Improved Lower Bounds for the Fourier Entropy/Influence Conjecture via Lexicographic Functions

Rani Hod

November 3, 2017

Abstract

Every Boolean function can be uniquely represented as a multilinear polynomial. The entropy and the total influence are two ways to measure the concentration of its Fourier coefficients, namely the monomial coefficients in this representation: the entropy roughly measures their spread, while the total influence measures their average level. The Fourier Entropy/Influence conjecture of Friedgut and Kalai from 1996 states that the entropy to influence ratio is bounded by a universal constant $C$.

Using lexicographic Boolean functions, we present three explicit asymptotic constructions that improve upon the previously best known lower bound $C > 6.278944$ by O'Donnell and Tan, obtained via recursive composition. The first uses their construction with the lexicographic function $\ell_{2/3}$ of measure $2/3$ to demonstrate that $C \geq 4 + 3 \log_2 3 > 6.377444$. The second generalizes their construction to biased functions and obtains $C > 6.413846$ using $\ell_{\Phi}$, where $\Phi$ is the inverse golden ratio. The third, independent, construction gives $C > 6.454784$, even for monotone functions.

Beyond modest improvements to the value of $C$, our constructions shed some new light on the properties sought in potential counterexamples to the conjecture. Additionally, we prove a Lipschitz-type condition on the total influence and spectral entropy, which may be of independent interest.

1 Introduction

Let $\text{true} = -1$ and $\text{false} = +1$. Throughout this paper, we write $[n] = \{1, 2, \ldots, n\}$ and $N = 2^n$ for an integer $n \geq 1$. It is well known that any function $f : \{\text{true, false}\}^n \to \mathbb{R}$ can be expressed as

$$f = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S,$$

where $\chi_S(x) = \prod_{i \in S} x_i$ for $S \subseteq [n]$ are the Fourier basis functions and

$$\hat{f}(S) = \langle f, \chi_S \rangle = \mathbb{E}[f(x) \chi_S(x)]$$

for $S \subseteq [n]$ are called the Fourier coefficients of $f$. When $f$ is a Boolean function, i.e., $f : \{\text{true, false}\}^n \to \{\text{true, false}\}$, we have $\sum_{S \subseteq [n]} \hat{f}(S)^2 = 1$ by Parseval, so we can treat the Fourier coefficients’ squares as a probability distribution $p_f$ on the $N$ subsets of $[n]$, which we call the spectral distribution of $f$.

The following two parameters of the function $f$ can be defined in terms of its spectral distribution.

*Department of Mathematics, Bar Ilan University, Ramat Gan, Israel. Email: rani.hod@math.biu.ac.il. The research was partially supported by the Israel Science Foundation (grants no. 402/13 and 1612/17), by the Binational US-Israel Science Foundation (grant no. 2014290), and by the Coleman–Soref postdoctoral fellowship.
**Definition.** The *total influence* (also called average sensitivity) of a Boolean function \( f \) is
\[
\mathbf{I}[f] = \mathbb{E}_{S \sim p_f} [ |S| ] .
\]

**Definition.** The *spectral entropy* of a Boolean function \( f \) is the (Shannon) entropy of its spectral distribution
\[
\mathbf{H}[f] = \mathbb{E}_{S \sim p_f} \left[ - \log_2 (p_f(S)) \right] = \sum_S p_f(S) \log_2 (p_f(S)) .
\]

In 1996 Friedgut and Kalai raised the following conjecture, known as the Fourier Entropy/Influence (FEI) conjecture:

**Conjecture 1.1 (FEI).** There exists a universal constant \( C > 0 \) such that for every Boolean function \( f \) with total influence \( \mathbf{I}[f] \) and spectral entropy \( \mathbf{H}[f] \) we have \( \mathbf{H}[f] \leq C \cdot \mathbf{I}[f] \).

Conjecture 1.1 was verified for various families of Boolean functions (e.g., symmetric functions [10], random functions [3], read-once formulas [1, 9], decision trees of constant average depth [11], read-\( k \) decision trees for constant \( k \) [11]) but is still open for the class of general Boolean functions.

The rest of this paper is organized as follows. In the remainder of Section 1 we describe past results and some rudimentary improvements. In Section 2 we introduce lexicographic functions and provide a formal proof of the approach described in Section 1.3. In Section 3 we generalize Proposition 1.2 to biased functions and get an improved lower bound. In Section 4 we build a limit-of-limits function that achieves an even better bound. In Section 5 we prove a Lipschitz-type condition used throughout the paper, namely that a small change in a Boolean function cannot result in a substantial change to its total influence and spectral entropy.

### 1.1 A baby example and two definitions

Here is an example of providing a lower bound on \( C \). For \( n \geq 1 \) consider the function
\[
\text{And}_n (x_1, \ldots, x_n) = x_1 \land x_2 \land \cdots \land x_n .
\]

It satisfies \( \mathbf{I}[\text{And}_n] = 2n/N \) and \( \mathbf{H}[\text{And}_n] \approx 8 \left( n - 1 + 1/\ln 4 \right) / N \), so any constant \( C \) in Conjecture 1.1 must satisfy
\[
C \geq \frac{\mathbf{H}[\text{And}_n]}{\mathbf{I}[\text{And}_n]} \approx 4 - \frac{4}{n} (1 - 1/\ln 4) .
\]

This is true for every \( n \), so by taking \( n \to \infty \) we establish that \( C \geq 4 \).

**Definition.** A Boolean function \( f \) is called *monotone* if changing an input bit \( x_i \) from *false* to *true* cannot change the output \( f(x) \) from *true* to *false*.

**Fact.** A Boolean function is monotone if and only if it can be expressed as a formula combining variables using conjunctions (\( \land \)) and disjunctions (\( \lor \)) only, with no negations.

**Definition.** Let \( f \) be a Boolean function on \( n \) variables. The *dual* function of \( f \), denoted \( f^\dagger \), is defined as
\[
f^\dagger (x_1, \ldots, x_n) = \neg f (\neg x_1, \ldots, \neg x_n) .
\]

\(^1\)More precisely, \( 0 < 8 \left( n - 1 + 1/\ln 4 \right) / N - \mathbf{H}[\text{And}_n] < 12n/N^2 \) for all \( n \geq 1 \).
Fact. For all $S \subseteq [n]$ we have $\hat{f}^{\dagger} (S) = (-1)^{|S|+1} \hat{f} (S)$.

Corollary. The spectral distributions $p_f$ and $p_{f^\dagger}$ are identical; in particular, $I [f^\dagger] = I [f]$ and $H [f^\dagger] = H [f]$. But $E [f^\dagger] = -E [f]$ and $Pr [f^\dagger (x) = \text{true}] = 1 - Pr [f (x) = \text{true}]$.

Remark. If $f$ is monotone then $f^\dagger$ is monotone too. Furthermore, given a monotone formula computing $f$, the formula obtained by swapping conjunctions and disjunctions computes $f^\dagger$.

Example. The dual of $\text{Or}_n (x_1, \ldots, x_n) = x_1 \lor x_2 \lor \cdots \lor x_n$ is $\text{And}_n$.

1.2 Past results and preliminary improvements

The current best lower bound on $C$ was achieved by O’Donnell and Tan [9]. Using recursive composition they showed the following bound:

Proposition 1.2. Let $g$ be a balanced Boolean function such that $H [g] > 0$. Then any constant $C$ in Conjecture [2.1] satisfies $C \geq H [g] / (I [g] - 1)$.

Remark. Any balanced Boolean function $g$ has $I [g] \geq 1$ since $p_g (\emptyset) = E [g] = 0$; in case of equality we must have $g = \chi_{\{i\}}$ for some $i \in [n]$ and thus $p_y$ is supported on a single set $S = \{i\}$ and its spectral entropy is zero.

By presenting a function on 6 variables with total influence $I = 13/8 = 1.625$ and entropy $H > 3.92434$, they established that $C > 3.92434/\frac{5}{8} = 6.278944$. Although the specific function presented in [9] happens to be biased, their result stands as there exists a balanced Boolean function $g_3$ on 6 variables with the same total influence and entropy:

$$g_3 (x_1, x_2, x_3, x_4, x_5, x_6) = (x_1 \lor x_2) \land (x_3 \lor x_4) \land (x_1 \lor x_3 \lor x_5) \land (x_3 \lor x_5 \lor x_6).$$

A slight improvement can be achieved by modifying the last clause of $g_3$. Indeed,

$$g'_3 (x_1, x_2, x_3, x_4, x_5, x_6) = (x_1 \lor x_2) \land (x_3 \lor x_4) \land (x_1 \lor x_3 \lor x_5) \land (x_2 \lor x_4 \lor x_6)$$

is balanced too, with the same total influence $I [g'_3] = I [g_3] = 13/8$ and a slightly higher entropy $H [g'_3] > 3.9669$, so we have $C > 3.9669/\frac{5}{8} = 6.34704$.

Moving to balanced functions on 8 variables, we find a monotone function $g_4$ that provides a better lower bound:

$$g_4 (x_1, \ldots, x_8) = (x_1 \lor x_2) \land (x_3 \lor x_4) \land (x_3 \lor x_5 \lor x_6) \land (x_1 \lor x_3 \lor x_5 \lor x_7) \land (x_3 \lor x_5 \lor x_7 \lor x_8)$$

with $I [g_4] = 53/32$ and $H [g_4] > 4.16885$ yields $C > 6.35253$.

A further search discovers a slightly superior function:

$$g'_4 (x_1, \ldots, x_8) = (x_1 \lor x_2) \land (x_3 \lor x_4) \land (x_3 \lor x_5 \lor x_6) \land (x_1 \lor x_3 \lor x_5 \lor x_7) \land (x_2 \lor x_3 \lor x_6 \lor x_8)$$

with $I [g'_4] = I [g_4] = 53/32$ and $H [g'_4] > 4.17635$ achieves $C > 6.36396$.

