



# When Majority Fails: Tight Bounds for Correlation Distillation Conjectures

Pritish Kamath  
pritch@alum.mit.edu

Ravi Kumar  
ravi.k53@gmail.com

Pasin Manurangsi  
pasin@google.com

Google Research

## Abstract

We study two conjectures posed in the analysis of Boolean functions  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , in both of which, the *Majority* function plays a central role: the “Majority is Least Stable” (Benjamini et al., 1999) and the “Non-Interactive Correlation Distillation for Erasures” (Yang, 2004; O’Donnell and Wright, 2012).

While both conjectures have been refuted in their originally stated form, we obtain a nearly tight characterization of the noise parameter regime in which each of the conjectures hold, for all  $n \geq 5$ . Whereas, for  $n = 3$ , both conjectures hold in all noise parameter regimes. We state refined versions of both conjectures that we believe captures the spirit of the original conjectures.

## 1 Introduction

The *Non-Interactive Correlation Distillation (NICD)* problem, parameterized by a joint distribution  $P$  over  $\mathcal{X} \times \mathcal{Y}$ , is to maximize for  $(\mathbf{x}, \mathbf{y}) \sim P^{\otimes n}$ , the correlation  $\mathbb{E}[f(\mathbf{x}) \cdot g(\mathbf{y})]$ , subject to  $\mathbb{E}[f(\mathbf{x})] = \mathbb{E}[g(\mathbf{y})] = 0$  for functions  $f : \mathcal{X}^n \rightarrow [-1, 1]$  and  $g : \mathcal{Y}^n \rightarrow [-1, 1]$ . Let  $\rho_n^*(P)$  denote this optimal value, and let  $\rho^*(P) := \sup_n \rho_n^*(P)$ . Informally, this corresponds to a two-player setting, where Alice receives a sequence  $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{X}^n$ , Bob receives a sequence  $\mathbf{y} = (y_1, \dots, y_n) \in \mathcal{Y}^n$ , where each  $(x_i, y_i) \stackrel{\text{i.i.d.}}{\sim} P$ . Each player is required to return a single bit, that is marginally uniform, while maximizing the probability that their returned bits agree.<sup>1</sup>

Specifically, we consider NICD for the following distributions (defined in Table 1):

1. the *Doubly Symmetric Binary Source (DSBS)* defined over  $\{-1, 1\} \times \{-1, 1\}$ , and
2. the *Symmetric Binary Erasure Source (SBES)* defined over  $\{-1, 1\} \times \{-1, 1, \star\}$ .

### 1.1 DSBS and the “Majority is Least Stable (MLS)” Conjecture

The *Doubly Symmetric Binary Source*  $\text{DSBS}_\rho$ , parameterized by a given  $\rho \in [0, 1]$ , is a joint distribution  $(x, y)$  over  $\{-1, 1\} \times \{-1, 1\}$  given in Table 1(a), where the marginal distributions are uniform, and the correlation between corresponding bits is  $\mathbb{E}[xy] = \rho$ . It is folklore (see Section 1.4) that the dictator functions, namely  $f(\mathbf{x}) = \text{Dict}(\mathbf{x}) := x_1$  and  $g(\mathbf{y}) = \text{Dict}(\mathbf{y}) := y_1$ , realize the optimal value of  $\rho^*(\text{DSBS}_\rho)$  for all values of  $\rho \in [0, 1]$ .

<sup>1</sup>The range of  $f$  and  $g$  is  $[-1, 1]$ , which can be interpreted as Alice returning a randomized rounding of  $f(\mathbf{x})$  to  $\{-1, 1\}$ , and similarly Bob returning a randomized rounding of  $g(\mathbf{y})$ .

	$Y = 1$	$Y = -1$
$X = 1$	$\frac{1+\rho}{4}$	$\frac{1-\rho}{4}$
$X = -1$	$\frac{1-\rho}{4}$	$\frac{1+\rho}{4}$

(a) DSBS $_{\rho}$

	$Y = 1$	$Y = \star$	$Y = -1$
$X = 1$	$\frac{p}{2}$	$\frac{1-p}{2}$	0
$X = -1$	0	$\frac{1-p}{2}$	$\frac{p}{2}$

(b) SBES $_{\rho}$

Table 1: (a) Doubly Symmetric Binary Source and (b) Symmetric Binary Erasure Source.

The special case when  $f = g$  is often studied through the notion of *noise stability*,  $\text{Stab}_{\rho}(f) := \mathbb{E}_{(\mathbf{x}, \mathbf{y})}[f(\mathbf{x})f(\mathbf{y})]$  for  $(\mathbf{x}, \mathbf{y}) \sim \text{DSBS}_{\rho}^{\otimes n}$  (see O’Donnell, 2014). So, in other words, the dictator function is the most noise stable.

Benjamini et al. (1999) studied the noise stability of *linear threshold functions* (LTF), namely, functions of the form  $f(\mathbf{x}) = \text{sign}(\sum_{i=1}^n w_i x_i)$ , and they posed the question of which (unbiased) LTF has the least noise stability. They conjectured that for odd  $n$ ,  $\text{Maj}_n(\mathbf{x}) := \text{sign}(\sum_{i=1}^n x_i)$  has the least noise stability among unbiased LTFs, for any  $\rho \in (0, 1)$ . This *MLS* conjecture was also highlighted by Filmus et al. (2014).

However, as stated, the MLS conjecture was shown to be false. Gopi (2013) obtained a counterexample for all odd  $n \geq 5$ , which Biswas and Sarkar (2023) later proved to be minimal by showing that the conjecture holds for  $n \leq 3$ .<sup>2</sup> Gopi also showed a partial positive result: for any fixed  $n$ , the conjecture holds in a small neighborhood of  $\rho$  around 1.

## 1.2 SBES and the “NICD for Erasures (NICD-E)” Conjecture

Yang (2004) studied the NICD problem for the *symmetric binary erasure source*. Parameterized by  $p \in [0, 1]$ , SBES $_{\rho}$  is a joint distribution  $(x, y)$  over  $\{-1, 1\} \times \{-1, 1, \star\}$  given in Table 1(b), where the marginal distribution of  $x$  is uniform, and  $y$  is obtained by “erasing”  $x$  with probability  $1 - p$ .

It is easy to see that for any choice of  $f : \mathcal{X}^n \rightarrow [-1, 1]$ , the optimal  $g^* : \mathcal{Y}^n \rightarrow [-1, 1]$  that maximizes  $\mathbb{E}[f(\mathbf{x}) \cdot g(\mathbf{y})]$  is given by  $g^*(\mathbf{y}) := \text{sign}(\mathbb{E}[f(\mathbf{x})|\mathbf{y}])$ . And thus, maximizing  $\mathbb{E}[f(\mathbf{x}) \cdot g(\mathbf{y})]$  becomes equivalent to maximizing  $\Phi_p(f)$ , where

$$\Phi_p(f) := \mathbb{E}_{\mathbf{y}} [|\mathbb{E}[f(\mathbf{x}) | \mathbf{y}]|] \quad \text{subject to} \quad \mathbb{E}[f(\mathbf{x})] = 0.$$

Note that it does not follow that  $\mathbb{E}[g^*(\mathbf{y})] = 0$ . Following prior literature, we ignore this constraint and instead only require that  $\mathbb{E}[f(\mathbf{x})] = 0$ . Although, note that for *odd*  $f$ , namely when  $f(\mathbf{x}) = -f(-\mathbf{x})$  holds for all  $\mathbf{x}$ , it follows that  $g^*$  is also unbiased.

For the low-erasure regime ( $p > 1/2$ ), O’Donnell and Wright (2012) showed that the dictator function  $\text{Dict}(\mathbf{x}) := x_1$  maximizes  $\Phi_p(f)$ . Furthermore, for the high-erasure regime ( $p < 1/2$ ) and odd  $n$ , they conjectured that the Majority function  $\text{Maj}_n$  maximizes  $\Phi_p(f)$ ; this *NICD-E* conjecture was also highlighted by Filmus et al. (2014).

Recently, Ivanisvili and Xie (2025) presented a counterexample<sup>3</sup> refuting this conjecture for  $n = 5$  and  $p = 0.4$ . Similar to (Gopi, 2013), they also provide a partial positive result showing that for all odd  $n$ , the conjecture holds in a small neighborhood of  $p$  around 0.

<sup>2</sup>As mentioned by Biswas and Sarkar (2023), the counterexample for  $n = 5$  were also discovered by Noam Berger, Daniel Kane, Steven Heilman, and Jain (2017).

