Coin flipping from a cosmic source: On error correction of truly random bits

Elchanan Mossel
Microsoft Research
mossel@microsoft.com

Ryan O'Donnell*
MIT Department of Mathematics
odonnell@theory.lcs.mit.edu

November 7, 2002

Abstract

We study a new problem related to coin flipping, coding theory, and noise sensitivity. Consider a source of truly random bits $x \in \{0,1\}^n$, and k parties, who have noisy versions of the source bits $y^i \in \{0,1\}^n$, where for all i and j, it holds that $\mathbf{P}[y^i_j = x_j] = 1 - \epsilon$, independently for all i and j. That is, each party sees each bit correctly with probability $1-\epsilon$, and incorrectly (flipped) with probability ϵ, independently for all bits and all parties. The parties, who cannot communicate, wish to agree beforehand on *balanced* functions $f_i : \{0,1\}^n \to \{0,1\}$ such that $\mathbf{P}[f_1(y^1) = \cdots = f_k(y^k)]$ is maximized. In other words, each party wants to toss a fair coin so that the probability that all parties have the same coin is maximized. The functions f_i may be thought of as an error correcting procedure for the source x.

When $k = 2, 3$ no error correction is possible, as the optimal protocol is given by $f_i(x^i) = y^i_1$. On the other hand, for large values of k, better protocols exist. We study general properties of the optimal protocols and the asymptotic behavior of the problem with respect to k, n and ϵ. Our analysis uses tools from probability, discrete Fourier analysis, convexity and discrete symmetrization.

*Supported by NSF grant CCR-99-12342.

1 Introduction

Consider a source of truly random bits $x \in \{0,1\}^n$, which is accessible to k parties. If the k parties want to use the source in order to obtain a common single random bit, they can easily do so by deciding beforehand to let the common bit be x_1. More generally, they can decide beforehand on any balanced function $f : \{0,1\}^n \to \{0,1\}$, and let the common bit be $f(x)$. We call a function f *balanced* if $\mathbf{P}_x[f(x) = 0] = \mathbf{P}_x[f(x) = 1] = 1/2$.

In this setting, there is no real advantage in taking the function f to be anything other than $f(x) = x_1$. The problem becomes more interesting when the parties receive *noisy* versions of the random bits. That is, party i receives y^i, where the bits of y^i satisfy $\mathbf{P}[y^i_j = x_j] = 1 - \epsilon$, independently for all i and j. We also assume that the parties cannot communicate. Yet, the parties want to toss the same fair coin given their noisy versions of the source. We will now allow each party i to use a different *balanced* function $f_i : \{0,1\}^n \to \{0,1\}$ as a coin-tossing procedure. We want to maximize $\mathbf{P}[f_1(y^1) = \cdots = f_k(y^k)]$.

This problem is motivated naturally by several models in cryptography. Think of a long one-time pad which is distributed to parties with a small probability of error. The parties still want to use this one-time pad as their key to an encryption algorithm, by dividing the one-time pad into blocks of length n, and applying some function on each block to obtain a shorter one-time pad which has high probability of being the same for all parties. One setting in which such a procedure should be useful is Ding and Rabin's "everlasting security" [DR01], a strong encryption algorithm in the bounded storage model. This model presupposes the existence of a satellite broadcasting a continuous stream of a huge number of random bits. It is natural to expect some error in any reception of this stream. A somewhat related cryptographic problem was studied by Maurer [M97];

Our problem is also of interest as a noncryptographic collective coin flipping problem. One example of such a problem is the full information model, introduced by Ben-Or and Linial [BL90] and studied extensively (see, e.g., the survey [D00]). In this problem, many parties try to agree on a single random bit; each generates a random coin toss, and there is a single protocol (function) taking all the coin tosses and producing a bit. The difficulty arises from the assumption that some parties are corrupt and can choose their coins adversarially. In our problem, the major difference is that the parties do not communicate any random bits, so they each must apply a protocol to a shared random string. And, instead of arbitrary corruptions, we assume random ones.

The question presented in this paper is also a natural question regarding error correcting for the broadcast channel (see e.g. [CT91]) with a truly random source. Naturally, when the source is truly random, error correction is impossible. However, here instead of requiring that all parties receive the information transmitted to them with high probability, we require that all parties attain the same information with high probability, and that this mutual information has high entropy.