1.3 Sequences of balanced monotone functions

Staring at $g_3$, $g_4$ and $g'_4$ for a moment (but not $g'_3$), we may see a common property: $x_3$ appears in all clauses except the first. Let us rewrite $g_3$ and $g_4$ in a slightly different form:

$$g_3 (x_1, \ldots, x_6) = (x_1 \lor x_2) \land (x_3 \lor (x_4 \lor (x_5 \lor x_6 \land x_1))),$$
$$g_4 (x_1, \ldots, x_8) = (x_1 \lor x_2) \land (x_3 \lor (x_4 \lor (x_5 \lor (x_6 \lor (x_7 \lor (x_8 \lor x_1)))))).$$
This generalizes easily to a sequence \((g_m)_{m \geq 1}\) of balanced (to be shown below) monotone Boolean functions:

\[
g_m(x_1, \ldots, x_{2m}) = (x_1 \lor x_2) \land (x_3 \lor (x_4 \land (x_5 \lor \cdots (x_{2m-1} \lor (x_{2m} \land x_1))\cdots)))
\]

whose first two members are

\[
g_1(x_1, x_2) = (x_1 \lor x_2) \land x_1 = x_1, \\
g_2(x_1, x_2, x_3, x_4) = (x_1 \lor x_2) \land (x_3 \lor (x_4 \land x_1)) = (x_1 \lor x_2) \land (x_3 \lor x_4) \land (x_1 \lor x_3).
\]

Denote by \(C_m = H[g_m] / (I[g_m] - 1)\) the lower bound on \(C\) implied by \(g_m\). The first fifteen members of the sequence are explored in Table 1. Note how even \(C_2 = 6\) is much better than the \(C \geq 4\) bound of Subsection 11.1.

| \(m\) | \(n\) | \(I[g_m]\) | \(H[g_m]\) | \(C_m\) |
|---|---|---|---|---|
| 1 | 2 | 3/2 = 1.5 | 0 | (not defined) |
| 2 | 4 | 3/2 = 1.5 | 3 | 6 |
| 3 | 6 | 13/8 = 1.625 | > 3.92434 | > 6.27894 |
| 4 | 8 | 53/32 = 1.65625 | > 4.16885 | > 6.35253 |
| 5 | 10 | 213/128 = 1.6640625 | > 4.23087 | > 6.37119 |
| 6 | 12 | 853/512 = (5 - 2^{-9})/3 | > 4.24643 | > 6.37588 |
| 7 | 14 | 3413/2048 = (5 - 2^{-11})/3 | > 4.25033 | > 6.37705 |
| 8 | 16 | 13653/2^{13} = (5 - 2^{-10})/3 | > 4.25130 | > 6.37734 |
| 9 | 18 | 54613/2^{15} = (5 - 2^{-10})/3 | > 4.25154 | > 6.37741 |
| 10 | 20 | (5 - 2^{-17})/3 | > 4.251608 | > 6.377437 |
| 11 | 22 | (5 - 2^{-19})/3 | > 4.251624 | > 6.3774422 |
| 12 | 24 | (5 - 2^{-21})/3 | > 4.2516278 | > 6.3774433 |
| 13 | 26 | (5 - 2^{-23})/3 | > 4.25162885 | > 6.37744365 |
| 14 | 28 | (5 - 2^{-25})/3 | > 4.25162908 | > 6.37744372 |
| 15 | 30 | (5 - 2^{-27})/3 | > 4.251629147 | > 6.377443745 |

Table 1: Parameters of the sequence \(g_m\) for \(m \leq 15\)

The three sequences seem to be increasing and bounded, so let us denote their respective hypothetical limits by \(I_*, H_*\) and \(C_*\). If indeed \(I[g_m] = \frac{1}{3}(5 - 2^{3-2m})\) for all \(m \in \mathbb{N}\) then \(I_* = 5/3\). A prescient guess for the value of \(H_*\) could be

\[
H_* = \frac{8}{3} + \log_2 3 > 4.251629167,
\]

for which we would get

\[
C_* = H_*/(2/3) = 4 + 3\log_3 3 > 6.377443751
\]

as a lower bound for \(C\). We will verify this guess in Section 2.

Recall that \(g'_3\) and \(g'_4\) gave rise to better lower bounds, respectively, than \(g_3\) and \(g_4\). It is tempting perhaps to consider a generalization \((g'_m)_{m \geq 3}\), define \(C'_m\) accordingly and examine the hypothetical limits \(I'_*, H'_*\) and \(C'_*\). It is indeed possible to do so, and we get \(I'[g_m] = I[g_m]\) while \(H'[g_m] > H[g_m]\), making \(C'_m > C_m\). Nevertheless, \(H[g'_m]\) and \(C'_m\) seem to converge towards the same \(H_*\) and \(C_*\), respectively, so there is no real benefit in pursuing this further.
It remains to verify that $g_m$ is indeed balanced for all $m \geq 1$. Let us write it as

$$g_m(x_1, \ldots, x_{2m}) = (x_1 \lor x_2) \land G_m(x_3, x_4, \ldots, x_{2m}, x_1),$$

where $G_m(y_1, y_2, \ldots, y_{2m-1})$ is defined recursively via

$$G_1(y_1) = y_1,$$

$$G_{m+1}(y_1, \ldots, y_{2m+1}) = y_1 \lor (y_2 \land G_m(y_3, \ldots, y_{2m+1})).$$

**Remark.** The function $G_m$ belongs to a class of monotone Boolean functions called lexicographic functions, as we will see in Section 2.1.

For simplicity of notation, we abbreviate and write $\Pr[f(x)]$ or even $\Pr[f]$ to denote $\Pr[f(x) = \text{true}]$. Since

$$\Pr[g_m(x)] = \Pr[x_1 \lor x_2] \cdot \Pr[G_m(x_3, \ldots, x_{2m}, x_1) \mid x_1 \lor x_2],$$

to prove $\Pr[g_m] = \frac{1}{2}$, it suffices to verify the following (see Appendix A for the calculation):

**Claim 1.3.** For all $m \geq 1$ we have $\Pr[G_m(x_3, \ldots, x_{2m}, x_1) \mid x_1 \lor x_2] = 2/3$.

## 2 A Tale of Two Thirds

Although each of $(C_m)_{m=2}^{15}$ from Table I is a valid, explicit lower bound on $C$, the asymptotic discussion in Subsection I.3 was more of a wishful thinking rather than a mathematically sound statement.

In this section we explore the class of lexicographic functions, develop tools to compute total influence and spectral entropy, and then rigorously calculate $I_s$, $H_s$ and $C_s$.

### 2.1 Lexicographic functions

**Definition.** Fix integers $n \geq 1$ and $0 \leq s \leq N$. Denote by $L_n(s) \subseteq \{\text{true, false}\}^n$ the initial segment of cardinality $s$ (with respect to the lexicographic order on $\{\text{true, false}\}^n$), and denote by $\ell_n(s) : \{\text{true, false}\}^n \rightarrow \{\text{true, false}\}$ its characteristic function

$$\ell_n(s)(x) = \begin{cases} \text{true}, & x \in L_n(s); \\ \text{false}, & x \notin L_n(s). \end{cases}$$

**Fact.** We have $\Pr[\ell_n(s)] = |L_n(s)|/N = s/N$ and $E[\ell_n(s)] = 1 - 2s/N$.

**Fact.** The function $\ell_n(s)$ is monotone and its dual is $(\ell_n(s))^\dagger = \ell_n(N-s)$.

**Example.** $\ell_n(1) = \text{And}_n$ and $\ell_n(N-1) = \text{Or}_n$.

**Fact.** If $s$ is even then $\ell_n(s)$ is isomorphic to $\ell_{n-1}(s/2)$ (when the latter is extended from $n-1$ to $n$ variables by adding an influenceless variable).

Let $0 < s < N$ be an odd integer, and let $s_1 s_2 \cdots s_n$ be its binary representation, where $s_1$ is the most significant bit and $s_n = 1$ is the least significant bit. Denote the corresponding $\{\text{true, false}\}^n$ representation of $s$ by $\vec{s} = ((-1)^{s_1}, \ldots, (-1)^{s_n})$.

By definition, to determine the value of $\ell_n(s)$ for an input $x$, we need to compare $x$ with $\vec{s}$ element by element. This gives a neat formula for $\ell_n(s)$:

$$\ell_n(s)(x_1, \ldots, x_n) = x_1 \circ_1 (x_2 \circ_2 (x_3 \circ_3 \cdots (x_{n-1} \circ_{n-1} x_n) \cdots)),$$

where $\circ_i = \begin{cases} \land, & s_i = 0; \\ \lor, & s_i = 1. \end{cases}$
**Remark.** The formula (1) shows that every monotone decision list, i.e., a monotone decision tree consisting of a single path, is isomorphic to a lexicographic function.

From (1) we derive an important property of lexicographic functions.

**Fact 2.1.** For $k \in [n]$, the value of $\ell_n \langle s \rangle (x)$ only depends on $x_k$ with probability $2^{1-k}$; that is, when $x_k = (-1)^{s_i}$ for all $i < k$.

**Remark.** This can be interpreted as saying that the average decision tree complexity of $\ell_n \langle s \rangle$ is $2 - (n + 2)/N$.

We extend the definition of lexicographic functions by writing $\ell_n \langle \mu \rangle = \ell_n \langle \langle \mu N \rangle \rangle$ for some $0 \leq \mu \leq 1$. Note that $|\mu N|$ is not necessarily odd, so the effective number of variables can be smaller.

**Example.** For any $n \geq 2$ we have $\ell_n \langle 3/4 \rangle (x) = x_1 \lor x_2$; that is, $\ell_n \langle 3/4 \rangle = \text{Or}_2$.

**Example.** For $n = 2m - 1$, we have $\ell_n \langle 2/3 \rangle = \ell_n \langle [2N/3] \rangle = \ell_n \langle (2N - 1)/3 \rangle$. Observe that the binary representation of the odd integer $s = (2N - 1)/3$ has $s_i = i \mod 2$ for $i \in [n]$ and thus

$$\ell_n \langle 2/3 \rangle (x) = x_1 \lor (x_2 \land (x_3 \lor (x_4 \land \cdots (x_{2m-2} \land x_{2m-1}) \cdots)))$$

that is, $\ell_n \langle 2/3 \rangle = G_m$.

Fix $0 < \mu < 1$ and consider the sequence $(\ell_n \langle \mu \rangle)_{n \geq 1}$. Whenever $\mu$ is a dyadic rational, $\ell_n \langle \mu \rangle$ converges to a fixed function $\ell \langle \mu \rangle$ (e.g., $\ell \langle 3/4 \rangle = \text{Or}_2$ in the example above). We would like to consider the limit object $\ell \langle \mu \rangle = \lim_{n \to \infty} \ell_n \langle \mu \rangle$ for other values of $\mu$ as well.

It may sound intimidating; after all, $\ell \langle \mu \rangle : \{\text{true, false}\}^N \to \{\text{true, false}\}$ is a Boolean function on $\mathbb{N}_0$ variables, which is quite a lot. Nevertheless, by Fact 2.1 $\ell \langle \mu \rangle$ only reads two input bits on average.

Moreover, we care about the total influence and spectral entropy of functions. By Lemmata 5.1 and 5.2 from Section 5, $\mathbf{I}[\ell \langle \mu \rangle] \xrightarrow{n \to \infty} \mathbf{I}[\ell \langle \mu \rangle]$ and $\mathbf{H}[\ell \langle \mu \rangle] \xrightarrow{n \to \infty} \mathbf{H}[\ell \langle \mu \rangle]$. Indeed, $\ell_n \langle \mu \rangle$ differs from $\ell_{n-1} \langle \mu \rangle$ (when considering the latter as a function on $n$ variables by adding an influenceless variable) in at most one place, and thus $(\mathbf{I}[\ell_n \langle \mu \rangle])_{n \geq 1}$ and $(\mathbf{H}[\ell_n \langle \mu \rangle])_{n \geq 1}$ are Cauchy sequences.

Needless to say, $\Pr[\ell_n \langle \mu \rangle] = [\mu N] / N \xrightarrow{n \to \infty} \mu = \Pr[\ell \langle \mu \rangle]$.

An even stronger statement holds (but will not be used or proved here): the spectral distributions of $\ell_n \langle \mu \rangle$ converge in distribution to a limit distribution $p_\mu$, which we call the spectral distribution of $\ell \langle \mu \rangle$. Note that $p_\mu$ is supported on finite subsets of $\mathbb{N}$. The expected cardinality and the entropy of $S \sim p_\mu$ are $\mathbf{I}[\ell \langle \mu \rangle]$ and $\mathbf{H}[\ell \langle \mu \rangle]$ respectively.