<sup>3</sup>Interestingly, they discovered the counterexample with help of a large language model. While their stated counterexample  $\text{sign}(x_1 - 3x_2 + x_3 - x_4 + 3x_5)$  appears highly asymmetric, it is equivalent (up to permutations and sign flips) to the counterexample discovered for the MLS conjecture, namely  $\text{sign}(2x_1 + 2x_2 + x_3 + x_4 + x_5)$ .

### 1.3 Our Contributions

In this work, we provide a unified understanding of both these seemingly unrelated conjectures, and nearly-tight characterization of the noise parameter regimes under which the respective conjectures are tight. Our main contributions are summarized as follows:

**When the conjectures are false?** We show, in [Theorem 9](#), that the counterexample for the MLS conjecture by [Gopi \(2013\)](#) also refutes the NICD-E conjecture for all  $n \geq 5$ . Furthermore, we show that this example refutes the MLS conjecture for all  $\rho \in (0, 1 - \frac{8+o(1)}{n^2})$  and refutes the NICD-E conjecture for all  $p \in (\frac{4+o(1)}{n^2}, \frac{1}{2})$ .

**When the conjectures are true?** We complement the counterexamples by obtaining nearly matching positive bounds in [Theorem 10](#): the MLS conjecture holds for  $\rho \in (1 - \frac{8-o(1)}{n^2}, 1)$  and the NICD-E conjecture holds for  $p \in (0, \frac{2-o(1)}{n^2})$ . Our result for the MLS conjecture is in fact stronger in that we show the MLS conjecture holds for all unate<sup>4</sup> functions, of which LTFs are a special case. Additionally, the NICD-E conjecture when restricted to unate functions holds for all  $p \in (0, \frac{4-o(1)}{n^2})$ . Using our techniques, we also show that the conjectures hold for  $n = 3$  for the entire regime of parameters ([Theorem 11](#)).<sup>5</sup>

### 1.4 Related Work & Discussion

For any distribution  $P$  over  $\mathcal{X} \times \mathcal{Y}$ , the maximal correlation coefficient  $\rho(P)$ , studied by [Hirschfeld \(1935\)](#); [Gebelein \(1941\)](#); [Rényi \(1959\)](#) is the maximum value of  $\mathbb{E}[f(x)g(y)]$  subject to  $\mathbb{E}[f(x)] = \mathbb{E}[g(y)] = 0$  and  $\mathbb{E}[f(x)^2] = \mathbb{E}[g(y)^2] = 1$  over  $f : \mathcal{X} \rightarrow \mathbb{R}$  and  $g : \mathcal{Y} \rightarrow \mathbb{R}$ , where the expectation is over  $(x, y) \sim P$ . The difference between  $\rho^*(P)$  and  $\rho(P)$  is that the range of  $f$  and  $g$  is  $[-1, 1]$  in the former, but  $\mathbb{R}$  in the latter.<sup>6</sup> Thus, it follows that  $\rho^*(P) \leq \rho(P)$ . [Witsenhausen \(1975\)](#) provided a construction of  $f$  and  $g$  to show that

$$\frac{2}{\pi} \arcsin(\rho(P)) \leq \rho^*(P) \leq \rho(P). \quad (1)$$

It is challenging to determine  $\rho^*(P)$ , and is not known exactly even for some simple distributions  $P$ . Only recently, it was shown that  $\rho^*(P)$  is even computable ([Ghazi et al., 2016](#)); subsequently, this was extended to correlations over larger alphabet size ([De et al., 2017, 2018](#); [Ghazi et al., 2018](#)).

**DSBS $_{\rho}$ .** The maximal correlation  $\rho(\text{DSBS}_{\rho}) = \rho$  and thus, we have  $\rho^*(\text{DSBS}_{\rho}) \leq \rho$ . Moreover the dictator functions achieve equality thereby establishing their optimality.

Witsenhausen’s construction underlying (1) is identical to  $f = g = \text{Maj}_n$ , in the limit as  $n \rightarrow \infty$ , it holds that  $\text{Stab}_{\rho}(\text{Maj}_n)$  approaches  $\frac{2}{\pi} \arcsin \rho$  from above ([O’Donnell, 2003](#)), and hence  $\text{Stab}_{\rho}(\text{Maj}_n) \geq \frac{2}{\pi} \arcsin \rho$  for all  $n$ .

Somewhat unrelated, [Mossel et al. \(2010\)](#) introduced the *invariance principle* and showed that for functions  $f$  with “low influence”, it holds that  $\text{Stab}_{\rho}(f) \leq \text{Stab}_{\rho}(\text{Maj}_n) + o_n(1)$ , i.e., the MLS conjecture holds for low-influence functions.

Given all the results thus far, a reasonable refined version of the MLS conjecture that we believe captures the spirit of the original formulation is as follows. This formulation was also stated by [Gopi \(2013\)](#).

<sup>4</sup>A Boolean function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  is *unate* if it is monotone, either increasing or decreasing, in each  $x_i$ .

<sup>5</sup>[Biswas and Sarkar \(2023\)](#) already showed this for the MLS conjecture; see the proof of [Theorem 11](#).

<sup>6</sup>Another difference is that in  $\rho(P)$ , the domain of  $f$  is a single  $x$  (similarly for  $g$ ), however,  $\rho(P)$  exhibits a “tensorization property”, namely,  $\rho(P) = \rho(P^{\otimes n})$ .

**Conjecture 1** (Refined MLS (Gopi, 2013)). *For all  $\rho \in (0, 1)$ , odd  $n$ , and unbiased LTF  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , it holds that  $\text{Stab}_\rho(f) \geq \frac{2}{\pi} \arcsin(\rho)$ .*

*Remark.* It is tempting to conjecture that the MLS conjecture holds for all unate functions, not just LTFs, given that our positive result shows this for  $\rho \in (1 - O(1/n^2), 1)$ . However, this turns out to be false. Benjamini et al. (1999) showed that the noise stability of the so-called Tribes function (a balanced monotone function) approaches 0 as  $n \rightarrow \infty$  for  $\rho = 1 - \Omega(1/\log n)$ .

**SBES<sub>p</sub>.** It is easy to see that the maximal correlation coefficient  $\rho(\text{SBES}_p) = \sqrt{p}$ . In this case, Witsenhausen’s construction corresponds to  $f = \text{Maj}_n$ , and so we get  $\Phi_p(f)$  approaches  $\frac{2}{\pi} \arcsin(\sqrt{p})$  as  $n \rightarrow \infty$ . Furthermore,  $\Phi_p(\text{Maj}_n)$  approaches this limit from below when  $p < 1/2$  (and approaches from above when  $p > 1/2$ ).<sup>7</sup>

Somewhat unrelated, Mossel (2010) extended the invariance principle of Mossel et al. (2010) and showed, as an application, that for any low-influence function  $f$ , it holds that  $\Phi_p(f) \leq \Phi_p(\text{Maj}_n) + o_n(1)$ .

Given all the results thus far, a reasonable refined version of the NICD-E conjecture that we believe captures the spirit of the original formulation is as follows—it is basically saying that for  $p < 1/2$ , Witsenhausen’s construction is optimal for NICD for SBES<sub>p</sub>.

**Conjecture 2** (Refined NICD-E). *For all  $p \in (0, 1/2)$ , odd  $n$ , and unbiased function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , it holds that  $\Phi_p(f) \leq \frac{2}{\pi} \arcsin(\sqrt{p})$ .*

*Remark.* Note however, that the formulation in terms of  $\Phi_p$  is not equivalent to the NICD setting because  $g(\mathbf{y}) := \text{sign}(\mathbb{E}[f(\mathbf{x})|\mathbf{y}])$  need not satisfy  $\mathbb{E}[g(\mathbf{y})] = 0$ . So a weaker variant, implied by the above conjecture, is that  $\mathbb{E}[f(\mathbf{x})g(\mathbf{y})] \leq \frac{2}{\pi} \arcsin(\sqrt{p})$  whenever  $\mathbb{E}[f(\mathbf{x})] = \mathbb{E}[g(\mathbf{y})] = 0$ .

## 2 Correlation Distillation Conjectures

Let  $[n]$  denote  $\{1, \dots, n\}$ . For  $\mathbf{x} = (x_1, \dots, x_n)$  and for  $S \subseteq [n]$ , let  $\mathbf{x}_S$  denote the subsequence of  $\mathbf{x}$  indexed by  $S$  and let  $\mathbf{1}_S$  denote the all-ones vector with coordinates indexed by  $S$ .