Finally, a basic motivation comes from the study of noise-sensitivity, see [KKL88] and [BKS99]. The functions f_i that maximize the probability $\mathbf{P}[f_1(y^1) = \cdots = f_k(y^k)]$ are in an intuitive sense stable to noise, and it turns out that when the number of parties k is 2 or 3, this intuition can be used in order to prove that the optimal functions are just the first-bit function.

1.1 Definitions and notation

We begin by defining the problem.

Definition 1.1

- **The model** *Let $k \geq 1$ be the number of* parties, *and $n \geq 1$ be the* block length. *Let $\epsilon \in (0, 1/2)$ be the* corruption probability. *Our space is the space of all sequences $(x, y^1, \ldots, y^k) \in \{0,1\}^{n\times(k+1)}$, where x represents the* source *and is chosen uniformly at random from $\{0,1\}^n$. For each i, y^i represents the bits that party i holds and it is assumed that for all $1 \leq i \leq k$ and $1 \leq j \leq n$, it holds that $\mathbf{P}[y^i_j = x_j] = 1 - \epsilon$, independently for all i and j. This is our probability space, and when we write $\mathbf{P}$ (or $\mathbf{E}$) we mean the probability (expected value) in this space.*

- **Balanced and antisymmetric functions** *Let $\mathcal{B}_n$ denote the set of* balanced *functions $f : \{0,1\}^n \to \{0,1\}$; i.e., those with $|f^{-1}(0)| = |f^{-1}(1)|$. Let $\mathcal{A}_n$ denote the set of* antisymmetric *n-bit boolean functions; i.e., those satisfying $f(\bar{x}_1\bar{x}_2\cdots\bar{x}_n) = \overline{f(x)}$, where the bar denotes flipping 0's and 1's, so $\bar{x} = 1 - x$. Note that $\mathcal{A}_n \subset \mathcal{B}_n$.*

- **Protocols** *A* protocol *consists of k functions $f_i \in \mathcal{B}_n$. An* antisymmetric protocol *consists of k functions $f_i \in \mathcal{A}_n$. For a protocol $(f_1, \ldots, f_k)$, we write $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$ for the probability that all functions agree, so*

$$\mathcal{P}(f_1, \ldots, f_k; \epsilon) = \mathbf{P}[f_1(y^1) = \cdots = f_k(y^k)].$$

We write $\mathcal{P}_k(f; \epsilon)$, in place of $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$, if $f_1 = \cdots = f_k = f$.

It turns out that restricting all f_i to be balanced is neither necessary nor sufficient for ensuring that the output bit, when agreed upon, is uniformly random — see Proposition G.3 for a counterexample to sufficiency. A sufficient condition is that every function be antisymmetric, since if all the function are antisymmetric, then

$$\mathbf{P}[f_1(y^1) = \cdots = f_k(y^k) = 1] = \mathbf{P}[f_1(\overline{y^1}) = f_k(\overline{y^k}) = 0] = \mathbf{P}[f(y^1) = \cdots = f(y^k) = 0],$$

where the first equality follows from the fact that f_i are antisymmetric and the second since $\mathbf{P}$ assigns the same probability to $(x, y^1, \ldots, y^k)$ as it does to $(\bar{x}, \overline{y^1}, \ldots, \overline{y^k})$. We are not aware of a weaker condition than antisymmetry that ensures that the output bit when agreed upon is uniformly random.

We end this section with a few more definitions. For $S \subseteq [n]$ and π a permutation of $[n]$, let $\pi_S : \{0,1\}^n \to \{0,1\}^n$ be defined by $\pi_S(x)_i = x_{\pi(i)}$ if $i \in S$, and $\pi_S(x)_i = \overline{x_{\pi(i)}}$ if $i \notin S$. Any π_S merely permutes coordinates, and flips the roles and 0 and 1 on some coordinates. It's therefore easy to see that $\mathcal{P}(f_1 \circ \pi_S, \ldots, f_k \circ \pi_S; \epsilon) = \mathcal{P}(f_1, \ldots, f_k; \epsilon)$ for any π_S.

In order to express uniqueness results cleanly, we abuse language in the following way: For particular k, n, and ϵ, we say that $(f_1, \ldots, f_k)$ is the unique best protocol "up to π_S" if the set of best protocols is exactly $\{(f_1 \circ \pi_S, \ldots, f_k \circ \pi_S) : S \subseteq [n], \pi \in S_n\}$.