### 2.2 Total influence and lexicographic functions

The edge isoperimetric inequality in the discrete cube (by Harper [5], with an addendum by Bernstein [2], and independently Lindsey [7]) gives a lower bound on the total influence of Boolean functions.

**Theorem 2.2.** Let $f$ be a Boolean function with $\Pr[f] = \mu \leq \frac{1}{2}$. Then $\mathbf{I}[f] \geq -2\mu \log_2 \mu$.

In fact, they proved that lexicographic functions are the minimizers of total influence.

**Theorem 2.3.** Fix integers $n \geq 1$ and $s \leq N/2$ and let $f$ be a Boolean function on $n$ variables with $\Pr[f] = s/N$. Then $\mathbf{I}[f] \geq \mathbf{I}[\ell_n \langle s \rangle]$.

---

2 That is, a rational number of the form $a/2^b$. 

6
**Remark.** Theorem 2.3 explains our interest in lexicographic functions: when seeking a function \( f \) with large entropy/influence ratio \( H[f]/I[f] \), it makes sense to minimize \( I[f] \).

In [6], Hart exactly computed the total influence of lexicographic functions:

**Proposition 2.4 (6 Theorem 1.5).** Fix integers \( n \geq 1 \) and \( 0 \leq s \leq N \). Then

\[
I[\ell_n(s)] = \frac{2sn}{N} - \frac{4}{N} \sum_{x=0}^{s-1} \text{wt}(x),
\]

where \( \text{wt}(x) \) is the Hamming weight of \( x \).

Let us rephrase Proposition 2.4 a bit.

**Claim.** Let \( s = \sum_{i=0}^{t} N/2^{k_i} \), where \( 1 \leq k_0 < k_1 < \cdots < k_t \) are the locations of 1 in the binary representation of \( s \). Then \( I[\ell_n(s)] = \sum_{i=0}^{t} (k_i - 2i) 2^{1-k_i} \).

**Proof.** By induction on \( t \). For details see Appendix A.

**Example.** For \( s = N/2^k \) we get \( I[\text{And}_k] = I[\ell_n(N/2^k)] = k2^{1-k} \), demonstrating the tightness of Theorem 2.2.

**Corollary 2.6.** Let \( \mu = \sum_{i=0}^{\infty} 2^{-k_i} \), where \( 1 \leq k_0 < k_1 < \cdots \) are the locations of 1 in the binary representation of \( \mu \). Then \( I[\ell(\mu)] = \sum_{i=0}^{\infty} (k_i - 2i) 2^{1-k_i} \).

This leads to the following observation:

**Fact.** For any \( 0 \leq \mu \leq 1 \) we have

\[
I \left[ \ell \left( \frac{1}{2} \pm \frac{\mu}{4} \right) \right] = 2 \cdot 2^{-1} + \sum_{i=1}^{\infty} ((k_{i-1} + 2) - 2i) 2^{1-(k_{i-1}+2)} = 1 + \frac{1}{4} I[\ell(\mu)].
\]

(2)

**Example.** For \( \mu = \frac{2}{3} \) we have

\[
I[\ell(2/3)] = I[\ell(1/2 + 1/6)] = 1 + \frac{1}{4} I[\ell(2/3)],
\]

hence \( I[\ell(2/3)] = \frac{4}{3} \). By duality we have \( I[\ell(1/3)] = \frac{4}{3} \) as well.

**Remark.** Compare the bound \( \frac{2}{3} \log_2 3 \approx 1.05664 \) obtained for \( \mu = \frac{1}{3} \) from Theorem 2.2 to \( I[\ell(1/3)] = 4/3 \approx 1.33333 \) computed above. In fact, Theorem 2.2 is only tight when \( \mu \) is a power of two.

Four-thirds is actually the maximum influence attainable by any lexicographic function, as the following claim shows:

**Claim 2.7.** For all \( 0 \leq \mu \leq 1 \) we have \( I[\ell(\mu)] \leq \frac{4}{3} \).

**Proof.** Case 1. \( \mu < \frac{1}{4} \). Writing \( \mu = \sum_i 2^{-k_i} \), we have \( k_i \geq i + 3 \) for all \( i \geq 0 \). Moreover, we cannot have \( k_i = i + 3 \) for all \( i \) since \( \mu < \sum_{j=3}^{\infty} 2^{-j} = \frac{1}{4} \). Denote by \( j \) the minimal \( i \) for which \( k_i > i + 3 \). Now, by Corollary 2.6

\[
I[\ell(\mu)] = \sum_{i=0}^{\infty} (k_i - 2i) 2^{1-k_i} = \cdots \leq 1 + 2^{-j-2} (1 + j) = \frac{5}{4} < \frac{4}{3},
\]

where the full calculation is in Appendix A.

---

3To be read as a finite sum when \( \mu \) is a dyadic rational.
Case 2. \( \mu > \frac{3}{4} \). Then \( I[\ell \langle \mu \rangle] = I\left[\ell \langle \mu \rangle\right]^{\dagger} = I[\ell \langle 1 - \mu \rangle] \leq \frac{5}{4} < \frac{4}{3} \).

Case 3. \( \frac{1}{2} \leq \mu \leq \frac{3}{4} \). Since \( I[\ell \langle \mu \rangle] \) is a continuous function of \( \mu \), it has a maximum in the closed interval \( \left[\frac{1}{3}, \frac{2}{3}\right] \), obtained at \( \mu = \mu_0 \). If \( I[\ell \langle \mu_0 \rangle] = \frac{4}{3} + \epsilon \) for some \( \epsilon > 0 \) then for \( \mu_1 = 4\mu_0 - 2 \) we have

\[
I[\ell \langle \mu_1 \rangle] = 4(I[\ell \langle \mu_0 \rangle] - 1) = \frac{4}{3} + 4\epsilon > I[\ell \langle \mu_0 \rangle],
\]

contradicting either the choice of \( \mu_0 \) or one of the two previous cases. \( \square \)

Remark. We have \( I[\ell \langle \mu \rangle] = \frac{4}{3} \) for other values of \( \mu \) besides \( \mu = \frac{1}{3} \) and \( \mu = \frac{2}{3} \), e.g.,

\[
I[\ell \langle 7/12 \rangle] = I[\ell \langle (2 + 1/3) / 4 \rangle] = 1 + \frac{1}{4}I[\ell \langle 1/3 \rangle] = \frac{4}{3}.
\]

2.3 Disjoint composition

We now present the main tool we use to compute total influence and spectral entropy for our construction.

Definition. For two Boolean functions \( f_1 \) and \( f_2 \) on \( n_1 \) and \( n_2 \) variables, resp., define the Boolean functions on \( n = n_1 + n_2 \) variables \( f_1 \cap f_2 \) and \( f_1 \cup f_2 \) as

\[
(f_1 \cap f_2)(x_1, x_2, \ldots, x_n) = f_1(x_1, \ldots, x_{n_1}) \land f_2(x_{n_1+1}, \ldots, x_n);
(f_1 \cup f_2)(x_1, x_2, \ldots, x_n) = f_1(x_1, \ldots, x_{n_1}) \lor f_2(x_{n_1+1}, \ldots, x_n),
\]

and denote by \( \iota = \ell \langle 1/2 \rangle \) the one variable identity function.

Remark. The class of functions built using \( \iota, \cap, \text{ and } \cup \) is called read-once monotone formulas. By \( \ell \) every lexicographic function is a read-once monotone formulas.

As mentioned in the introduction, it was shown by [1, 9] that read-once formulas satisfy Conjecture [11] with the constant \( C \leq 10 \).

Definition. Let \( h : [0, 1] \to [0, 1] \) be the binary entropy function, defined by

\[
h(p) = -p \log_2 p - (1 - p) \log_2 (1 - p)
\]

for \( 0 < p < 1 \) and \( h(0) = h(1) = 0 \). We also make extensive use of its variant

\[
\tilde{h}(p) = h(4p(1 - p)) = h((1 - 2p)^2).
\]

Fact. Both \( h \) and \( \tilde{h} \) are symmetric about \( p = \frac{1}{2} \).

The following proposition is an easy corollary of [11] Lemmata 5.7 and 5.8. Alternatively, it is a special case of Lemma [3.1] in Section 3, which is an adaptation of [9] Proposition 3.2.

Proposition 2.8. Let \( f_1 \) and \( f_2 \) be Boolean functions and let \( p_i = \Pr[f_i] \) for \( i = 1, 2 \). Then

\[
I[f_1 \cap f_2] = p_2I[f_1] + p_1I[f_2];
\]

\[
H[f_1 \cap f_2] = p_2\left(H[f_1] - \tilde{h}(p_1)\right) + p_1\left(H[f_2] - \tilde{h}(p_2)\right) + \psi(p_1, p_2),
\]

where

\[
\psi(p, q) = \tilde{h}(pq) + 4pq(h(p) + h(q) - h(pq)).
\]

\footnote{If the maximum is attained multiple times, pick one arbitrarily.}
Remark. Via the De Morgan equality $f_1 \sqcup f_2 = (f_1^\dagger \cap f_2^\dagger)^\dagger$, this also yields

$$I[f_1 \sqcup f_2] = (1 - p_2) I[f_1] + (1 - p_1) I[f_2];$$
$$H[f_1 \sqcup f_2] = (1 - p_2) \left( H[f_1] - \hat{h} (p_1) \right) + (1 - p_1) \left( H[f_2] - \hat{h} (p_2) \right) + \psi (1 - p_1, 1 - p_2).$$

Proposition 2.8 gets simplified significantly when one of the functions is balanced, using the following observation (see Appendix A for the calculation):

Claim 2.9. Let $0 < p < 1$. Then $\psi (p, 1/2) = 2h (p)$.

Corollary 2.10. Let $f$ be a Boolean function and let $p = \Pr [f]$. Then

$$I[f \cap i] = \frac{1}{2} I[f] + p,$$
$$I[f \sqcup i] = \frac{1}{2} I[f] + 1 - p,$$

and

$$H[f \cap i] = H[f \sqcup i] = \frac{1}{2} H[f] - \frac{1}{2} \hat{h} (p) + 2h (p).$$

2.4 A first lower bound

We could use Claim 2.5 to compute the total influence of $G_m = \ell_2 m^{-1} (2/3)$, but we also need its spectral entropy, so we use its recursive definition and Corollary 2.10. Since we are interested in asymptotics, we prefer working directly with $G = \ell (2/3)$, which satisfies the “equation” $G = i \sqcup (i \cap G)$.