### 2.1 Majority is Least Stable (MLS)

**Conjecture 3** (MLS, Benjamini et al. (1999)). *For any odd integer  $n$ , any LTF  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  and  $\rho \in [0, 1]$ , it holds that  $\text{Stab}_\rho[f] \geq \text{Stab}_\rho[\text{Maj}_n]$ .*

Gopi (2013) proposed a counterexample to the conjecture for all odd  $n \geq 5$ . To define this counterexample, let  $h = (n - 1)/2$ , and partition  $[n]$  into  $A = [h + 1] = \{1, \dots, h + 1\}$  and  $B = [n] \setminus A = \{h + 2, \dots, n\}$ . Let  $e = \mathbf{1}_A \circ -\mathbf{1}_B$ . Define  $g_n : \{-1, 1\}^n \rightarrow \{-1, 1\}$  as

$$g_n(x) = \begin{cases} -\text{Maj}_n(x) & x \in \{e, -e\}, \\ \text{Maj}_n(x) & \text{otherwise.} \end{cases} \quad (2)$$

In other words,  $g_n$  is almost the same as  $\text{Maj}_n$  except at two antipodal points at the boundary that are negated. As noted in (Gopi, 2013),  $g_n$  is indeed an unbiased LTF, as it can be expressed as  $\text{sign}(a \cdot \sum_{i \in A} x_i + b \cdot \sum_{i \in B} x_i)$  for any odd positive integers  $a, b$  such that  $\frac{b}{a} \in \left(1 + \frac{1}{h}, 1 + \frac{1}{h-2}\right)$ . (For example,  $a = 2h - 1 = n - 2$  and  $b = 2h - 3 = n - 4$ .) It was shown that this is indeed a counterexample to the MLS conjecture, as stated more formally below.

<sup>7</sup>This is easy to show, and is similar in spirit to Condorcet’s Jury Theorem. We omit the proof.

**Theorem 4** (Gopi (2013)). *For every odd  $n \geq 5$ , there is  $\rho_n > 0$  such that  $\text{Stab}_\rho[g_n] < \text{Stab}_\rho[\text{Maj}_n]$  for all  $\rho \in (0, \rho_n)$ .*

Later, Biswas and Sarkar (2023) show that the MLS conjecture holds for all  $\rho \in (0, 1)$  when  $n \leq 3$ . Thus, the above counterexample is minimal. Additionally, they also independently discovered a different counterexample and proved a statement similar to Theorem 4. Nevertheless, an interesting question remains: Could the counterexample be extended to hold for all  $\rho \in (0, 1)$ ? Gopi (2013) partially answered this question, by showing that, for any fixed  $n$ , there exists a small neighborhood of  $\rho$  around 1 in which the MLS conjecture holds:

**Theorem 5** (Gopi (2013)). *For every odd  $n$ , there exists  $\zeta_n > 0$  such that, for all  $\rho \in [1 - \zeta_n, 1]$ , it holds that  $\text{Stab}_\rho[f] \geq \text{Stab}_\rho[\text{Maj}_n]$  for all LTF  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ .*

However, this does not answer, e.g., if for a fixed  $\rho = 0.99$ , does there exist  $n$  for which the MLS conjecture is false?

## 2.2 NICD for Erasures (NICD-E)

**Conjecture 6** (NICD-E, O’Donnell and Wright (2012)). *For any odd integer  $n$ , any unbiased  $f$  and  $p \in (0, 1/2)$ , it holds that  $\Phi_p[f] \leq \Phi_p[\text{Maj}_n]$ .*

Recently, Ivanišvili and Xie (2025) gave a (numerical) counterexample to this conjecture, for a specific value of  $n, p$ :

**Theorem 7** (Ivanišvili and Xie (2025)). *For  $n = 5$  and  $p = 0.4$ , there is an unbiased  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  such that  $\Phi_p[f] > \Phi_p[\text{Maj}_n]$ .*

And similar to Theorem 5, they showed that the NICD-E conjecture holds for  $p$  close to 0:

**Theorem 8** (Ivanišvili and Xie (2025)). *For any odd integer  $n$ , there exists  $\gamma_n > 0$  such that, for all unbiased  $f$  and  $p \in [0, \gamma_n]$ , it holds that  $\Phi_p[f] \leq \Phi_p[\text{Maj}_n]$ .*

## 3 Our Results

Our first result is to unify these two lines of work and show that Gopi’s counterexample violates the NICD-E conjecture as well for all  $n \geq 5$ , and furthermore provides a strengthening of Theorem 4 by showing that the counterexample refutes the MLS conjecture for almost all  $\rho$  (i.e.,  $\rho < 1 - \Omega(1/n^2)$ ):

**Theorem 9.** *Let  $\varepsilon_n := \frac{4}{(n-3)^2+6} = \frac{4+o(1)}{n^2}$ . There is a family  $\{g_n\}$  of unbiased LTFs such that the following holds for all odd integers  $n \geq 5$ :*

(i)  $\text{Stab}_\rho[g_n] < \text{Stab}_\rho[\text{Maj}_n]$  for all  $\rho \in (0, 1 - 2\varepsilon_n)$ .

(ii)  $\Phi_p[g_n] > \Phi_p[\text{Maj}_n]$  for all  $p \in (\varepsilon_n, \frac{1}{2})$ .

We complement this result and show almost matching bounds: (i) the MLS conjecture holds all unate functions (not just LTFs) for  $\rho$  close to 1 and (ii) the NICD-E conjecture holds for  $p$  close to 0, providing quantitative improvements over Theorems 5 and 8 where an explicit bound on the regime of parameters is not provided.

**Theorem 10.** *Let  $\gamma_n := \frac{4}{(n+7)(n-1)} = \frac{4-o(1)}{n^2}$ . For every odd  $n \geq 5$  and any unate unbiased  $f$ , we have the following:*

(i)  $\text{Stab}_\rho[\text{Maj}_n] \leq \text{Stab}_\rho[f]$  for all  $\rho \in (1 - 2\gamma_n, 1)$ .

(ii)  $\Phi_p[\text{Maj}_n] \geq \Phi_p[f]$  for all  $p \in (0, \gamma_n)$ .

Moreover, for every odd  $n \geq 5$  and any (possibly non-unate) unbiased  $f$ , we have:

(iii)  $\Phi_p[\text{Maj}_n] \geq \Phi_p[f]$  for all  $p \in (0, \gamma'_n)$  for  $\gamma'_n := \frac{2}{(n+2)(n-1)} \geq \frac{2}{n^2}$ .

In all cases, the inequality is in fact strict as long as  $f$  is not  $\text{Maj}_n$  (up to sign flips).

Finally, using similar techniques, we show that both conjectures hold unconditionally for  $n = 3$ .

**Theorem 11.** *The following statements hold:*

1.  $\text{Stab}_\rho[\text{Maj}_3] \leq \text{Stab}_\rho[f]$  holds for all unbiased LTF  $f$  and  $\rho \in (0, 1)$ , and

2.  $\Phi_p[\text{Maj}_3] \geq \Phi_p[f]$  holds for all unbiased  $f$  and  $p \in (0, 1/2)$ .

In both cases, the inequality is in fact strict as long as  $f$  is not  $\text{Maj}_3$  (up to sign flips).

While the case of MLS for  $n = 3$  was also shown by [Biswas and Sarkar \(2023\)](#), their proof involved an exhaustive enumeration over functions on 3 variables, and evaluating which ones are LTFs, whereas our proof is free of case analysis, and uses techniques developed to show [Theorem 10](#).

