1)] \cdot \sqrt{n_2}.$$

Thus by Lemma 12 (with $B = \lfloor fn \rfloor - \lfloor fn_1 \rfloor$), and for n large enough that $\lfloor fn \rfloor - \lfloor fn_1 \rfloor \geq \frac{\sqrt{n_2}}{2}$,

$$\sum_{i=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \mathcal{L}_{\mathbf{X}}(y_0)(i) \geq$$

$$\left[\sum_{i=\lfloor fn_1 \rfloor}^{\lfloor fn \rfloor} \mathcal{L}^{n_2}(y)(i)\right] - c_5 \cdot \sqrt{n_2} \geq$$

$$c_5 \cdot \sqrt{n_2} \cdot (\frac{3}{2} - o(1) - 1) = \omega(1);$$

for n large enough the above expression is ≥ 1, as desired.

□

Since $\mathcal{L}_{\mathbf{X}}(y')$ is an integer vector whenever y' is an integer vector and $\mathbf{X}$ is a legal response sequence to y', this implies that at least one coordinate of $\mathcal{L}_{\mathbf{X}}(y_0)$ is at least 1.

- Since $\mathcal{L}_{\mathbf{X}}(y_0)(i)$ describes, for each legal response sequence $\mathbf{X}$ to y and each nonnegative integer i, the number of messages with error count i after Carole gives the sequence of answers corresponding to $\mathbf{X}$, and
- since any allowed sequence of answers by Carole corresponds to a legal response sequence to y,

we now have Theorem 0.

APPENDIX A

GUIDE TO CITED RESULTS

All unmarked references in this appendix are to the version of Siegel's paper [Siegel, 2001] appearing in "Journal of Algorithms," vol. 38.

A.1 Correcting Typos

There are some typos in Siegel's paper on median bounds [Siegel, 2001] that are of relevance to our arguments, which we list below. We will violate some rules of punctuation for clarity.

(i) Near the top of p.201, in the statement of Theorem 2.4 (beginning on p.200), in the expression "b1′) The density f'", the expression "f'" should be changed to "f".

(ii) In the second-to-last sentence of the proof discussion for Corollary 2.3 (p.203), the expression "$T^+_\bullet(t)$" should be changed to "$T^+_\bullet$".

(iii) In the second sentence of the proof of Corollary 2.3 (appendix A.3, p.221), the expression "$T^+_\bullet(t)$" should be changed to "$T^+_\bullet$".

(iv) In the fifth centered equation line on p.221, the expression "$\ln(R-b)$" should be changed to "$\ln(R-r)$".

(v) In the second-to-last centered equation line on p.221, a "?" appears immediately after the "$\geq$" sign (at least in our .pdf reader). This "?" should be deleted.

(vi) In the first centered equation line on p.222, a "?" appears immediately after the "$\geq$" sign (at least in our .pdf reader). This "?" should be deleted.

A.2 Generalizing Corollary 2.3

Our Lemma 2 is a tiny generalization of Siegel's formulation [Siegel, 2001, Corollary 2.3, p. 203]. This generalization was noted by Cooper and Ellis [Cooper and Ellis, 2010, Theorem 12]. Though stated slightly differently here, our statement is equivalent to that of Cooper and Ellis. Siegel's formulation is proven in his paper [Siegel, 2001, Appendix A.3, pp. 221-222]. The necessary prerequisites for this proof are discussed in Appendix A.3.

Siegel's proof can be converted to a proof of our more general statement by the following trivial steps:

(i) Correct the typos in Siegel's paper according to Appendix A.1.

(ii) In the second-to-last sentence of the proof discussion on p.203, change b to $\lceil b \rceil$.

(iii) In the proof of Corollary 2.3 (beginning on p.221), change b to $\lceil b \rceil$ and r to $\lceil r \rceil$.

(iv) In the fifth centered equation line on p.221, change the final $=$ to $\leq$.

(v) Replace the third sentence after the fifth centered equation line on p.221 (beginning "[s]ubstituting w for $w_\bullet$") with the following sentences: "Substituting w for $w_\bullet$ gives $\frac{e^{-wt}}{1-e^{-wt}} = \frac{B-\lceil b \rceil}{\lceil b \rceil}$, which has the unique solution $t = \rho := \frac{1}{w}(\ln(B) - \ln(B - \lceil b \rceil))$. Using the fact that $\lceil b \rceil \geq B(1-e^{-w})$ gives $\frac{1}{w}(\ln(B)-\ln(B-\lceil b \rceil)) \geq 1 > \mu$, and thereby fulfills requirement (a2)."

(vi) On pp.221-222, change all instances of $(2-t)$ to $(2\rho-t)$ and change all instances of $(0, 1)$ to $(0, \rho)$.