We already know $I[G] = I [\ell (2/3)] = \frac{4}{3}$, whereas for the entropy we have

$$H[G] = H[i \sqcup (i \cap G)] = \frac{1}{2} H[i \cap G] - \frac{1}{2} \hat{h} (1/3) + 2h (1/3)$$
$$= \frac{1}{2} \left( \frac{1}{2} H[G] - \frac{1}{2} \hat{h} (2/3) + 2h (2/3) \right) - \frac{1}{2} \hat{h} (1/3) + 2h (1/3)$$
$$= \frac{1}{4} H[G] + 3h (1/3) - \frac{3}{4} \hat{h} (2/3)$$
and we can solve for

$$H[G] = \frac{4}{3} \left( 3h (1/3) - \frac{3}{4} \hat{h} (1/9) \right) = 2 \log_2 3.$$

Note that it is possible to fully compute the total influence of $G_m$:

$$I[G_m] = 2 \Pr [G_m] = \frac{4}{3} (1 - 4^{-m})$$

and to write an expression for its spectral entropy:

$$H[G_{m+1}] = \sum_{i=0}^{m} 4^{-i} \left[ h \left( 2 \left( 1 - 4^{i-m} \right) /3 \right) + 2h \left( (1 - 4^{i-m}) /3 \right) \right.\right.$$ 
$$\left. - \frac{1}{4} \hat{h} \left( 2 \left( 1 - 4^{i-m} \right) /3 \right) - \frac{1}{2} \hat{h} \left( (1 - 4^{i-m}) /3 \right) \right],$$

but it is far easier to use the exponentially fast convergence $H[G_m] \xrightarrow{m \to \infty} H[G]$, rather than find an exact closed expression for $H[G_m]$.
Remark. Similarly, it is possible to exactly compute the total influence and spectral entropy of $\ell \langle p \rangle$ for any rational $p$. Indeed, every rational number has a recurrent binary representation, yielding linear equations in $I[\ell \langle p \rangle]$ and $H[\ell \langle p \rangle]$.

Approximating $I[\ell \langle p \rangle]$ and $H[\ell \langle p \rangle]$ for an irrational $p$ can be done, with exponentially decreasing error, via writing $p$ as a limit of a sequence of dyadic rationals (e.g., truncated binary representations of $p$).

Remark. In a certain sense, $\ell \langle 2/3 \rangle$ is the simplest infinite lexicographic function. Indeed, denote by $\lambda(p)$ the length of the recurring part in the binary expansion of a rational $p$. We have $\lambda(p) = 1$ if and only if $p$ is a dyadic rational. If $p$ is a dyadic multiple of $1/m$ for a positive odd integer $m$, then $\lfloor \log_2 m \rfloor \leq \lambda(p) \leq \text{ord}_m 2$, where $\text{ord}_m 2$ is the multiplicative order of 2 modulo $m$. In particular, $\lambda(p) \leq 2$ if and only if $p$ is a dyadic multiple of $1/3$.

Recall that $g_m$ is the conjunction of two functions: $\text{Or}_2(x_1, x_2)$ and $G_m(x_3, \ldots, x_{2m}, x_1)$. By Fact 2.4, these are almost independent since the shared variable $x_1$ has exponentially small influence on $G_m$.

When considering the limit object $g = \lim_m g_m$, the dependence disappears and we have $g = \text{Or}_2 \cap G = \ell \langle 3/4 \rangle \cap \ell \langle 2/3 \rangle$, so we can calculate its total influence and entropy using Proposition 2.8 (full details in Appendix A):

$$I_* = I[g] = I[\text{Or}_2 \cap G] = \frac{2}{3}I[\text{Or}_2] + \frac{3}{4}I[G] = \frac{2}{3} \cdot 1 + \frac{3}{4} \cdot \frac{4}{3} = \frac{5}{3},$$

$$H_* = H[g] = H[\text{Or}_2 \cap G] = \cdots = \frac{8}{3} + \log_2 3,$$

establishing our first lower bound:

**Theorem 2.11.** Any constant $C$ in Conjecture 1.1 satisfies

$$C \geq C_* = H_*/(I_* - 1) = 4 + 3\log_4 3 > 6.377443751,$$

even when restricted to monotone functions.

One technicality in the discussion above is that Proposition 1.2 supposedly only takes a finite function, so we cannot apply it directly to $g$, and we formally need to apply it to $g_m$ and let $m \to \infty$. The slight dependence on $x_1$ prevents us from computing the total influence and spectral entropy of $g_m$ via a direct application of Proposition 2.8, we can, however, consider the slight perturbation $\tilde{g}_m = \text{Or}_2 \cap G_m$, for which Proposition 2.8 gives $I[\tilde{g}_m] \approx 4/3$ and $H[\tilde{g}_m] \approx \frac{8}{3} + \log_2 3$.

Note that $\Pr[\tilde{g}_m] = \frac{3}{4}\Pr[G_m] = \frac{1}{2} (1 - 4^{-m})$, so $\tilde{g}_m$ is now slightly biased, and cannot be used in Proposition 1.2. To fix that, we only need to change a single entry of $\tilde{g}_m$ from false to true to get $g_m$ (or a different balanced function). Once again, Lemmata 5.1 and 5.2 of Section 5 tell us that such a minuscule modification has little effect on the entropy and total influence, which vanishes in the limit.

### 3 NAND on the run

In this section we review O’Donnell and Tan’s proof of Proposition 1.2 and apply it, in the biased case, to the function

$$\tau(x_1, x_2) = \neg(x_1 \land x_2).$$

That is, we can write $p = a/b$ for co-prime positive integers $a$ and $b = 2^\ell m$. 

---

10
3.1 Generalizing the composition method

Here is a sketch of the proof of Proposition 1.2 as done in [9, Lemma 5.1]. A sequence of balanced Boolean functions is built by recursively composing independent copies of $g$. Although both the total influence and entropy of the sequence grow to infinity, the limit of their entropy/influence ratios is $H[g] / (I[g] - 1)$. For these functions to be balanced, the base function $g$ ought to be balanced.

The same strategy could work for a biased function $g$ as well, assuming its satisfies a condition that we shall immediately see.

**Definition.** Fix an integer $n \geq 1$. A bias is a vector $\vec{\eta} = (\eta_1, \ldots, \eta_n)$ such that $-1 < \eta_i < 1$ for $i \in [n]$. Every bias $\vec{\eta}$ induces a product measure on \{true, false\}^n in which $E[x_i] = \eta_i$ for $i \in [n]$ and they are pairwise independent. Denote this distribution by $x \sim \vec{\eta}$.

Oftentimes we have $\eta_i = \eta$ for all $i \in [n]$, and we denote this by $x \sim \eta$.

**Example.** The zero bias induces the uniform distribution.

**Definition.** A Boolean function $f$ on $n$ variables is called $\eta$-balanced if $E_{x \sim \eta}[f] = \eta$.

**Example.** Balanced functions are 0-balanced.

**Example.** We seek a probability $0 \leq p \leq 1$ such that $\tau$ is $\eta$-balanced for $\eta = 1 - 2p$, i.e.,

$$p = \Pr_{x \sim \eta}[\tau(x)] = 1 - p^2.$$

The polynomial $x^2 + x - 1 = 0$ has exactly two real roots in $[-1, 1]$, which is

$$\Phi = \frac{\sqrt{5} - 1}{2} \approx 0.618034,$$

equalReciprocalGoldenRatio.

Thus, $\tau$ is $(1 - 2\Phi)$-balanced.

Two changes are required to make the proof of Proposition 1.2 work when the base function $g$ is $\eta$-balanced for $\eta \neq 0$:

1. Computing total influence and spectral entropy under a bias. This is provided by Lemma 3.1, adapted from [9, Proposition 3.2].

2. Instead of uniform input bits, we need to start from $\eta$-biased bits. These would be provided by lexicographic functions.

3.2 Biased Fourier analysis

Let us quickly recall biased Fourier analysis of Boolean functions.

**Definition.** Let $f$ be a Boolean function on $n$ variables. For $S \subseteq [n]$, denote by $\tilde{\chi}_S$ the $\vec{\eta}$-biased basis function

$$\tilde{\chi}_S = \prod_{i \in S} \frac{x_i - \eta_i}{\sqrt{1 - \eta_i^2}}$$

and denote by $\tilde{f}(S)$ the $\vec{\eta}$-biased Fourier coefficients of $f$

$$\tilde{f}(S) = \langle f, \tilde{\chi}_S \rangle = E_{x \sim \eta}[f(x) \tilde{\chi}_S(x)].$$
Since $\chi_S$ for $S \subseteq [n]$ form an orthonormal basis of $\{\text{true, false}\}^n$ under the $\eta$-biased product measure, we still have $\sum_S \tilde{f}(S)^2 = 1$ and we can speak of the $\eta$-biased spectral distribution $\tilde{p}_f(S) = \tilde{f}(S)^2$ of $f$, and consequently, the $\eta$-biased total influence $\tilde{I}[f]$ and $\eta$-biased spectral entropy $\tilde{H}[f]$.

**Example.** Let $\eta = 1 - 2\Phi = -\Phi^3$. Given that $\sqrt{1 - \eta^2} = 2 \sqrt{\Phi (1 - \Phi)} = 2\Phi^{3/2}$, the $\eta$-biased spectral distribution of $\tau$ is:

$$\tilde{p}_\tau(S) = \begin{cases} 
\Phi^6, & S = \emptyset; \\
4\Phi^5, & |S| = 1; \\
4\Phi^6, & |S| = 2,
\end{cases}$$

so its $\eta$-biased total influence and spectral entropy are (full details in Appendix A):

$$\tilde{I}[\tau] = \Phi^6 \cdot 0 + 2 \cdot 4\Phi^5 \cdot 1 + 4\Phi^6 \cdot 2 = 8 (\Phi^5 + \Phi^6) = 8\Phi^4 \approx 1.16718,$n

$$\tilde{H}[\tau] = \cdots = 8 (1 - 2\Phi) + 10 (4\Phi - 3) \log_2 \Phi \approx 1.77611.$n

### 3.3 Composition lemma

To simplify the notation of Lemma 3.1, we introduce a variant of the total influence and entropy definitions.

**Definition.** Let $f$ be a Boolean function and let $S \sim p_f$. The unbiased total influence and unbiased entropy of $f$, denoted respectively by $I^+[f]$ and $H^+[f]$, are

$$I^+[f] = E[|S| \mid S \neq \emptyset] = \frac{I[f]}{\text{Var}[f]},$$

$$H^+[f] = H[S \mid S \neq \emptyset] = \frac{H[f] - h(\text{Var}[f])}{\text{Var}[f]} = \frac{H[f] - \bar{h}(\text{Pr}[f])}{\text{Var}[f]},$$

where $\text{Var}[f] = \text{Pr}[S \neq \emptyset] = 1 - E[f]^2 = 4\text{Pr}[f] (1 - \text{Pr}[f])$.

**Example.** For $f = \text{And}_n$ we have

$$I^+[\text{And}_n] = \frac{I[\text{And}_n]}{\text{Var}[\text{And}_n]} = \frac{2n/N}{4/N (1 - 1/N)} = \frac{n}{2 (1 - 1/N)};$$

$$H^+[\text{And}_n] = H[S \mid S \neq \emptyset] = H[\text{Uniform}(N-1)] = \log_2 (N-1) \approx n.$n

**Lemma 3.1 ([9 Proposition 3.2]).** Let $F$ be a Boolean function on $k$ variables and let $g_1, \ldots, g_k$ be Boolean functions on $n$ variables. Define a Boolean function $f = F \circ (g_1, \ldots, g_k)$ on $kn$ variables by

$$f(x_1, \ldots, x_{kn}) = F(g_1(x_1, \ldots, x_n), g_2(x_{n+1}, \ldots, x_{2n}), \ldots, g_k(x_{(k-1)n+1}, \ldots, x_{kn})).$$

Then

$$I[f] = \sum_{i=1}^{k} \tilde{I}_i[F] I^+[g_i];$$

$$H[f] = \sum_{i=1}^{k} \tilde{I}_i[F] H^+[g_i] + \tilde{H}[F],$$

where $\tilde{p}_F$ is the $\eta$-biased spectral distribution of $F$ for the bias $\eta = (E[g_1], \ldots, E[G_k])$ and $\tilde{I}_i[F] = \text{Pr}_{S \sim \tilde{p}_F} [i \in S]$ for $i \in [k]$. In particular, when $g_i = g$ for all $i \in [k]$ we get

$$I[f] = \tilde{I}[F] I^+[g];$$

$$H[f] = \tilde{I}[F] H^+[g] + \tilde{H}[F].$$
3.4 A second lower bound

Define a sequence of functions \((F_m)_m\geq 0\) by \(F_0 = \ell \langle \Phi \rangle\) and \(F_{m+1} = \tau \circ (F_m, F_m, F_m)\) for all \(m \geq 0\). Recall that \(\tau\) is \((1 - 2\Phi)\)-balanced and thus \(\Pr[F_m] = \Phi\) for all \(m \geq 0\).