## 4 When the conjectures are false? (Proof of [Theorem 9](#))

### 4.1 Useful Properties of Binomials

We start by a couple of useful lemmata on binomial distributions. First is a lemma that relates  $\text{Bin}(h+1, p)$  and  $\text{Bin}(h, p)$ :

**Lemma 12.** For any  $p \in [0, 1]$  and  $h \in \mathbb{N}$ , we have

$$\mathbb{E}_{\substack{u \sim \text{Bin}(h+1, p) \\ u' \sim \text{Bin}(h, p)}} \left[ (-1)^{\mathbb{1}[u > u']} \right] = (1 - 2p) \Pr_{v, v' \sim \text{Bin}(h, p)} [v = v'] = (1 - 2p) \Pr_{w, w' \sim \text{Bin}(h, 1-p)} [w = w']$$

*Proof.* Write  $u = v + m$  where  $v \sim \text{Bin}(h, p)$  and  $m \sim \text{Ber}(p)$  are independent. Since  $u', v'$  are identically distributed, we have

$$\begin{aligned} \mathbb{E}_{\substack{u \sim \text{Bin}(h+1, p) \\ u' \sim \text{Bin}(h, p)}} \left[ (-1)^{\mathbb{1}[u > u']} \right] &= \mathbb{E}_{\substack{v, u' \sim \text{Bin}(h, p) \\ m \sim \text{Ber}(p)}} \left[ (-1)^{\mathbb{1}[v+m > u']} \right] \\ &= \mathbb{E}_{\substack{v, u' \sim \text{Bin}(h, p) \\ m \sim \text{Ber}(p)}} \left[ \frac{1}{2} \left( (-1)^{\mathbb{1}[v+m > u']} + (-1)^{\mathbb{1}[u'+m > v]} \right) \right]. \end{aligned}$$

Notice that the term  $(-1)^{\mathbb{1}[v+m > u']} + (-1)^{\mathbb{1}[u'+m > v]} = 0$  unless  $u' = v$ . Thus, the RHS equals

$$\Pr_{u', v'} [u' = v'] \cdot \mathbb{E}_{m \sim \text{Ber}(p)} \left[ (-1)^{\mathbb{1}[m > 0]} \right] = \Pr_{u', v'} [u' = v'] \cdot (1 - 2p),$$

which yields the first equality.

The second equality follows simply by substituting  $w = h - v, w' = h - v'$ . □

The second lemma gives a lower bound on the probability that two i.i.d. binomial random variables are equal. The form of the lower bound might look peculiar, but this is exactly what we will need subsequently.

**Lemma 13.** For any  $h \in \mathbb{N}$  such that  $h \geq 2$  and any  $q \in \left(\frac{1}{(h-1)^2+2}, \frac{1}{2}\right)$ , we have

$$(1-2q) \Pr_{w,w' \sim \text{Bin}(h,q)}[w = w'] + q^{2h+1} - (1-q)^{2h+1} > 0.$$

*Proof.* We have

$$\begin{aligned} \Pr_{w,w' \sim \text{Bin}(h,q)}[w = w'] &= \sum_{i=0}^h \Pr_w[w = i] \cdot \Pr_{w'}[w' = i] = \sum_{i=0}^h \left( \binom{h}{i} q^i (1-q)^{h-i} \right)^2 \\ &\geq (1-q)^{2h} + h^2 q^2 (1-q)^{2h-2} + q^4 (1-q)^{2h-4}, \end{aligned}$$

where the inequality follows by considering only the terms corresponding to  $i \in \{0, 1, 2\}$ , and using the rough lower bound  $\binom{h}{2} \geq 1$ .

Using  $q \leq 1/2$  and the above inequality, we can thus bound the desired quantity as follows:

$$\begin{aligned} &(1-2q) \Pr_{w,w' \sim \text{Bin}(h,q)}[w = w'] + q^{2h+1} - (1-q)^{2h+1} \\ &\geq (1-2q) \left( (1-q)^{2h} + h^2 q^2 (1-q)^{2h-2} + q^4 (1-q)^{2h-4} - \sum_{j=0}^{2h} q^j (1-q)^{2h-j} \right) \\ &\geq (1-2q) \left( (h^2 - 1)q^2 (1-q)^{2h-2} - q(1-q)^{2h-1} - (2h-3)q^3 (1-q)^{2h-3} \right) \\ &\geq (1-2q) \left( (h^2 - 2h + 2)q^2 (1-q)^{2h-2} - q(1-q)^{2h-1} \right) \\ &= (1-2q)q(1-q)^{2h-2} \left( (h^2 - 2h + 3)q - 1 \right) \\ &> 0, \end{aligned}$$

where the last inequality is due to our assumption that  $q \in \left(\frac{1}{(h-1)^2+2}, \frac{1}{2}\right)$ .  $\square$

## 4.2 Characterization of the Gaps

We now establish a simple formula of the differences  $\Phi_q[g_n] - \Phi_q[\text{Maj}_n]$  and  $\text{Stab}_{1-2q}[\text{Maj}_n] - \text{Stab}_{1-2q}[g_n]$  based on binomial distributions. In fact, the two quantities are equal, up to an appropriate scaling, as stated below.

**Theorem 14.** For all odd  $n \geq 3$  and  $q \in (0, 1)$ , it holds that

$$\begin{aligned} 2^{n-2} \cdot (\Phi_q[g_n] - \Phi_q[\text{Maj}_n]) &= 2^{n-3} \cdot (\text{Stab}_{1-2q}[\text{Maj}_n] - \text{Stab}_{1-2q}[g_n]) \\ &= (1-2q) \Pr_{z,z' \sim \text{Bin}\left(\frac{n-1}{2}, q\right)}[z = z'] + q^n - (1-q)^n. \end{aligned}$$

Note that [Theorem 9](#) immediately follows from [Theorem 14](#) and [Lemma 13](#) (using  $h = \frac{n-1}{2}$ ).

*Proof of Theorem 14.* We will separately derive the formula for  $\text{Stab}$  and  $\Phi$ . Let  $\delta_n : \{-1, 1\}^n \rightarrow \{-2, 0, 2\}$  be given as  $\delta_n(\mathbf{x}) = \text{Maj}_n(\mathbf{x}) - g_n(\mathbf{x})$ . From definition of  $g_n$  (2), it follows that  $\delta_n(e) = 2$ ,  $\delta_n(-e) = -2$  and  $\delta_n(\mathbf{x}) = 0$  for all other  $\mathbf{x}$ .

**Derivation for  $\text{Stab}_\rho$ .** Let  $\rho = 1 - 2q$ . For  $(\mathbf{x}, \mathbf{y}) \sim \text{DSBS}_\rho$ , we can write the difference between stabilities as

$$\begin{aligned} \text{Stab}_\rho[\text{Maj}_n] - \text{Stab}_\rho[g_n] &= \mathbb{E}[\text{Maj}_n(\mathbf{x})\text{Maj}_n(\mathbf{y}) - g_n(\mathbf{x})g_n(\mathbf{y})] \\ &= \mathbb{E}[\text{Maj}_n(\mathbf{x})\text{Maj}_n(\mathbf{y}) - (\text{Maj}_n(\mathbf{x}) - \delta_n(\mathbf{x}))(\text{Maj}_n(\mathbf{y}) - \delta_n(\mathbf{y}))] \\ &= 2\mathbb{E}[\delta_n(\mathbf{x})\text{Maj}_n(\mathbf{y})] - \mathbb{E}[\delta_n(\mathbf{x})\delta_n(\mathbf{y})]. \end{aligned}$$

Since  $\delta_n$  is only supported at  $e$  and  $-e$ , it is easy to see that

$$\mathbb{E}[\delta_n(\mathbf{x})\delta_n(\mathbf{y})] = \frac{1}{2^{n-3}} \left( \left( \frac{1+\rho}{2} \right)^n - \left( \frac{1-\rho}{2} \right)^n \right) = \frac{1}{2^{n-3}} ((1-q)^n - q^n).$$

To compute the first expectation, using the fact that both  $\delta_n$  and  $\text{Maj}_n$  are odd functions, we have

$$\mathbb{E}[\delta_n(\mathbf{x})\text{Maj}_n(\mathbf{y})] = \mathbb{E}[\delta_n(\mathbf{x})\text{Maj}_n(\mathbf{y}) \mid x_1 = 1] = \frac{1}{2^{n-2}} \mathbb{E}[\text{Maj}_n(\mathbf{y}) \mid \mathbf{x} = e],$$

where the second inequality is due to the fact that, conditioned on  $x_1 = 1$ ,  $\delta_n(\mathbf{x})$  is only non-zero when  $\mathbf{x} = e$ . Now, let  $n_A = |\{i \in A \mid y_i = 1\}|$  and  $n_B = |\{i \in B \mid y_i = -1\}|$ . We have that  $\text{Maj}_n(\mathbf{y}) = (-1)^{\mathbb{1}[n_A > n_B]}$ . Combining all these, we have

$$\text{Stab}_\rho[\text{Maj}_n] - \text{Stab}_\rho[g_n] = \frac{1}{2^{n-3}} \left( \mathbb{E} \left[ (-1)^{\mathbb{1}[n_A > n_B]} \right] + q^n - (1-q)^n \right)$$

Notice that, conditioned on  $\mathbf{x} = e$ , we have  $n_A \sim \text{Bin} \left( h+1, \frac{1+\rho}{2} \right) = \text{Bin}(h+1, 1-q)$  and  $n_B \sim \text{Bin} \left( h, \frac{1+\rho}{2} \right) = \text{Bin}(h, 1-q)$ . Thus, from [Lemma 12](#), we have

$$\text{Stab}_\rho[\text{Maj}_n] - \text{Stab}_\rho[g_n] = \frac{1}{2^{n-3}} \left( (1-2q) \Pr_{w, w' \sim \text{Bin}(h, q)} [w = w'] + q^n - (1-q)^n \right),$$

as desired.