(vii) In the last centered equation line on p.221 and the third-to-last centered equation line on p.222, in the fourth equation and the second-to-last equation, change $\frac{e^{-w}}{1-e^{-w}}$ to $\frac{B-\lceil b \rceil}{\lceil b \rceil}$.

(viii) Delete the first line of text below the last centered equation line on p.221. In the subsequent line, change "$t = 1$" to "$t = \rho$."

A few additional points should be noted during the "conversion process":

(i) Siegel's proof (and ours) depends on other results within his paper. See Appendix A.3 for a guide to the necessary prior results.

(ii) The fourth centered equation line on p.221 holds with r replaced by any nonnegative integer $\leq R$ in this equation as well as the Definition of $F_\circ$ (on the second centered equation line on p.221). It can be verified by two integrations by parts.

(iii) Shortly below the third centered equation line on p.221, reference is made to condition b1′, which is stated at the top of p.201 in the statement of Theorem 2.4. Verification that the function f satisfies condition b1′ is left to the reader.

In fact, we will verify that f satisfies condition b1′ when r is replaced by any nonnegative integer $s \leq r$ in the Definition of $F_\circ$ (on the second centered equation line on p.221) and μ is replaced by any nonnegative real number in the Definition of b1′ (at the top of p.201 in the statement of Theorem 2.4, which begins on p.200). For $t \in (0, 2\mu)$,

$$f(t) = F'(t) = \sum_{j=s}^{R} \binom{R}{j} [w_\circ \cdot j \cdot (1 - \exp(-w_\circ t))^{j-1} (\exp(-w_\circ t))^{R-j+1} -$$

$$w_\circ \cdot (R - j) \cdot (1 - \exp(-w_\circ t))^{j} (\exp(-w_\circ t))^{R-j}] =$$

$$w_\circ \cdot \sum_{j=s}^{R} [R \cdot \binom{R-1}{j-1} \exp(-w_\circ t)(1 - \exp(-w_\circ t))^{j-1} (\exp(-w_\circ t))^{(R-1)-(j-1)} -$$

$$R \cdot \binom{R-1}{j} \exp(-w_\circ t)(1 - \exp(-w_\circ t))^{j} (\exp(-w_\circ t))^{(R-1)-j}] =$$

$$w_{\circ}R \cdot \Big(\Big[\sum_{j=s-1}^{R-1} \binom{R-1}{j} \exp(-w_{\circ}t)(1-\exp(-w_{\circ}t))^{j}(\exp(-w_{\circ}t))^{(R-1)-j} \Big] - \Big[R \cdot \sum_{j=s}^{R} \binom{R-1}{j} \exp(-w_{\circ}t)(1-\exp(-w_{\circ}t))^{j}(\exp(-w_{\circ}t))^{(R-1)-j} \Big] \Big) =$$

$$w_{\circ}R \cdot \binom{R-1}{s-1}(1-\exp(-w_{\circ}t))^{s-1}(\exp(-w_{\circ}t))^{R-s+1};$$

thus

$$\frac{f'(t)}{f(t)} = (s-1)w_{\circ} \cdot \frac{\exp(-w_{\circ}t)}{1-\exp(-w_{\circ}t)} - (R-s+1)w_{\circ},$$

so

$$\frac{f'(t)}{f(t)} + \frac{f'(2\mu-t)}{f(2\mu-t)} = (s-1)w_{\circ} \cdot \Big[\frac{\exp(-w_{\circ}t)}{1-\exp(-w_{\circ}t)} + \frac{\exp(-w_{\circ}(2\mu-t))}{1-\exp(-w_{\circ}(2\mu-t))}\Big] - 2(R-s+1)w_{\circ}. \quad \text{(A.1)}$$

Now as a function of t, the right-hand side of equation (A.1) has derivative

$$(s-1)w_{\circ} \cdot \Big[-w_{\circ}\frac{\exp(-w_{\circ}t)}{(1-\exp(-w_{\circ}t))^2} + w_{\circ}\frac{\exp(-w_{\circ}(2\mu-t))}{(1-\exp(-w_{\circ}(2\mu-t)))^2}\Big] = (s-1)w_{\circ}^2 \cdot \Big[-\frac{1}{(1-\exp(-w_{\circ}t))^2} + \frac{1}{(1-\exp(-w_{\circ}(2\mu-t)))^2}\Big],$$

which is strictly increasing in t for $t \in (0, 2\mu)$. Thus the right-hand side of equation (A.1) has at most two zeroes in $(0, 2\mu)$ as a function of t.

A.3 Understanding Corollary 2.3

We briefly trace the prerequisites for Siegel's proof of his Corollary 2.3 [Siegel, 2001, Corollary 2.3, p. 203].