Via Lemma 3.1 we can compute the asymptotic entropy/influence ratio of \(F_m\) (see Appendix A for details):

Claim 3.2. \(\lim_{m \to \infty} \frac{H[F_m]}{I[F_m]} = \frac{H[\ell \langle \Phi \rangle] + (3 + 2\Phi) \hat{H}[\tau] - (4 + 2\Phi) \hat{h}(\Phi)}{I[\ell \langle \Phi \rangle]}\).

Theorem 3.3. Any constant \(C\) in Conjecture 1.1 satisfies \(C > 6.413846\).

Proof. Plug in Claim 3.2 the value \(\hat{H}[\tau] = 8 (1 - 2\Phi) + 10 (4\Phi - 3) \log_2 \Phi > 1.7761\) computed above, and the approximations \(I[\ell \langle \Phi \rangle] < 1.2976895, H[\ell \langle \Phi \rangle] > 2.4239395\).

Remark. Using any \((1 - 2p)\)-balanced base function \(g\) for \(0 < p < 1\), the same computation yields the lower bound:

\[ C \geq \frac{H[\ell \langle p \rangle]}{I[\ell \langle p \rangle]} + \frac{4p (1-p) \hat{H}[g] - \hat{I}[g] \hat{h}(p)}{I[\ell \langle p \rangle] \left( \hat{I}[g] - 4p (1-p) \right)}. \]

If \(g\) is balanced, i.e., \(p = \frac{1}{2}\), then this is plainly \(H[g] / (I[g] - 1)\), recovering Proposition 1.2.

4 To Infinity, and Beyond

In both Theorem 2.11 and Theorem 3.3 the notion of limit was used twice:

1. In creating an infinite lexicographic function \((\ell \langle 2/3 \rangle\) and \(\ell \langle \Phi \rangle\), respectively); and
2. When taking the asymptotic entropy/influence ratio for the sequence of functions defined by recursive composition.

In this section we use limits a countable number times for a superior construction.

4.1 Limit of limits

The basic step is inspired by the NAND function \(\tau\) of Section 3.

Fix a Boolean function \(\lambda\), and define a function \(\kappa\) using the equation \(\kappa = (\lambda \sqcap \kappa)\). Formally, we define a sequence \((\kappa_m)_m\geq 0\) of functions via \(\kappa_0 = \lambda\) and \(\kappa_{m+1} = (\lambda \sqcap \kappa_m)\) and let \(\kappa = \lim_{m \to \infty} \kappa_m\).

Proposition 4.1. Write \(p = \Pr[\lambda]\) and \(q = \Pr[\kappa]\). Then \(\kappa\) satisfies \(q = 1 / (1 + p)\) and

\[
\begin{align*}
\hat{I}^+ [\kappa] &= \hat{I}^+ [\lambda] / q \\
\hat{H}^+ [\kappa] &= \left( \hat{H}^+ [\lambda] + \frac{h(p)}{1 - p} \right) / q.
\end{align*}
\]

Proof. Apply Proposition 2.8 (see Appendix A for the computation).

Remark. Note that when \(\lambda\) is monotone, \(\kappa\) is monotone as well.
Fix some Boolean function $F_0$ and let $z = \Pr[F_0]$, $I = I[F_0]$ and $H = H[F_0]$. Define a sequence $(F_m)_{m \geq 1}$ using the equation $F_{m+1} = (F_m \cap F_{m+1})^\dagger$, and let $q_m = \Pr[F_m]$. By Proposition 4.1

\[
q_{m+1} = \frac{1}{1 + q_m}, \tag{3a}
\]

\[
I^+[F_{m+1}] = \frac{I^+[F_m]}{q_{m+1}}, \tag{3b}
\]

\[
H^+[F_{m+1}] = \frac{H^+[F_m] + \tilde{h}(q_m)}{1 - q_m} q_{m+1}. \tag{3c}
\]

These values naturally depend on the initial parameters $(z, I, H)$. Nevertheless, the sequence $q_m$ converges to the same limit $\lim_m q_m = \Phi$ regardless of the choice of $0 \leq z \leq 1$.

In fact, $q_m$ is a linear rational function of $z$, as the following claim states.

**Claim 4.2.** For all $m \geq 0$ we have

\[
q_m = \frac{b_{m-1}z + b_m}{b_mz + b_{m+1}}, \tag{4}
\]

where

\[
b_m = \frac{\Phi^m - (-\Phi)^m}{\sqrt{5}} \tag{5}
\]

is the $m$th Fibonacci number.

**Proof.** By induction on $m$ (see Appendix A for details). \qed

**Remark.** Binet’s formula [5] naturally extends the Fibonacci sequence to $\mathbb{Z}$. Note that for all $m \in \mathbb{Z}$ we have $b_{-m} = (-1)^{m+1} b_m$.

This can be used to define $q_m$ for negative $m$ as well. Note that for $m < 0$ we are no longer promised that $0 \leq q_m \leq 1$. In particular, $q_{-k}$ is undefined for $z = b_k/b_{k+1}$.

**Corollary.** For all $n \geq 0$ we have

\[
\pi_m = \frac{1}{zb_m + b_{m+1}}, \tag{6}
\]

where $\pi_m$ is the cumulative product $\pi_m = \prod_{k=1}^{m} q_k$. In particular,

\[
\frac{1}{2} \Phi^{m-1} \leq \frac{1}{b_{m+2}} = \frac{1}{b_{m} + b_{m+1}} \leq \pi_m \leq \frac{1}{b_{m+1}} \leq \Phi^{m-1}. \tag{7}
\]

For notational convenience, write $\pi_0 = \pi_0/q_0 = 1/z$ and $\pi_{-2} = \pi_{-1}/q_{-1} = 1/(1 - z)$. The next claim computes the entropy/influence ratio of $F_m$ via (3b) and (3c):

**Claim 4.3.** For all $m \geq 0$ we have

\[
\frac{H[F_m]}{I[F_m]} = \frac{1}{I} \left(H - \tilde{h}(z) + \beta_m(z) + z(1 - z)(\pi_{m-1} + \pi_{m-2}) \tilde{h}(q_m)\right),
\]

where $\beta_m : [0, 1] \to \mathbb{R}$ is

\[
\beta_m(z) = 4z(1 - z) \sum_{k=-2}^{m-3} \tilde{h}(q_{k+2}) \pi_k
\]

\[
= \begin{cases} 
0, & m = 0; \\
4zh(z), & m = 1; \\
4zh(z) + 4(1 - z)h(q_1) + 4z(1 - z) \sum_{k=0}^{m-3} h(q_{k+2}) \pi_k, & m \geq 2,
\end{cases}
\]
Proof. See Appendix A. □

Remark. Using (4) and (6) we can write $\beta_m$ as an explicit function of $z$ for all $m \in \mathbb{N}$.

Asymptotically the term $(\pi_{m-1} + \pi_{m-2}) \tilde{h} (q_m)$ vanishes, and we obtain

$$\lim_{m \to \infty} \frac{H [F_m]}{I [F_m]} = \frac{H - \tilde{h} (z) + \beta (z)}{I},$$

(8)

where

$$\beta (z) = \lim_{m \to \infty} \beta_m (z) = 4z (1 - z) \sum_{k=-2}^{\infty} h (q_{k+2}) \pi_k$$

$$= 4z \tilde{h} (z) + 4 (1 - z) \tilde{h} (q_1) + 4z (1 - z) \sum_{k=0}^{\infty} h (q_{k+2}) \pi_k.$$ 

By (7), $\beta_m \xrightarrow{m \to \infty} \beta$ exponentially fast to $\beta$, and thus $H [F_m]/I [F_m]$ converges very quickly as well. This can be seen visually in Figure 1.

**Theorem 4.4.** Any constant $C$ in Conjecture 1.1 satisfies

$$C \geq \beta (1/2) > 6.4547837,$$

even when restricted to monotone functions.

Proof. Select $F_0 = \iota$ with parameters $z = 1/2$, $I = 1$ and $H = 0$. Alternatively, since for $F_0 = \iota$ we have $F_1 = \ell (2/3)$, we could start with $F_0 = \ell \langle 2/3 \rangle$ and get the same bound. □

4.2 Afterthoughts

1. One may ask herself whether it would suffice to define a limit function $T$ using the equation $T = (T \cap T)^\dagger$. This is equivalent to the composition construction of Theorem 3.3 but we get a monotone function and are no longer limited to using $T_0 = \ell (\Phi)$. It is possible to show, via a computation quite similar to the one in previous subsection (see Appendix A), that

$$\lim_{m \to \infty} \frac{H [T_m]}{I [T_m]} = \frac{H - \tilde{h} (z) + \gamma (z)}{I}$$

for a function $\gamma$ slightly smaller than $\beta$. Picking $T_0 = \ell \langle \Phi \rangle$ is actually quite far from being optimal here; $T_0 = \iota$ or $T_0 = \ell \langle 5/8 \rangle$ yield a lower bound of $\approx 6.44539$, while $T_0 = \ell \langle 2/3 \rangle$ seems to attain the best lower bound $\approx 6.453111$ achievable using this method. This comes close, but is still less than $\beta (1/2)$, since $\gamma (2/3) < \beta (2/3)$.

2. Recall that the decision tree complexity of an infinite lexicographic function is just two bits. By simple induction, it can be shown that the average decision tree complexity of $F_m$ is $2^m d$ for all $m \geq 0$, where $d$ is the average decision tree complexity of $F_0$.

In particular, the average decision tree complexity of the sequence $(F_m)_{m \geq 0}$ is unbounded, and thus the construction is not subject to the upper bound on constant average depth decision trees of [11]. Each $F_m$ is still computable by a read-once formula, though.
3. The half circle shape of $\beta(z) = 4z(1-z)\sum_{k=-2}^{\infty} h(q_{k+2})\pi_k$ is mostly dictated by the variance term $4z(1-z)$, which is symmetric about $z = \frac{1}{2}$. One may thus guess that $\max_z \beta(z) = \beta(1/2)$. Surprisingly, the maximum of $\beta$ is obtained at $z^* \approx 0.50168825$, giving a meager improvement of $0.006\%$ over $\beta(1/2)$.