**Derivation for  $\Phi_p$ .** We are interested in  $\Phi_p[g_n] - \Phi_p[\text{Maj}_n]$  for  $p = q$ . Let  $S \subseteq [n]$  be generated by including each element of  $[n]$  independently with probability  $p$ , and let  $\mathbf{x}_S \in \{-1, 1\}^S$  be uniformly at random. We also write  $\bar{S}$  to denote  $[n] \setminus S$ . Note that  $\Phi_p$  can be written as

$$\Phi_p[f] := \mathbb{E}_S \left[ \mathbb{E}_{\mathbf{x}_S} \left[ \left[ \mathbb{E}_{\mathbf{x}_{\bar{S}}} [f(\mathbf{x})] \right] \right] \right]$$

Thus, we can write the desired quantity as

$$\Phi_p[g_n] - \Phi_p[\text{Maj}_n] = \mathbb{E}_S \mathbb{E}_{\mathbf{x}_S} \left[ \frac{1}{2^{n-|S|}} \cdot h_S(\mathbf{x}_S) \right],$$

where

$$h_S(\mathbf{x}_S) := \left| \sum_{\mathbf{x}_{\bar{S}} \in \{-1, 1\}^{\bar{S}}} g_n(\mathbf{x}_S \circ \mathbf{x}_{\bar{S}}) \right| - \left| \sum_{\mathbf{x}_{\bar{S}} \in \{-1, 1\}^{\bar{S}}} \text{Maj}_n(\mathbf{x}_S \circ \mathbf{x}_{\bar{S}}) \right|.$$

For notational convenience, write  $e_S$  as the restriction of  $e$  on  $S$ . Now, if  $\mathbf{x}_S \notin \{e_S, -e_S\}$ , then we immediately have  $g_n(\mathbf{x}_S \circ \mathbf{x}_{\bar{S}}) = \text{Maj}_n(\mathbf{x}_S \circ \mathbf{x}_{\bar{S}})$ , and  $h_S(\mathbf{x}_S) = 0$ . Furthermore, notice that if  $S = \emptyset$ , we always have  $h_S(\mathbf{x}_S) = 0$ . Thus, the above term can be written as

$$\Phi_p[g_n] - \Phi_p[\text{Maj}_n] = \mathbb{E}_S \left[ \frac{1}{2^n} \cdot (h_S(e_S) + h_S(-e_S)) \right] = \frac{1}{2^{n-1}} \mathbb{E}_S [h_S(e_S)], \quad (3)$$

where in the second equality, we use the fact that  $g_n$  and  $\text{Maj}_n$  are odd and thus  $h_S$  is even.

Finally, to determine the value of  $h_S(e_S)$ , observe that, unless  $S = \emptyset$ , we have

$$\sum_{\mathbf{x}_{\bar{S}} \in \{-1,1\}^{\bar{S}}} g_n(e_S \circ \mathbf{x}_{\bar{S}}) = -2 + \sum_{\mathbf{x}_{\bar{S}} \in \{-1,1\}^{\bar{S}}} \text{Maj}_n(e_S \circ \mathbf{x}_{\bar{S}}).$$

This means that  $h_S(e_S)$  is completely determined by the sign of  $\sum_{\mathbf{x}_{\bar{S}} \in \{-1,1\}^{\bar{S}}} \text{Maj}_n(\mathbf{x}_S \circ \mathbf{x}_{\bar{S}})$ :

$$h_S(e_S) = \begin{cases} 0 & \text{if } S = \emptyset \text{ or } S = [n], \\ -2 & \text{if } \sum_{\mathbf{x}_{\bar{S}} \in \{-1,1\}^{\bar{S}}} \text{Maj}_n(e_S \circ \mathbf{x}_{\bar{S}}) > 0, \\ 2 & \text{otherwise.} \end{cases}$$

Let  $n_A := |S \cap A|$ ,  $n_B := |S \cap B|$ . Finally, notice that  $\text{Maj}_n(e_S \circ \mathbf{x}_{\bar{S}})$  is simply  $\text{sign}(n_A - n_B + \sum_{i \in \bar{S}} x_i)$ . Thus, we have

$$h_S(e_S) = \begin{cases} 0 & \text{if } S = \emptyset \text{ or } S = [n], \\ 2 \cdot (-1)^{\mathbb{1}[n_A > n_B]} & \text{otherwise.} \end{cases}$$

Plugging this back into (3), we have

$$\Phi_p[g_n] - \Phi_p[\text{Maj}_n] = \frac{1}{2^{n-2}} \left( \mathbb{E} \left[ (-1)^{\mathbb{1}[n_A > n_B]} \right] + p^n - (1-p)^n \right).$$

Notice that  $n_A \sim \text{Bin}(h+1, p)$ ,  $n_B \sim \text{Bin}(h, p)$  are independent. Thus, applying Lemma 12, we get

$$\Phi_p[g_n] - \Phi_p[\text{Maj}_n] = \frac{1}{2^{n-2}} \left( (1-2p) \Pr_{w, w' \sim \text{Bin}(h, p)} [w = w'] + p^n - (1-p)^n \right). \quad \square$$

## 5 When the conjectures are true? (Proof of Theorem 10)

For any function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , let  $E_f := \{\mathbf{x} \in \{-1, 1\}^n \mid f(\mathbf{x}) \neq \text{Maj}_n(\mathbf{x})\}$  and  $\mu_f := \Pr_{\mathbf{x} \sim \{-1, 1\}^n} [f(\mathbf{x}) \neq \text{Maj}_n(\mathbf{x})] = \frac{|E_f|}{2^n}$ .

### 5.1 Useful Lemmata on Gaps

We first obtain a lower bound on the difference between the first-level Fourier coefficients of Majority and  $f$  in terms of their disagreement. We state the lemma below without defining Fourier coefficients for simplicity.

**Lemma 15.** For any function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , it holds that

$$\sum_{i \in [n]} \mathbb{E}_{\mathbf{x} \sim \{-1, 1\}^n} [x_i \cdot (\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))] \geq 2 \cdot \mu_f.$$

*Proof.* We have

$$\sum_{i \in [n]} \mathbb{E}_{\mathbf{x} \sim \{-1, 1\}^n} [x_i \cdot (\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))] = \mathbb{E}_{\mathbf{x} \sim \{-1, 1\}^n} \left[ \left( \sum_{i \in [n]} x_i \right) \cdot (\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \right]$$

$$\geq 2 \cdot \Pr_{\mathbf{x} \sim \{-1,1\}^n} [\mathbf{x} \in E_f] = 2 \cdot \mu_f,$$

where in the inequality we use the fact that  $\left(\sum_{i \in [n]} x_i\right)$  and  $\text{Maj}_n(x)$  have the same sign, (and that the former is a non-zero integer), which means that  $\left(\sum_{i \in [n]} x_i\right) \cdot (\text{Maj}_n(x) - f(x)) \geq 2$  if  $f(x) \neq \text{Maj}_n(x)$ .  $\square$

We next prove a simple algebraic fact that we will use later.

**Lemma 16.** For  $n \geq 5$ ,

$$q < \frac{4}{(n+7)(n-1)} \implies \left(2(1-q)^{n-1} - q(1-q)^{n-2} \frac{(n-1)^2}{2} - q^2 \cdot 2 \binom{n}{3}\right) > 0.$$

*Proof.* From  $q < \frac{4}{(n+7)(n-1)}$ , we have

$$\begin{aligned} & 2(1-q)^{n-1} - q(1-q)^{n-2} \frac{(n-1)^2}{2} - q^2 \cdot 2 \binom{n}{3} \\ &= (1-q)^{n-2} \left(2(1-q) - q \frac{(n-1)^2}{2}\right) - q^2 \cdot 2 \binom{n}{3} \\ &= (1-q)^{n-2} \left(2 - q \frac{(n+1)^2}{2}\right) - q^2 \cdot 2 \binom{n}{3} \\ &\stackrel{(*)}{\geq} (1 - (n-2)q) \left(2 - q \frac{(n+1)^2}{2}\right) - q^2 \cdot 2 \binom{n}{3} \\ &= 2 - q \frac{(n+1)^2}{2} - 2q(n-2) + q^2(n-2) \frac{(n+1)^2}{2} - q^2 \cdot 2 \binom{n}{3} \\ &= 2 - q \cdot \frac{(n+7)(n-1)}{2} + q^2 \cdot (n-2) \cdot \left(\frac{(n+1)^2}{2} - \frac{n(n-1)}{3}\right) \\ &\stackrel{(**)}{>} 2 - q \cdot \frac{(n+7)(n-1)}{2} > 0, \end{aligned}$$

where in (\*), we applied Bernoulli's inequality,  $(1-q)^{n-2} \geq 1 - (n-2)q$ , to the first term and in (\*\*), we used that the numerator of the dropped term is positive for  $n \geq 5$ .  $\square$

For any  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  we will consider the ‘‘gap’’ between  $f$  and  $\text{Maj}_n$  that has a very specific form. In particular, the gap will have the following binomial-expansion-like form:

$$G_q[f] = \sum_{k=1}^n \left[ q^k (1-q)^{n-k} \cdot \sum_{S \in \binom{[n]}{k}} G_S(f) \right], \quad (4)$$

for some  $G_S(f) \in \mathbb{R}$ . We prove a generic bound on  $G_q[f]$  in terms of the subset expressions  $G_S(\cdot)$ .