1.2 Main results

Methods of discrete Fourier analysis (see [KKL88, BKS99, MO02] for background) give an exact solution to our problem in the cases $k = 2, 3$, and the best protocol, up to π_S, is for all parties to use the function $f(x) = x_1$. We attribute the case $k = 2$ in the following theorem to folklore.

Theorem 1.2 *For all k, n, ϵ, if we wish to maximize the expression*

$$\mathbf{E}[\#(i,j) : f_i(y^i) = f_j(y^j)], \tag{1}$$

the unique best protocol up to π_S is given by $f_1 = \cdots = f_k = f$, where $f(x) = x_1$. In particular, if $k = 2$ or $k = 3$, then for all n and ϵ, the unique best protocol up to π_S for maximizing $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$ is given by $f_1 = \cdots = f_k = f$, where $f(x) = x_1$.

In general we do not know how to find the optimal protocol. However, we can prove some general properties of the protocols which maximize $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$. Recall that a function f is *monotone* if for all $x, y \in \{0,1\}^n$, we have $f(x) \leq f(y)$ whenever $x \preceq y$ (in the sense $x_i \leq y_i$ for all i). For $x, y \in \{0,1\}^n$, we write $x \preceq_L y$ if $\sum_{i=1}^m x_i \leq \sum_{i=1}^m y_i$ for every $m = 1 \ldots n$. We call a function f *left-monotone* if $f(x) \leq f(y)$ whenever $x \preceq_L y$. Note that the partial order induced by $\preceq_L$ is a refinement of the partial order induced by $\preceq$; in particular, every left-monotone function is monotone.

The following theorem is based on convexity and on the Steiner symmetrization principle (see e.g. [T93] for background).

Theorem 1.3 *For all k, n, and ϵ, any protocol which maximizes $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$ among all protocols satisfies $f_1 = \cdots = f_k = f$, where f is left-monotone (up to π_S). This theorem remains true if the phrase "protocol" is everywhere replaced by "antisymmetric protocol".*

So far we haven't ruled out the possibility that the optimal protocol always consists of taking just one bit. For r an odd number, let MAJ_r denote the majority function on the first r bits; i.e., $\mathsf{MAJ}_r(x)$ is 1 if $\sum_{i=1}^r x_i > r/2$ and $\mathsf{MAJ}_r(x) = 0$ if $\sum_{i=1}^r x_i < r/2$. Using a coupling argument, we prove the following result:

Theorem 1.4 *For all n odd, and all ϵ, there exists a $K = K(n, \epsilon)$ such that for $k \geq K$, the unique best protocol up to π_S is given by $f_1 = \cdots = f_k = \mathsf{MAJ}_n$. Moreover, as $k \to \infty$,*

$$\mathcal{P}_k(\mathsf{MAJ}_n; \epsilon) = \Theta\left((1 - \mathbf{P}[Bin(n, \epsilon) > n/2])^k\right), \tag{2}$$

where $Bin(n, \epsilon)$ is a binomial variable with parameters n and ϵ. (This should be compared to $\Theta((1-\epsilon)^k)$ for the function $f(x) = x_1$.) When n is even, a similar result is true; in place of MAJ_n, one should take any balanced funciton f which has $f(x) = 1$ whenever $|\{i : x_i = 1\}| > |\{i : x_i = 0\}|$, and $f(x) = 0$ whenever $|\{i : x_i = 0\}| > |\{i : x_i = 1\}|$.

A dual result is obtained by fixing n and k, and letting ϵ be either close to 0 or close to 1/2.

Theorem 1.5 *For all k and n, there exist $0 < \epsilon' = \epsilon'(n,k) < \epsilon'' = \epsilon''(n,k) < 1/2$, such that for all $0 < \epsilon < \epsilon'$, or $\epsilon'' < \epsilon < 1/2$, the unique best protocol up to π_S is given by $f_1 = \cdots = f_k = f$, where $f(x) = x_1$; i.e., $f = \mathsf{MAJ}_1$.*

It may now seem like the optimal protocol consists of either taking all functions to be MAJ_n or all functions to be MAJ_1. This is not the case however, as a computer-assisted proof shows that sometimes MAJ_r is better than MAJ_1 and MAJ_n for $1 < r < n$. See Proposition G.2.