Nevertheless, it seems this cannot be used to improve the bound of Theorem 4.4. Indeed, any change in $z$ will have a negative effect on (8) by increasing both $I$ and $\tilde{h}(z)$, so to gain anything we need the initial function $F_0$ to provide a large entropy/influence ratio, which is what we were seeking all along.

Furthermore, any balanced function $F_0$ beating $\beta(1/2)$ must have $H > (I-1)\beta(1/2)$, so we could have used it in Proposition 1.2 directly!

4. The function $\beta$ can be simplified a bit further. Observe that

$$h(q_{m+1}) = h\left(\frac{1}{1+q_m}\right) = \frac{1}{1+q_m}\log_2 (1+q_m) + \frac{q_m}{1+q_m}\log_2 \frac{1+q_m}{q_m} = \log_2 (1+q_m) - q_m q_{m+1} \log_2 q_m = -\log_2 q_{m+1} - q_m q_{m+1} \log_2 q_m,$$

hence we can write

$$\beta(z) = 4zh(z) - 4z(1-z)\sum_{m=0}^{\infty} (\pi_m \log_2 q_{m+1} + \pi_{m+2} \log_2 q_m).$$
5 A Lipschitz-type condition for total influence and entropy

In this section we show that changing a single entry in a Boolean function has a negligible effect on its total influence and entropy.

Lemma 5.1. Let \( f \) and \( g \) be Boolean functions on \( n \) variables differing in a single entry \( x = x_0 \). Then \( |I[f] - I[g]| \leq 2n/N \).

Proof. We use an equivalent definition of total influence as the average sensitivity

\[
I[f] = E[S_f(x)] = \frac{1}{N} \sum_x S_f(x),
\]

where \( S_f(x) \) is the number of neighbors \( y \sim x \) in the Boolean cube such that \( f(x) \neq f(y) \). Indeed, we have

\[
|S_f(x) - S_g(x)| \leq \begin{cases} 
  n, & x = x_0; \\
  1, & x \sim x_0; \\
  0, & \text{otherwise.}
\end{cases}
\]

Thus,

\[
|I[f] - I[g]| = \frac{1}{N} \left| \sum_x S_f(x) - S_g(x) \right| \leq \frac{1}{N} \sum_x |S_f(x) - S_g(x)| \leq \frac{2n}{N}. \quad \square
\]

Remark. This is tight. Indeed, \( \text{Or}_n \) differs from the all-true function in a single entry \( x = \text{false}^n \) and \( I[\text{Or}_n] - I[\text{true}] = 2n/N - 0 = 2n/N \).

Lemma 5.2. Let \( f \) and \( g \) be Boolean functions on \( n \) variables differing in a single entry \( x = x_0 \). Then \( |H[f] - H[g]| \leq 12n/\sqrt{N} \).

Proof. This is trivial for \( n = 1 \) so assume \( n \geq 2 \). Also assume without loss of generality that the differing entry is \( x_0 = \text{false}^n \); that is, \( g = f + 2\delta \), where

\[
\delta(x) = \frac{1}{2} + \frac{1}{2} \text{Or}_n(x) = \begin{cases} 
  1, & x = x_0; \\
  0, & x \neq x_0.
\end{cases}
\]

Write \( a = \frac{N}{2} (f + g) = N (f + \delta) \), so \( f = \frac{1}{N} a - \delta \) and thus \( \hat{f}(S) = \frac{1}{N} (\hat{a}(S) - 1) \). Similarly we have \( g = \frac{1}{N} a + \delta \) and \( \hat{g}(S) = \frac{1}{N} (\hat{a}(S) + 1) \). In particular, we have

\[
\frac{\sum_{S \subseteq [n]} \hat{a}(S)}{N} = N \left( \frac{\sum_{S \subseteq [n]} \hat{f}(S)}{N} \right) - N = 0; \\
\frac{\sum_{S \subseteq [n]} \hat{a}(S)^2}{N} = \sum_{S \subseteq [n]} \left( N^2 \hat{f}(S)^2 + 2N \hat{f}(S) + 1 \right) = N (N - 1).
\]

Fourier coefficients of Boolean functions on \( n \) variables are known to reside in \( \{ \pm 2k/N \}_{k=0}^{N/2} \); thus the Fourier coefficients of \( a \) belong to \( \{ \pm (2k - 1) \}_{k=1}^{N/2} \). For \( k \in [N/2] \) let

\[
\Delta_k = \sum_{S : \hat{a}(S) = 2k-1} \text{sgn} \left( \hat{a}(S) \right) = |\{ S : \hat{a}(S) = 2k - 1 \}| - |\{ S : \hat{a}(S) = 1 - 2k \}|.
\]
and observe that
\[
\frac{N}{2} \sum_{k=1}^{N/2} \Delta_k (2k - 1) = \sum_{S \subseteq [n]} \hat{a} (S) = 0,
\] (9a)
\[
\frac{N}{2} \sum_{k=1}^{N/2} |\Delta_k| \leq \sum_{S \subseteq [n]} 1 = N,
\] (9b)
\[
\frac{N}{2} \sum_{k=1}^{N/2} |\Delta_k| (2k - 1)^2 \leq \sum_{S \subseteq [n]} \hat{a} (S)^2 = N \left( N - 1 \right),
\] (9c)
\[
\frac{N}{2} \sum_{k=1}^{N/2} |\Delta_k| (2k - 1) \leq \sqrt{\sum_{k=1}^{N/2} |\Delta_k| \sum_{k=1}^{N/2} |\Delta_k| (2k - 1)^2} \leq N \sqrt{N - 1} < N^{3/2},
\] (9d)
where Cauchy–Schwartz was used to derive (9d) from (9b) and (9c).

We express the difference of entropies in terms of \((\Delta_k)_{k \in [N/2]}\) (details in Appendix A):
\[
H[f] - H[g] = 8 \frac{N}{2} \sum_{k=2}^{N/2} \Delta_k k^2 \log_2 \frac{k}{k - 1} + \frac{8}{N^2} \sum_{k=2}^{N/2} \Delta_k (2k - 1) \log_2 (k - 1).
\] (10)

To bound the first term, note that the function \(\xi (x) = x^2 \log_2 (x/ (1 - x)) - (2x - 1)/ \ln 4\) is decreasing and positive for \(x > 1\), and \(\xi (2) = 4 - 3/ \ln 4 \approx 1.836\). Now
\[
\left| \sum_{k=2}^{N/2} \Delta_k k^2 \log_2 \frac{k}{k - 1} \right| = \left| \sum_{k=2}^{N/2} \Delta_k \left( \xi (k) + \frac{2k - 1}{\ln 4} \right) \right| \\
\leq \frac{1}{\ln 4} \left| \sum_{k=2}^{N/2} \Delta_k (2k - 1) \right| + \left| \sum_{k=2}^{N/2} \Delta_k \xi (k) \right| \\
\leq \frac{1}{\ln 4} |\Delta_1| + \sum_{k=2}^{N/2} |\Delta_k| \xi (2) \\
\leq \max \left\{ \frac{1}{\ln 4}, \xi (2) \right\} \sum_{k=1}^{N/2} |\Delta_k| \\
\leq \max \left\{ \frac{1}{\ln 4}, \xi (2) \right\} N = \xi (2) N < 2N.
\] (11)

To bound the second term, note that the function \(\zeta (x) = x \log_2 \frac{x-1}{2}\) is increasing, positive
and convex for \( x > 3, \) so

\[
2 \sum_{k=2}^{N/2} \Delta_k (2k-1) \log_2 (k-1) = 2 \sum_{k=2}^{N/2} \Delta_k \zeta (2k-1) \leq 2 \sum_{k=2}^{N/2} |\Delta_k| \zeta (2k-1)
\]

[Jensen’s inequality] \leq 2 \zeta \left( \sum_{k=2}^{N/2} |\Delta_k| (2k-1) \right)

\[
|\text{by (9d)}| \leq 2 \zeta (N^{3/2}) = 2N^{3/2} \log_2 \frac{N^{3/2} - 1}{2} < 2N^{3/2} (3n/2 - 1) = 3N^{3/2} n - 2N^{3/2}.
\]  

(12)

Combining (10), (11) and (12), we get

\[
|H[f] - H[g]| \leq \frac{8}{N^2} \cdot 2N + \frac{4}{N^2} \left( 3N^{3/2} n - 2N^{3/2} \right) \leq \frac{12n}{N},
\]

establishing the proof of the lemma.

\[ \square \]

**Corollary.** Let \( f \) and \( g \) be Boolean functions on \( n \) variables, and let \( \epsilon = \Pr [f(x) \neq g(x)] \). Then \( |I[f] - I[g]| \leq 2\epsilon n \) and \( |H[f] - H[g]| < 12\epsilon n \sqrt{N} \).

**Remark.** One may wonder how tight is Lemma 5.2, since the largest distance obtained from natural examples seems to be

\[
H[\text{Or}_n] - H[\text{true}] < 8 \left( n - 1 + 1/\ln 4 \right)/N - 0 < 8n/N.
\]

There exists, however, a algebraic construction in which the entropy difference is greater than \( 8/(3\sqrt{N}) \), so the lemma is tight up to the logarithmic factor \( n \):

Niho [8] considered functions from \( \GF(N) \) to \( \GF(2) \) of the form \( f(\alpha) = \text{Tr}(\alpha^r) \), where \( \text{Tr} : \GF(N) \to \GF(2) \) is the trace operator and \( r \) is some integer. These can naturally be interpreted as Boolean functions from \{false, true\} to \{false, true\}.

The case when \( n \equiv 0 \mod 4 \) and \( r = 2\sqrt{N} - 1 \) was analyzed in [8, Theorem 3-6]. The Fourier spectrum of the resulting function \( f \) has four possible values, as summarized in Table 2. Plugging these numbers in (10) shows that indeed \( H[f + 2\delta] - H[f] > 8/(3\sqrt{N}) \).

| Value of \( f \) | \(-1/\sqrt{N}\) | 0 | \(1/\sqrt{N}\) | \(2/\sqrt{N}\) |
|-----------------|-----------------|---|-----------------|-----------------|
| Multiplicity    | \(1/(N - \sqrt{N})\) | \(1/(N - \sqrt{N})\) | \(\sqrt{N}\) | \(1/(N - \sqrt{N})\) |

Table 2: Spectrum of Niho’s function \( f(\alpha) = \text{Tr}(\alpha^{2\sqrt{N} - 1}) \) for \( n \equiv 0 \mod 4 \).

### 6 Concluding Remarks and Open Problems

1. The key element repeating in all our constructions is lexicographic functions:

(a) In Theorem 2.11 they allowed us to create a balanced function \( \ell(2/3) \cap \ell(3/4) \) of positive entropy and small total influence that we could plug in Proposition 1.2

(b) In Theorem 5.3 using \( \ell(\Phi) \) we converted uniform bits to \((1 - 2\Phi)\)-biased bits that we could plug in the biased variant of Proposition 1.2 with the base function \( \tau \);

(c) In Theorem 4.4 the constructed sequence had \( \ell(2/3) \) as either its first or second member.
This is far from being a coincidence, as lexicographic functions are the minimizers of total influence (for a given bias) by Theorem 2.3. It seems plausible to attempt proving Conjecture 1.1 for the class of monotone Boolean functions (or perhaps all Boolean functions) by proving an upper bound on what can be done using lexicographic functions.