Corollary 2.3 is stated and discussed on p.203, and proven in Appendix A3 on pp.221-222. This proof depends on Theorem 2.4, which is stated on pp.200-201 and proven in Appendix A.2 on pp.219-220. Theorem 2.4, in turn, depends on Theorem 2.1 and Lemma 2.1, which are stated and proven on pp.190-191.

APPENDIX B

WRITING CONVENTIONS

The following information may be useful to the reader:

Level of Formality Each of our formal statements is made either

- within a proof (which begins with "*Proof.*" and ends with "□") or
- in *italics*, headed by one of the words "**Definition**," "**Convention**," "**Game**," "**Fact**," "**Lemma**," or "**Theorem**," and terminated by the first non-italic word after the heading.

Any other statements should not be taken as definitions or as contributing substantively to any of our arguments; they are for purely explanatory purposes.

Definitions of Symbols

- Some symbol definitions change from section to section as additional constraints are progressively added to them. Definitions given in examples, or in the statement or proof of a result do not extend beyond that comment or result. Labeled definitions and conventions, on the other hand, should be considered to persist throughout the thesis until they are changed.
- All Landau notation in this thesis (big-O, little-o, big-Ω, little-ω, big-Θ) is as $n \to \infty$ and $n \in \mathbb{Z}^+$.
- Given a real number j', the "sign of j'" means $\begin{cases} -1 & \text{if } j' < 0 \\ 0 & \text{if } j' = 0 \\ 1 & \text{if } j' > 0 \end{cases}$.
- A real number j' is "between" real numbers i' and k' if either $i' \leq j' \leq k'$ or $k' \leq j' \leq i'$.
- We occasionally discuss random variables without explicitly defining the probability space on which they are defined. We will invariably assume that the probability space is equal to $(S, 2^S, \Pr)$, where S is a finite set

and Pr is a probability measure on 2^S. For a proposition P with domain S we will use "Pr$\{P\}$" to denote Pr$(\{\sigma \in S \; s.t. \; P(\sigma) \; is \; true\})$.

APPENDIX C

STATEMENT ON COLLABORATION

This work forms part of a collaboration with James Williamson. His forthcoming M.S. thesis [Williamson, 2012] and this paper cover complementary topics, each extending a different portion of the results of Cooper and Ellis [Cooper and Ellis, 2010]. We remained in communication with each other as I wrote this paper, harmonizing a good deal of our notation and coordinating our research objectives to prevent duplication of effort. His thesis refers to several results from this paper. The main Theorem 16 of this paper is essentially a stronger version of his main Theorem 3 [Williamson, 2012], although his argument differs from mine.

BIBLIOGRAPHY

[Berlekamp, 1964] Berlekamp, E. (1964). *Block Coding with Noiseless Feedback*. PhD thesis, Massachusetts Institute of Technology.

[Cooper et al., 2007] Cooper, J., Doerr, B., Spencer, J., and Tardos, G. (2007). Deterministic random walks on the integers. *European Journal of Combinatorics*, *28*:2072–2090.

[Cooper and Ellis, 2010] Cooper, J. and Ellis, R. (2010). Linearly bounded liars, adaptive covering codes, and deterministic random walks. *Journal of Combinatorics*, *1*(3-4):307–344.

[Delsarte and Piret, 1986] Delsarte, P. and Piret, P. (1986). Do most binary linear codes achieve the goblick bound on the covering radius? (corresp.). *Information Theory, IEEE Transactions on*, 32(*6*):826–828.

[Du and Hwang, 2006] Du, D. and Hwang, F. (2006). *Pooling designs and nonadaptive group testing*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 1st edition.

[Ellis and Yan, 2004] Ellis, R. and Yan, C. (2004). Ulam's pathological liar game with one half-lie. *International Journal of Mathematics and Mathematical Sciences*, (29-32):1523–1532.

[Rényi, 1961] Rényi, A. (1961). On a problem in information theory. *Magyar Tudományos Akadémia Matematikai Kutató Intézeténk Közleményei*, *6*:505–516. The original article appears in Hungarian.

[Siegel, 2001] Siegel, A. (2001). Median bounds and their application. *Journal of Algorithms*, *38*(1):184–236.

[Spencer, 1992] Spencer, J. (1992). Ulam's searching game with a fixed number of lies. *Theoretical Computer Science*, *95*:307–321.

[Spencer and Winkler, 1992] Spencer, J. and Winkler, P. (1992). Three thresholds for a liar. *Combinatorics, Probability and Computing*, *1*(1):81–93.

[Ulam, 1976] Ulam, S. (1976). *Adventures of a Mathematician*. Charles Scribner's Sons, New York.

[Williamson, 2012] Williamson, J. (Forthcoming 2012). Analysis of the application of the liar machine to the q-ary pathological liar game with a focus on lower discrepancy bounds.
Unpublished Master's Thesis. Title given is the working title. By a collaborator.