**Lemma 17.** Suppose there is a constant  $c > 0$  such that:

- (i) For  $i \in [n]$ ,  $G_{\{i\}}(f) = c \cdot \mathbb{E}_{\mathbf{x}}[x_i \cdot (\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))]$ .
- (ii) For  $i \neq j \in [n]$ ,  $G_{\{i,j\}}(f) \geq c \cdot \left(\frac{1}{4} \mathbb{E}_{\mathbf{x}}[(x_i + x_j)(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \mid x_i = x_j] - \Pr[\mathbf{x} \in E_f \mid x_i \neq x_j]\right)$ .
- (iii) For  $S$  such that  $|S| \geq 3$ ,  $G_S(f) \geq -2c \cdot \Pr[\mathbf{x} \in E_f]$ .

Then,  $G_q[f] > 0$  for  $q < \frac{4}{(n+7)(n-1)}$ .

*Proof.* We lower bound the LHS using our assumptions on each term on the RHS. For the  $k = 1$  term, we sum over all singletons and apply [Lemma 15](#) to obtain

$$\sum_{i \in [n]} G_{\{i\}}(f) = c \cdot \sum_{i \in [n]} \mathbb{E}_{\mathbf{x}}[x_i \cdot (\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))] \geq 2c \cdot \mu_f.$$

Thus, the contribution to  $G_q[f]$  from the  $k = 1$  term is at least  $2c \cdot \mu_f \cdot q(1 - q)^{n-1}$ .

For the term corresponding to  $k = 2$ , we sum over all  $\binom{n}{2}$  pairs. Note that the bound in our assumption has two components. Considering the first component, we obtain

$$\begin{aligned} c \cdot \sum_{i < j} \frac{1}{4} \mathbb{E}_{\mathbf{x}}[(x_i + x_j)(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \mid x_i = x_j] &= \frac{c}{2} \cdot \sum_{i < j} \mathbb{E}_{\mathbf{x}}[(x_i + x_j)(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))] \\ &= c \cdot \frac{n-1}{2} \sum_{i \in [n]} \mathbb{E}_{\mathbf{x}}[x_i(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))] \\ &\geq c(n-1)\mu_f, \end{aligned}$$

where the first equality follows by observing that if  $x_i \neq x_j$  then  $x_i + x_j = 0$ , so we can drop the conditioning on  $x_i = x_j$  by gaining a factor of 2, and the last step follows from [Lemma 15](#). Now we consider the second component.

$$\begin{aligned} -c \cdot \sum_{i < j} \Pr[\mathbf{x} \in E_f \mid x_i \neq x_j] &= -2c \cdot \sum_{i < j} \Pr[\mathbf{x} \in E_f \text{ and } x_i \neq x_j] \\ &= -2c \binom{n}{2} \cdot \Pr_{\mathbf{x}, S=\{i,j\}}[\mathbf{x} \in E_f \text{ and } x_i \neq x_j] \\ &= -2c \binom{n}{2} \cdot \Pr_{\mathbf{x}, S=\{i,j\}}[x_i \neq x_j \mid \mathbf{x} \in E_f] \cdot \mu_f \\ &\geq -\frac{c}{2}(n^2 - 1) \cdot \mu_f, \end{aligned}$$

where the last inequality follows from the fact that, for any  $\mathbf{x} \in \{-1, 1\}^n$ , the probability (over random  $\{i, j\} \in \binom{[n]}{2}$ ) that  $x_i \neq x_j$  is at most  $\frac{\binom{n+1}{2}\binom{n-1}{2}}{\binom{n}{2}} = \frac{(n^2-1)}{4\binom{n}{2}}$ .

Thus, the contribution to  $G_q[f]$  from the  $k = 2$  term is at least

$$q^2(1 - q)^{n-2} \left( c(n-1)\mu_f - c\frac{n^2-1}{2}\mu_f \right) = -c \cdot q^2(1 - q)^{n-2} \frac{(n-1)^2}{2} \mu_f.$$

For the terms corresponding to  $k \geq 3$ , we lower bound the contribution to  $G_q[f]$  as follows:

$$\begin{aligned} \sum_{k=3}^n q^k(1 - q)^{n-k} \sum_{S \in \binom{[n]}{k}} G_S(f) &\geq -2c \sum_{k=3}^n q^k(1 - q)^{n-k} \sum_{S \in \binom{[n]}{k}} \Pr[x \in E_f] \\ &\geq -2c \cdot \mu_f \sum_{k=3}^n q^k(1 - q)^{n-k} \binom{n}{k} \\ &\geq -2c \cdot \mu_f q^3 \binom{n}{3}, \end{aligned}$$

where the last inequality is due to the union bound:  $\sum_{k=3}^n q^k(1-q)^{n-k} \binom{n}{k}$  is the probability that there are at least 3 successes from  $n$  i.i.d. Bernoulli random variables with success probability  $q$ .

Summing the above contributions and applying [Lemma 16](#) with the choice of  $q$ , we obtain

$$G_q[f] \geq c \cdot \mu_f \cdot q \left( 2(1-q)^{n-1} - q(1-q)^{n-2} \frac{(n-1)^2}{2} - q^2 \cdot 2 \binom{n}{3} \right) > 0. \quad \square$$

## 5.2 Bounds on the Gaps

**Derivation for  $\text{Stab}_\rho$ .** We assume here that  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  is monotone (but not necessarily unbiased). Let  $\rho = 1 - 2q$ . Recall that

$$\text{Stab}_\rho[f] - \text{Stab}_\rho[\text{Maj}_n] = \sum_{k=1}^n q^k(1-q)^{n-k} \sum_{S \in \binom{[n]}{k}} G_S(f),$$

where

$$\begin{aligned} G_S(f) &= \mathbb{E}_{\mathbf{x}} [f(\mathbf{x})f(\mathbf{x}^{\oplus S}) - \text{Maj}_n(\mathbf{x})\text{Maj}_n(\mathbf{x}^{\oplus S})] \\ &= \mathbb{E}_{\mathbf{x}} [|\text{Maj}_n(\mathbf{x}) - \text{Maj}_n(\mathbf{x}^{\oplus S})| - |f(\mathbf{x}) - f(\mathbf{x}^{\oplus S})|]. \end{aligned}$$

Here, we are using that sampling  $(\mathbf{x}, \mathbf{y}) \sim \text{DSBS}_\rho^{\otimes n}$  is equivalent to sampling  $\mathbf{x}$  uniformly, a subset  $S$  by including each element with probability  $\frac{1-\rho}{2} = q$  and setting  $\mathbf{y} = \mathbf{x}^{\oplus S}$  (flipping all coordinates in  $S$ ). We wish to apply [Lemma 17](#) to this expansion. We have the following bounds.