2. It is possible that we can improve on Theorems 3.3 and 4.4 by finding a base function $g$ better than $\tau$. Of course, $g$ should be $\eta$-balanced for some $-1 < \eta < 1$; that is, $\eta$ should be a fixed point of $E_g(\rho) = E_{\rho \sim \rho}[g]$. Preferably, $\eta$ should be an attractive fixed point of $E_g$, so $g$ needs to be a non-monotone function. By exhaustive search, we have determined that no function on $n \leq 4$ variables will do better than $\tau$.

Nevertheless, all constructions based on disjoint composition belong to the class of read-once formulas, and thus cannot provide a lower bound better than $10$.

3. One remaining gap worth closing is the asymptotic behavior of the Lipschitz constant for the spectral entropy. Recall that Niho’s function gave a lower bound of $\Omega(1/\sqrt{N})$, whereas the upper bound provided by Lemma 5.2 is $O(n/\sqrt{N})$. We believe the upper bound is not tight.

Acknowledgements

The author wishes to thank Nathan Keller for fruitful discussions and suggestions, and Ohad Klein for useful comments.

References

[1] Sourav Chakraborty, Raghav Kulkarni, Satyanarayana V. Lokam, and Nitin Saurabh, Upper bounds on Fourier entropy. In: Proc. of the 21st International Computing and Combinatorics Conference (COCOON), pp. 771-782, 2015.

[2] A.J. Bernstein, Maximally connected arrays on the $n$-cube. SIAM J. Appl. Math. 15, pp. 1485–1489, 1967.

[3] Bireswar Das, Manjish Pal, Vijay Visavaliya, The Entropy Influence conjecture revisited. Electronic Colloquium on Computational Complexity (ECCC) 18, #146, 2011.

[4] Ehud Friedgut and Gil Kalai, Every monotone graph property has a sharp threshold. Proc. Amer. Math. Soc. 124(10), pp. 2993–3002, 1996.

[5] L.H. Harper, Optimal assignment of numbers to vertices. J. Soc. Indust. Appl. Math. 12, pp. 131–135, 1964.

[6] Sergiu Hart, A note on the edges of the $n$-cube. Discr. Math. 14, pp. 157–163, 1976.

[7] John H. Lindsey II, Assignment of numbers to vertices. Amer. Math. Monthly 7, pp. 508–516, 1964.

[8] Yoji Niho, Multi-valued cross-correlation functions between two maximal linear recursive sequences. Ph.D. thesis, University of Southern California, Los Angeles, 1972.

[9] Ryan O’Donnell and Li-Yang Tan, A composition theorem for the Fourier Entropy Influence conjecture. In: Proc. of the 40th International Colloquium on Automata, Language and Programming (ICALP), pp. 780–791, 2013.
A Boring Calculations

A.1 From Section 1

Proof of Claim 1.3. By induction on $m$. It is true for $m = 1$ since
\[
\Pr[x_1 \mid x_1 \lor x_2] = \frac{\Pr[x_1 \land (x_1 \lor x_2)]}{\Pr[x_1 \lor x_2]} = \frac{1/2}{3/4} = \frac{2}{3}.
\]
Assuming correctness for $m$, we have
\[
\begin{align*}
\Pr[G_{m+1}(x_3, \ldots, x_{2m+2}, x_1) \mid x_1 \lor x_2] \\
= \Pr[x_3 \lor (x_4 \land G_m(x_5, \ldots, x_{2m+2}, x_1)) \mid x_1 \lor x_2] \\
= 1 - \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{2} \cdot \Pr[G_m(x_5, \ldots, x_{2m+2}, x_1) \mid x_1 \lor x_2]\right) \\
= 1 - \frac{1}{2} \left(1 - \frac{1}{2} \cdot \frac{2}{3}\right) = \frac{1}{2} + \frac{1}{6} = \frac{2}{3}.
\end{align*}
\]

A.2 From Section 2

Proof of Claim 2.5. By induction on $t$. It trivially holds for $t = 0$. Assuming correctness for $t$, let $s' = s - N/2^{k_0}$. By Proposition 2.4
\[
\begin{align*}
I[\ell_n(s)] &= \frac{2sn}{N} - 4 \sum_{x=0}^{s-1} \text{wt}(x) \\
&= \frac{2n}{N} \cdot \frac{N}{2^{k_0}} + \frac{2s'n}{N} - 4 \sum_{x=0}^{N/2^{k_0}-1} \text{wt}(x) - 4 \sum_{x=N/2^{k_0}}^{s-1} \text{wt}(x) \\
&= \frac{n}{2^{k_0-1}} - \frac{2s'n}{N} - 4 \sum_{x=0}^{s'-1} \left(1 + \text{wt}(x)\right) \\
&= \frac{n}{2^{k_0-1}} - \frac{n-k_0}{2^{k_0-1}} + I[\ell_n(s')] - 4s' \cdot \frac{1}{N} \\
&= k_0 2^{1-k_0} + \sum_{i=1}^{t+1} (k_i - 2i + 2) 2^{1-k_i} - \frac{4}{N} \sum_{i=1}^{t+1} N/2^{k_i} \\
&= \sum_{i=0}^{t+1} (k_i - 2i) 2^{1-k_i}.
\end{align*}
\]

Proof of Claim 2.7. (A full computation of $I[\ell(\mu)]$ for the case $\mu < \frac{1}{4}$) Recall that $k_i = i + 3$
for \( i < j \) and \( k_i \geq i + 4 \) for \( i \geq j \). Thus,

\[
\begin{align*}
I[\ell \langle \mu \rangle] & = \sum_{i=0}^{\infty} (k_i - 2i) 2^{1-k_i} \\
& = \sum_{i=0}^{j-1} (i + 3 - 2i) 2^{1-i-3} + \sum_{i=j}^{\infty} (k_i - 2i) 2^{-1-k_i} \\
& \leq \frac{1}{4} \sum_{i=0}^{j-1} (3 - i) 2^{-i} + 2 \sum_{i=j}^{\infty} (k_i - 2j) 2^{-k_i} \\
& \leq \frac{3}{4} \sum_{i=0}^{j-1} 2^{-i} - \frac{1}{4} \sum_{i=0}^{j-1} i2^{-i} + 2 \sum_{i=j+4}^{\infty} (i - 2j) 2^{-i} \\
& = \frac{3}{2} \left( 1 - 2^{-j} \right) - \frac{1}{4} \left( 2 - \sum_{i=j}^{\infty} i2^{-i} \right) + 2 \sum_{i=j+4}^{\infty} i2^{-i} - 4j \sum_{i=j+4}^{\infty} 2^{-i} \\
& = 1 - 3 \cdot 2^{-j-1} + \frac{1}{4} (j + 1) 2^{1-j} + 2 (j + 4 + 1) 2^{1-j-4} - 4j \cdot 2^{1-j-4} \\
& = 1 + 2^{-j-2} (j + 1) \leq \frac{5}{4} < \frac{4}{3},
\end{align*}
\]

where in the second to last equality we used the identity \( \sum_{i=j}^{\infty} i2^{-i} = (j + 1) 2^{1-j} \).

**Proof of Claim 2.9** We have

\[
2 - 2h(p/2) = 2 + 2 \frac{p}{2} \log_2 \frac{p}{2} + 2 \left( 1 - \frac{p}{2} \right) \log_2 \left( 1 - \frac{p}{2} \right)
= p \log_2 p + (2 - p) \log_2 (2 - p)
\]

so

\[
\psi(p, 1/2) = \tilde{h}(p/2) + 2p (h(p) + 1 - h(p/2))
= (1 - p)^2 \log_2 \left( 1 - p \right)^2 - (2p - p^2) \log_2 (2p - p^2)
+ 2ph(p) + p^2 \log_2 p + p (2 - p) \log_2 (2 - p)
= 2ph(p) - 2 (1 - p)^2 \log_2 (1 - p) - p (2 - p) \log_2 p + p^2 \log_2 p
= 2ph(p) - 2 (1 - p) ((1 - p) \log_2 (1 - p) + p \log_2 p) = 2h(p).
\]

A full computation of \( H_s \):

\[
H_s = H[g] = H[\text{Or}_2 \cap G]
= \frac{2}{3} (H[\text{Or}_2] - \tilde{h}(3/4)) + \frac{3}{4} (H[G] - \tilde{h}(2/3)) + \psi(3/4, 2/3)
= \frac{2}{3} (2 - h(1/4)) + \frac{3}{4} (2 \log_2 3 - h(1/9)) + \tilde{h}(1/2) + 2 (h(3/4) + h(2/3) - h(1/2))
= \frac{4}{3} + \left( 2 - \frac{2}{3} \right) h(1/4) + \frac{3}{4} \cdot \frac{8}{3} + 0 + 2h(1/3) - 2 \cdot 1
= \frac{4}{3} + \frac{4}{3} \left( 2 - \frac{3}{4} \log_2 3 \right) + 2 \left( \log_2 3 - \frac{2}{3} \right) = \frac{8}{3} + \log_2 3.
\]
A.3 From Section 3

A full computation of $H [\tau]$ for the bias $\eta = 1 - 2\Phi$:

$$H [\tau] = -\Phi^6 \log_2 \Phi^6 - 2 \cdot 4\Phi^5 \log_2 (4\Phi^5) - 4\Phi^6 \log_2 (4\Phi^6)$$
$$= -\Phi^5 (6\Phi \log_2 \Phi + 40 \log_2 \Phi + 16 + 8\Phi + 24\Phi \log_2 \Phi)$$
$$= -8\Phi^5 (2 + \Phi) - 10 (3\Phi + 4) \Phi^5 \log_2 \Phi$$
$$= 8 (1 - 2\Phi) + 10 (4\Phi - 3) \log_2 \Phi.$$  

Proof of Claim 3.2 Recall that earlier we computed $I [\tau] = 8\Phi^4$. By Lemma 3.1

$$I [F_{m+1}] = I [\tau] I^+ [F_m] = I [\tau] \frac{I [F_m]}{\text{Var} [F_m]}$$
$$= \frac{8\Phi^4}{4\Phi^3} \cdot I [F_m] = 2\Phi \cdot I [F_m].$$

Furthermore,

$$\frac{H^+ [F_{m+1}]}{I^+ [F_{m+1}]} = \frac{H [F_{m+1}] - \tilde{h} (\Phi)}{I [F_{m+1}]}$$
$$= \frac{I [\tau] \cdot H^+ [F_m]}{I [\tau] \cdot I^+ [F_m]} + \frac{\hat{H} [\tau] - \tilde{h} (\Phi)}{I [F_{m+1}]}.$$

hence

$$\frac{H^+ [F_m]}{I^+ [F_m]} = \frac{H^+ [F_0]}{I^+ [F_0]} + \left( H [\tau] - \tilde{h} (\Phi) \right) \frac{1}{I [F_{m+1}]}$$
$$= \frac{H^+ [\ell (\Phi)]}{I^+ [\ell (\Phi)]} + \frac{\hat{H} [\tau] - \tilde{h} (\Phi)}{I [\ell (\Phi)]} \cdot \sum_{k=1}^{m} \frac{1}{(2\Phi)^k}.\]$$