- For  $|S| = 1$ , let  $S = \{i\}$ . We have  $G_{\{i\}}(f) = \mathbb{E}_{\mathbf{x}} [2x_i(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))]$ , since for any monotone function  $h : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , it holds that  $\mathbb{E}_{\mathbf{x}} [h(\mathbf{x}) - h(\mathbf{x}^{\oplus i})] = \mathbb{E}_{\mathbf{x}} [2x_i h(\mathbf{x})]$ . Thus  $G_{\{i\}}[f]$  satisfies [Lemma 17](#)(i) with  $c = 2$ .
- For  $|S| = 2$ , let  $S = \{i, j\}$ ; we have two cases. If  $x_i = x_j$ , then

$$\begin{aligned} & \frac{1}{2} \cdot \mathbb{E}_{\mathbf{x}} [|\text{Maj}_n(\mathbf{x}) - \text{Maj}_n(\mathbf{x}^{\oplus S})| - |f(\mathbf{x}) - f(\mathbf{x}^{\oplus S})| \mid x_i = x_j] \\ &= \frac{1}{4} \cdot \mathbb{E}_{\mathbf{x}} [(x_i + x_j)(\text{Maj}_n(\mathbf{x}) - \text{Maj}_n(\mathbf{x}^{\oplus S})) - (x_i + x_j)(f(\mathbf{x}) - f(\mathbf{x}^{\oplus S})) \mid x_i = x_j] \\ &= \frac{1}{2} \mathbb{E}_{\mathbf{x}} [(x_i + x_j) \cdot (\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \mid x_i = x_j], \end{aligned}$$

where the first step uses the monotonicity of  $\text{Maj}_n$  and  $f$ . If  $x_i \neq x_j$ , notice that  $\text{Maj}_n(\mathbf{x}) = \text{Maj}_n(\mathbf{x}^{\oplus S})$ . Thus, we have

$$\begin{aligned} & \frac{1}{2} \cdot \mathbb{E}_{\mathbf{x}} [|\text{Maj}_n(\mathbf{x}) - \text{Maj}_n(\mathbf{x}^{\oplus S})| - |f(\mathbf{x}) - f(\mathbf{x}^{\oplus S})| \mid x_i \neq x_j] \\ & \geq -\Pr_{\mathbf{x}} [\mathbf{x} \in E_f \text{ or } \mathbf{x}^{\oplus S} \in E_f \mid x_i \neq x_j] \\ & = -2 \Pr_{\mathbf{x}} [\mathbf{x} \in E_f \mid x_i \neq x_j]. \end{aligned}$$

Putting these bounds together,

$$G_{\{i,j\}}(f) \geq \frac{1}{2} \mathbb{E}_{\mathbf{x}} [(x_i + x_j) \cdot (\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \mid x_i = x_j] - 2 \Pr_{\mathbf{x}} [\mathbf{x} \in E_f \mid x_i \neq x_j].$$

This satisfies [Lemma 17](#)(ii) also with  $c = 2$ .

- For  $|S| \geq 3$ , we have

$$\begin{aligned} G_S(f) &= \mathbb{E}_{\mathbf{x}} [|\text{Maj}_n(\mathbf{x}) - \text{Maj}_n(\mathbf{x}^{\oplus S})| - |f(\mathbf{x}) - f(\mathbf{x}^{\oplus S})|] \\ &\geq -2 \cdot \Pr_{\mathbf{x}} [\mathbf{x} \in E_f \text{ or } \mathbf{x}^{\oplus S} \in E_f] \geq -4 \cdot \Pr[\mathbf{x} \in E_f], \end{aligned}$$

This satisfies [Lemma 17](#)(iii) also with  $c = 2$ .

From [Lemma 17](#), we conclude  $\text{Stab}_\rho[f] - \text{Stab}_\rho[\text{Maj}_n] > 0$  as desired.

**Derivation for  $\Phi_p$ .** We assume here that  $f$  is unbiased. Let  $g^*$  be the optimal function in the definition of  $\Phi_p[f]$  defined as<sup>8</sup>  $g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}}) := \text{sign}\left(\mathbb{E}_{\mathbf{x}_{\bar{S}} \in \{-1,1\}^{\bar{S}}}[f(\mathbf{x}_S \circ \mathbf{x}_{\bar{S}})]\right)$ . Note that  $\Phi_p[\text{Maj}_n] \geq \mathbb{E}_{\mathbf{x}, S}[\text{Maj}_n(\mathbf{x}) \cdot g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}})]$ . Therefore,

$$\Phi_p[\text{Maj}_n] - \Phi_p[f] \geq \mathbb{E}_{\mathbf{x}, S} [(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \cdot g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}})] = \sum_{k=1}^n q^k (1-q)^{n-k} \sum_{S \in \binom{[n]}{k}} G_S(f),$$

where  $q = p$  and

$$G_S(f) = \mathbb{E}_{\mathbf{x}} [(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \cdot g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}})],$$

and we note that the  $k = 0$  term disappears since  $f, \text{Maj}_n$  are both unbiased.

As noted in ([Ivanisvili and Xie, 2025](#)), we may assume<sup>9</sup> w.l.o.g. that the first-level Fourier coefficient is non-negative, i.e.,  $\mathbb{E}_{\mathbf{x}}[x_i \cdot f(\mathbf{x})] \geq 0$ . Observe that this implies  $g^*(x_i \circ \mathbf{0}_{[n] \setminus \{i\}}) = x_i$ .

**Lemma 18.** For any event  $C$ ,  $\mathbb{E}_{\mathbf{x}}[(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \cdot g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}}) \mid C] \geq -2 \Pr[\mathbf{x} \in E_f \mid C]$ .

*Proof.* Note that since  $|g^*| \leq 1$ , we have that  $(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \cdot g^*(\cdot)$  evaluates to 0 if  $x \notin E_f$ , and is lower bounded by  $-2$  if  $x \in E_f$ . The proof follows.  $\square$

We wish to apply [Lemma 17](#) to this expansion. First, consider the case of  $|S| = 1$ , and let  $S = \{i\}$ . Since  $g^*(x_i \circ \mathbf{0}_{[n] \setminus \{i\}}) = x_i$ , we have  $G_{\{i\}}(f) = \mathbb{E}_{\mathbf{x}}[x_i(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x}))]$ ; this satisfies [Lemma 17](#)(i) with  $c = 1$ .

For  $|S| > 1$ , we will separate into two cases based on whether  $f$  is monotone.

*When  $f$  is monotone.* In this case, we have that

- For  $|S| = 2$ , let  $S = \{i, j\}$ . We have two cases. If  $x_i = x_j$ , for an unbiased monotone function, notice that we simply have  $g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}}) = \text{sign}(x_i + x_j) = \frac{x_i + x_j}{2}$ . Thus,

$$\begin{aligned} &\frac{1}{2} \mathbb{E}_{\mathbf{x}} [(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \cdot g^*(\mathbf{x}_{\{i,j\}} \circ \mathbf{0}_{[n] \setminus \{i,j\}}) \mid x_i = x_j] \\ &= \frac{1}{4} \mathbb{E}_{\mathbf{x}} [(x_i + x_j)(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \mid x_i = x_j]. \end{aligned}$$

If  $x_i \neq x_j$ , applying [Lemma 18](#), we have

$$\frac{1}{2} \mathbb{E}_{\mathbf{x}} [(\text{Maj}_n(\mathbf{x}) - f(\mathbf{x})) \cdot g^*(\mathbf{x}_{\{i,j\}} \circ \mathbf{0}_{[n] \setminus \{i,j\}}) \mid x_i \neq x_j] \geq -\Pr[\mathbf{x} \in E_f \mid x_i \neq x_j].$$

Putting these bounds together, we have [Lemma 17](#)(ii) with  $c = 1$ .

<sup>8</sup>If the inner expectation  $\mathbb{E}_{\mathbf{x}_{\bar{S}} \in \{-1,1\}^{\bar{S}}}[f(\mathbf{x}_S \circ \mathbf{x}_{\bar{S}})]$  is zero, then we may select  $g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}})$  arbitrarily.

<sup>9</sup>This follows directly from monotonicity. Even without monotonicity, this simply follows by “flipping coordinate  $i$ ” that violates this condition, which does not change  $\Phi_p[f]$ ; see the proof of [Lemma 19](#) for a formalization.

- For  $|S| \geq 3$ , we have by [Lemma 18](#) that  $G_S(f) \geq -2\Pr[\mathbf{x} \in E_f]$ . Thus, we have [Lemma 17\(iii\)](#) with  $c = 1$ .

From [Lemma 17](#), we conclude  $\Phi_p[\text{Maj}_n] - \Phi_p[f] > 0$  as desired.

When  $f$  is not monotone. Finally, consider the case where  $f$  is non-monotone. In this case, we simply apply [Lemma 18](#) to obtain a tail bound on all terms, including the second term,

$$\sum_{k=2}^n q^k(1-q)^{n-k} \sum_{S \in \binom{[n]}{k}} G_S(f) \geq \sum_{k=2}^n q^k(1-q)^{n-k} \sum_{S \in \binom{[n]}{k}} (-2\mu_f) \geq -2\mu_f \cdot q^2 \binom{n}{2} = -\mu_f \cdot q^2 n(n-1).$$

Combining with the first term, we get

$$\begin{aligned} \Phi_p[\text{Maj}_n] - \Phi_p[f] &\geq q(1-q)^{n-1} \cdot 2\mu_f - \mu_f \cdot q^2 n(n-1) = \mu_f q(2(1-q)^{n-1} - qn(n-1)) \\ &\geq \mu_f q(2(1-(n-1)q) - qn(n-1)) > 0, \end{aligned}$$

where we used Bernoulli's inequality in the penultimate step and our assumption  $0 < q < \frac{2}{(n+2)(n-1)}$  in the last step.