Asymptotically the $(2\Phi)^{-m}$ term disappears, and we have

$$\lim_{m \to \infty} \frac{H^+ [F_m]}{I^+ [F_m]} = \lim_{m \to \infty} \frac{H^+ [F_m] - \tilde{h} (\Phi)}{I^+ [F_m]} = \lim_{m \to \infty} \frac{H^+ [F_m]}{I^+ [F_m]} - 0 = \frac{H^+ [\ell (\Phi)]}{I^+ [\ell (\Phi)}} + \frac{H [\tau] - \tilde{h} (\Phi)}{(2\Phi - 1) I [\ell (\Phi)]}$$
$$= \frac{H [\ell (\Phi)] - \tilde{h} (\Phi)}{I [\ell (\Phi)]} + \frac{H [\tau] - \tilde{h} (\Phi)}{(2\Phi - 1) I [\ell (\Phi)]} = \frac{H [\ell (\Phi)] + (3 + 2\Phi) \hat{H} [\tau] - (4 + 2\Phi) \tilde{h} (\Phi)}{I [\ell (\Phi)]}.$$

A.4 From Section 4

Proof of Proposition 4.1 We have

$$\Pr [\kappa] = 1 - \Pr [\lambda \cap \kappa] = 1 - p \Pr [\kappa],$$

so we can solve for

$$q = \Pr [\kappa] = \frac{1}{1 + p}.$$
Next, by Proposition 2.8 and duality we have
\[ \mathbf{I} [\kappa] = \mathbf{I} [\lambda \cap \kappa] = q \mathbf{I} [\lambda] + p \mathbf{I} [\kappa] \]
and we can solve for \( \mathbf{I} [\kappa] = q \mathbf{I} [\lambda] / (1 - p) \), yielding
\[ \mathbf{I}^+ [\kappa] = \frac{\mathbf{I} [\kappa]}{4q(1 - q)} = \frac{q \mathbf{I} [\lambda]}{4q \cdot pq \cdot (1 - p)} = \frac{1}{q} \mathbf{I}^+ [\lambda] \cdot \]
Last, we compute \( \mathbf{H} [\kappa] \). Note that since \( q = 1 - pq \) we have \( h (pq) = h (q) \) and \( \tilde{h} (pq) = \tilde{h} (q) \), so
\[ \psi (p, q) = \tilde{h} (pq) + 4pq (h (p) + h (q) - h (pq)) \]
\[ = \tilde{h} (q) + 4pqh (p) \cdot \]
Now, by Proposition 2.8
\[ \mathbf{H} [\kappa] = \mathbf{H} [\lambda \cap \kappa] = q \left( \mathbf{H} [\lambda] - \tilde{h} (p) \right) + p \left( \mathbf{H} [\kappa] - \tilde{h} (q) \right) + \psi (p, q) \]
\[ = p \mathbf{H} [\kappa] + q \left( \mathbf{H} [\lambda] - \tilde{h} (p) \right) - ph (q) + \tilde{h} (q) + 4pqh (p) \]
so we can solve for
\[ \mathbf{H} [\kappa] = \frac{q \left( \mathbf{H} [\lambda] - \tilde{h} (p) \right) + (1 - p) \tilde{h} (q) + 4pqh (p)}{1 - p} \]
\[ = \frac{1 - p}{p} \left( \mathbf{H} [\lambda] - \tilde{h} (p) + 4ph (p) \right) + \tilde{h} (q) \cdot \]
yielding
\[ \mathbf{H}^+ [\kappa] = \frac{\mathbf{H} [\kappa] - \tilde{h} (q)}{4q(1 - q)} = \frac{\mathbf{H} [\lambda] - \tilde{h} (p) + 4ph (p)}{4(1 - p) \cdot pq} \]
\[ = \frac{\mathbf{H} [\lambda] - \tilde{h} (p)}{4p(1 - p) q + (1 - p) q} \]
\[ = \left( \mathbf{H}^+ [\lambda] + \frac{h (p)}{1 - p} \right) / q. \]
\]
Proof of Claim 4.2: For \( m = 0 \) indeed \( q_m = (1 \cdot z + 0) / (0 \cdot z + 1) = z \). Now, assuming the claim holds for \( q_m \),
\[ q_{m+1} = \frac{1}{1 + q_m} = \frac{b_m z + b_{m+1}}{(b_m z + b_{m+1}) (1 + q_m)} \]
\[ = \frac{b_m z + b_{m+1}}{b_m z + b_{m+1} + b_{m-1} z + b_m} \]
\[ = \frac{b_m z + b_{m+1}}{b_m z + b_{m+1} + b_{m+2}}. \]
\]
Proof of Claim 4.3: By (3b) and (3c),
\[ \frac{\mathbf{H}^+ [F_{k+1}]}{\mathbf{I}^+ [F_{k+1}]} = \frac{\mathbf{H}^+ [F_k]}{\mathbf{I}^+ [F_k]} + \frac{h (q_k)}{(1 - q_k) \mathbf{I}^+ [F_k]} \]
\[ = \frac{\mathbf{H}^+ [F_k]}{\mathbf{I}^+ [F_k]} + \frac{h (q_k) \pi_k}{(1 - q_k) \mathbf{I}^+ [F_0]} \]
\[ = \frac{\mathbf{H}^+ [F_k]}{\mathbf{I}^+ [F_k]} + \frac{h (q_k) \pi_k - 2}{\mathbf{I}^+ [F_0]}. \]
\]
thus

\[ \begin{align*}
\mathbf{H}^{+} [F_m] - \mathbf{I}^{+} [F_m] &= \mathbf{H}^{+} [F_0] + \sum_{k=0}^{m-1} \mathbf{H}^{+} [F_{k+1}] - \mathbf{I}^{+} [F_k] \\
&= \mathbf{H}^{+} [F_0] + \sum_{k=-2}^{m-3} \frac{h(q_{k+2}) \pi_k}{\mathbf{I}^{+} [F_0]}
\end{align*} \]

and

\[ \begin{align*}
\mathbf{H} [F_m] - \mathbf{I} [F_m] &= \mathbf{H}^{+} [F_m] + \mathbf{I}^{+} [F_m] \\
&= \mathbf{H}^{+} [F_m] + \frac{\tilde{h}(q_m)}{4q_m(1-q_m)I^{+} [F_m]} \\
&= \mathbf{H}^{+} [F_0] + \frac{\tilde{h}(q_m) \pi_{m-2}}{4q_m I^{+} [F_0]} + \sum_{k=-2}^{m-3} \frac{h(q_{k+2}) \pi_k}{I^{+} [F_0]} \\
&= \frac{H - \tilde{h}(z)}{I} + \frac{4z(1-z)}{I} \left( \frac{1 + q_m - 1}{4} \tilde{h}(q_m) \pi_{m-2} + \sum_{k=-2}^{m-3} h(q_{k+2}) \pi_k \right) \\
&= \frac{1}{I} \left( H - \tilde{h}(z) + z(1 - z)(\pi_{m-1} + \pi_{m-2}) \tilde{h}(q_m) + \beta_m(z) \right). \quad \Box
\]
A.4.1 Lower bound obtained from $T = (T \cap T)^\dagger$

First we prove an analogue of Proposition 4.1:

Given a Boolean function $\lambda$, define $\kappa = (\lambda \cap \lambda)^\dagger$. Writing $p = \Pr[\lambda]$ and $q = \Pr[\kappa]$ we have $q = 1 - p^2$. By Proposition 2.8 we have

\[
\begin{align*}
I[\kappa] &= I[\lambda \cap \lambda] = 2pI[\lambda], \\
H[\kappa] &= H[\lambda \cap \lambda] = 2p \left( H[\lambda] - \tilde{h}(p) \right) + \psi(p, p) \\
&= 2p \left( H[\lambda] - \tilde{h}(p) \right) + \tilde{h}(p^2) + 4p^2 \left( 2h(p) - h(p^2) \right) \\
&= 2p \left( H[\lambda] - \tilde{h}(p) \right) + \tilde{h}(q) + 8p^2 h(p) - 4p^2 h(q).
\end{align*}
\]

Let $\tilde{C}_m = \left( H[T_m] - \tilde{h}(t_m) \right) / I[T_m]$, where $t_m = \Pr[T_m]$. Now

\[
\tilde{C}_{m+1} - \tilde{C}_m = \frac{8p^2 h(p) - 4p^2 h(q)}{I[T_{m+1}]} = \frac{4ph(p)}{I[T_m]} - \frac{4(1-q)h(q)}{I[T_{m+1}]}
\]

so

\[
\lim_{m \to \infty} \frac{H[T_m]}{I[T_m]} = \tilde{C}_0 + \sum_{k=0}^{\infty} \left( \tilde{C}_{k+1} - \tilde{C}_k \right) = \frac{H[T_0] - \tilde{h}(t_0) + \gamma(t_0)}{I[T_0]},
\]

where

\[
\gamma(z) = 4 \sum_{k=0}^{\infty} t_k h(t_k) - 4 \sum_{k=1}^{\infty} \frac{(1-t_k) h(t_k)}{2^k \prod_{i=0}^{k-1} t_i}
\]

\[
= 4zh(z) + 4 \sum_{k=1}^{\infty} \frac{(2t_k - 1) h(t_k)}{2^k \prod_{i=0}^{k-1} t_i}.
\]
A.5 From Section 5

The difference in entropies is:

\[ H[f] - H[g] = \sum_S \hat{g}(S)^2 \log_2 \left( \hat{g}(S)^2 \right) - \hat{f}(S)^2 \log_2 \left( \hat{f}(S)^2 \right) \]

\[ = \frac{2}{N^2} \sum_S (N\hat{g}(S))^2 \log_2 |\hat{g}(S)| - (N\hat{f}(S))^2 \log_2 |\hat{f}(S)| \]

\[ = \frac{2}{N^2} \sum_S \left[ (\hat{a}(S) + 1)^2 \log_2 (|\hat{a}(S) + 1|/N) \right. \]

\[ - (\hat{a}(S) - 1)^2 \log_2 (|\hat{a}(S) - 1|/N) \]

\[ = \frac{2}{N^2} \sum_S \left[ (1 + \hat{a}(S))^2 \log_2 (|1 + \hat{a}(S)|/N) \right. \]

\[ - (1 - \hat{a}(S))^2 \log_2 (|1 - \hat{a}(S)|/N) \]

\[ = \frac{2}{N^2} \sum_{k=1}^{N/2} \Delta_k \left[ (2k)^2 \log_2 (2k/N) - (2k - 2)^2 \log_2 (2(k - 1)/N) \right] \]

\[ = \frac{8}{N^2} \sum_{k=1}^{N/2} \Delta_k \left[ k^2 \log_2 k - (k - 1)^2 \log_2 (k - 1) \right] \]

\[ - \frac{8 (n - 1)}{N^2} \sum_{k=1}^{N/2} \Delta_k \left[ k^2 - (k - 1)^2 \right] \]

[by (9a)]

\[ = \frac{8}{N^2} \sum_{k=1}^{N/2} \Delta_k \left[ k^2 \log_2 k - (k - 1)^2 \log_2 (k - 1) \right] \]

[note the index k]

\[ = \frac{8}{N^2} \sum_{k=2}^{N/2} \Delta_k \left[ k^2 \log_2 \frac{k}{k - 1} \right] + \frac{8}{N^2} \sum_{k=2}^{N/2} \Delta_k \left( 2k - 1 \right) \log_2 (k - 1). \]