## 6 The Case of $n = 3$ (Proof of [Theorem 11](#))

While the previous sections established bounds for general  $n$ , the case of  $n = 3$  is small enough that we can evaluate the veracity of both conjectures. In fact, we use techniques from [Section 5](#).

**Derivation for  $\text{Stab}_\rho$ .** Assume that  $f$  is monotone and odd. We can use the same expression for  $G_S(f)$  for the case  $k = 1, 2$  as in [Section 5](#). Finally, notice that for  $k = 3$  (i.e.,  $S = \{1, 2, 3\}$ ), we simply have  $G_S(f) = 0$  since  $f$  is odd. As a result, following the inequalities in [Lemma 17](#) but without the  $k \geq 3$  term, we get

$$G_q[f] \geq c \cdot \mu_f \cdot q \left( 2(1-q)^{n-1} - q(1-q)^{n-2} \frac{(n-1)^2}{2} \right) = c \cdot \mu_f \cdot q(1-q)(2-4q) > 0,$$

where in the first equality we use  $n = 3$  and in the last inequality we use  $0 < q < 1/2$ .

**Derivation for  $\Phi_\rho$ .** In this case, we will use the following lemma.

**Lemma 19.** For any unbiased  $f : \{-1, 1\}^3 \rightarrow \{-1, 1\}$ , there exists another unbiased function  $f' : \{-1, 1\}^3 \rightarrow \{-1, 1\}$  such that  $\Phi_\rho[f'] = \Phi_\rho[f]$  and there is an optimum  $g^*$  in the definition of  $\Phi_\rho[f]$  that satisfies:

- (i)  $g^*(x_i \circ \mathbf{0}_{[n] \setminus \{i\}}) = x_i$  for all  $i \in [3]$  and  $\mathbf{x} \in \{-1, 1\}^3$  and
- (ii)  $g^*(\mathbf{x}_S \circ \mathbf{0}_{\bar{S}}) = (x_i + x_j)/2$  for all  $i \neq j$  and  $\mathbf{x} \in \{-1, 1\}^3$  such that  $x_i = x_j$ .

Before we prove [Lemma 19](#), we show how to use this to finish the proof. From the lemma, we may assume w.l.o.g. that  $f$  itself has  $g^*$  that satisfies the two properties. Note that these are the only two properties used in derivation of the monotone case of  $\Phi_\rho[f]$  in [Section 5](#). Again, similar to the derivation for  $\text{Stab}$  above, by recognizing that the  $k = 3$  term cancels out and following the inequalities in [Lemma 17](#) but without the  $k \geq 3$  term, we get

$$G_q[f] \geq c \cdot \mu_f \cdot q \left( 2(1-q)^{n-1} - q(1-q)^{n-2} \frac{(n-1)^2}{2} \right) \geq 0.$$

We will now prove [Lemma 19](#).

*Proof of Lemma 19.* We pick  $f'$  of the form  $f' = f(\mathbf{x} \oplus a)$  for some  $a \in \{-1, 1\}^3$  that maximizes  $\mathbb{E}_{\mathbf{x}}[x_i \cdot f'(\mathbf{x})]$ , ties broken arbitrarily. Note that this ensures that  $\Phi_\rho[f'] = \Phi_\rho[f]$ . Furthermore, our choice of  $f'$  ensures that  $\mathbb{E}_{\mathbf{x}}[x_i \cdot f'(\mathbf{x})] \geq 0$  for all  $i \in [n]$ . That is, property (i) holds.

Now, suppose that property (ii) does not hold. This means that there exists  $\mathbf{x} \in \{-1, 1\}^3$  and  $i \neq j$  such that  $x_i = x_j$  with  $\mathbb{E}_{\mathbf{x}_{\overline{S}} \in \{-1, 1\}^{\overline{S}}} [f'(\mathbf{x}_S \circ \mathbf{x}_{\overline{S}})] \cdot (x_i + x_j) < 0$ . By symmetry, we may assume that  $i = 1, j = 2$ , and  $x_i = x_j = 1$ . This implies that  $f'(1, 1, 1) = f'(1, 1, -1) = -1$ .

However, since  $\mathbb{E}_{\mathbf{x}}[x_i \cdot f'(\mathbf{x})] \geq 0$  and  $f'$  is unbiased, plugging in  $i = 1$  and  $i = 2$ , respectively, implies that

$$f'(1, -1, -1) = g(1, -1, 1) = 1 \text{ and } f'(-1, 1, 1) = f'(-1, 1, -1) = 1,$$

respectively. Finally, since  $f'$  is unbiased, we also have  $f'(-1, -1, 1) = f'(-1, -1, -1) = -1$ . This means that  $f'(\mathbf{x}) = -x_1x_2$ . However, in this case, we may take  $f'(\mathbf{x}) = x_1x_2$  instead, which satisfies both properties (i) and (ii).  $\square$

## References

- Itai Benjamini, Gil Kalai, and Oded Schramm. Noise sensitivity of Boolean functions and applications to percolation. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 90(1):5–43, 1999.
- Aniruddha Biswas and Palash Sarkar. On the “Majority is Least Stable” conjecture. *IPL*, 179, 2023.
- Anindya De, Elchanan Mossel, and Joe Neeman. Noise stability is computable and approximately low-dimensional. In *CCC*, pages 10:1–10:11, 2017.
- Anindya De, Elchanan Mossel, and Joe Neeman. Non interactive simulation of correlated distributions is decidable. In *SODA*, pages 2728–2746, 2018.
- Yuval Filmus, Hamed Hatami, Steven M. Heilman, Elchanan Mossel, Sushant Sachdeva, Andrew Wan, and Karl Wimmer. Real analysis in computer science: A collection of open problems, 2014. URL <https://simons.berkeley.edu/sites/default/files/openprobsmerged.pdf>.
- Hans Gebelein. Das statistische problem der korrelation als variations- und eigenwertproblem und sein zusammenhang mit der ausgleichsrechnung. *J. Appl. Math. Mech. / Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM)*, 21(6):364–379, 1941.
- Badih Ghazi, Pritish Kamath, and Madhu Sudan. Decidability of non-interactive simulation of joint distributions. In *FOCS*, pages 545–554, 2016.
- Badih Ghazi, Pritish Kamath, and Prasad Raghavendra. Dimension reduction for polynomials over Gaussian space and applications. In *CCC*, pages 28:1–28:37, 2018.
- Sivakanth Gopi. Stability of linear threshold functions. Undergraduate thesis, IIT Bombay, 2013. URL [https://gopisivakanth.github.io/files/UndergraduateThesis\\_IITBombay.pdf](https://gopisivakanth.github.io/files/UndergraduateThesis_IITBombay.pdf).
- H. O. Hirschfeld. A connection between correlation and contingency. *Math. Proc. Cambridge Phil. Soc.*, 31(4):520–524, 1935.
- Paata Ivanisvili and Xinyuan Xie. Counterexample to majority optimality in NICD with erasures. *arXiv*, 2510.20013, 2025.

- Vishesh Jain. A counterexample to the “majority is least stable” conjecture. *arXiv*, 1703.07657, 2017.
- Elchanan Mossel. Gaussian bounds for noise correlation of functions. *GAF*, 19:1713–1756, 2010.
- Elchanan Mossel, Ryan O’Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. Math.*, 171(1):295–341, 2010.
- Ryan O’Donnell. *Computational Applications Of Noise Sensitivity*. PhD thesis, MIT, 2003.
- Ryan O’Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 1st edition, 2014.
- Ryan O’Donnell and John Wright. A new point of NP-hardness for unique games. In *STOC*, pages 289–306, 2012.
- Alfréd Rényi. On measures of dependence. *Acta Mathematica Academiae Scientiarum Hungarica*, 10:441–451, 1959.
- H. S. Witsenhausen. On sequences of pairs of dependent random variables. *SIAM J. Appl. Math.*, 28(1):100–113, 1975.
- Ke Yang. On the (im)possibility of non-interactive correlation distillation. In *LATIN*, pages 222–231, 2004.