Despite Theorem 1.4, it is not true that as $k \to \infty$, the success probability of the best protocol goes to 0 exponentially fast in k (treating ϵ as fixed). In fact, if we allow n to be an unbounded function of k, then the best protocol's success probability is at least inverse-polynomially large in k.

Theorem 1.6 *Fix ϵ. Then there exists a sequence (n_k) such that*

$$\mathcal{P}_k(\mathsf{MAJ}_{n_k}; \epsilon) \geq \Omega\left(\frac{1}{k^{2.01/(1-2\epsilon)^2}}\right).$$

It suffices to have $n_k = O(k^{4.01/(1-2\epsilon)^2})$.

Finally, it is natural to ask if the optimal function is always MAJ_r for some $1 \leq r \leq n$ (assuming, say, n is odd).

Conjecture M: *For a particular k, ϵ, and odd n, there is a function $f \in \mathcal{A}_n$ which is not a majority function such that $\mathcal{P}_k(f; \epsilon) > \mathcal{P}_k(\mathsf{MAJ}_m; \epsilon)$ for all odd $m \leq n$.*

Conjecture O: *For any given k, ϵ, and odd n, there is an odd $m \leq n$ such that the best antisymmetric function for the parties is MAJ_m.*

In fact, we know of no counterexample to Conjecture O even if we allow the parties to use any balanced function (which could allow for a biased output). Some evidence that resolving this conjecture could possibly be hard: One, it is not true that for any non-majority function f, and any fixed k, there is a majority function which dominates f over all ϵ — we have a computer-verified counterexample. Two, for certain k, ϵ, $\mathcal{P}_k(\mathsf{MAJ}_n, \epsilon)$ is not even unimodal as a function of n. E.g., for $k = 12$, $\epsilon = 0.1$, the success probability decreases between MAJ_1 and MAJ_3, increases up to MAJ_{11}, and then decreases again out to MAJ_{17} (and appears to continue decreasing from this point on).

We conclude the introduction with a road map to the following sections. In Section 2 we prove Theorem 1.2 using Fourier analysis. In Section 3 we prove the Theorem 1.3 using Steiner symmetrization and convexity. In Section 4 we prove Theorems 1.4, 1.5, and 1.6 — the arguments in this section are mostly probabilistic. Finally, in Section 5 we discuss the results of some computer analysis, and pose two more open problems.

2 Fourier methods

In this section, we make a usual notational switch; the bits 0 and 1 will be denoted by $+1$ and -1, respectively. Note that $f : \{+1,-1\}^n \to \{+1,-1\}$ is balanced iff $\mathbf{E}[f] = 0$. Given the functions of the parties $f_1, \ldots, f_k : \{+1,-1\}^n \to \{+1,-1\}$, we view these as functions in the larger space $\{+1,-1\}^{n\times(k+1)}$ in the natural way: $f_i(x, y^1, \ldots, y^k) = f_i(y^i)$. Our probability space gives rise to a natural inner product on functions $f, g : \{+1,-1\}^{n\times(k+1)} \to \mathbb{R}$:

$$\langle f, g\rangle = \mathop{\mathbf{E}}_{x,y^1,\ldots,y^k}[f(x, y^1, \ldots, y^k)g(x, y^1, \ldots, y^k)]. \tag{3}$$

Lemma 2.1

$$\sum_{i,j=1}^{k} \langle f_i, f_j\rangle = 2\mathbf{E}[\#(i,j) : f_i(y^i) = f_j(y^j)] - k^2.$$

Proof: Since the f_i are ± 1 valued functions,

$$\begin{aligned}\sum_{i,j=1}^{k} \langle f_i, f_j\rangle &= \sum_{i,j=1}^{k} \big(\mathbf{P}[f_i(y^i) = f_j(y^j)] - \mathbf{P}[f_i(y^i) \neq f_j(y^j)]\big) \\ &= \sum_{i,j=1}^{k} \big(2\mathbf{P}[f_i(y^i) = f_j(y^j)] - 1\big) = 2\mathbf{E}[\#(i,j) : f_i(y^i) = f_j(y^j)] - k^2.\end{aligned}$$

□

Now, in order to maximize the quantity in (1), we analyze the scalar products $\langle f_i, f_j\rangle$. In order to analyze scalar products, it is useful to work with the Fourier basis. We refer the reader to [KKL88, BKS99, MO02] for background. For a set $S \subseteq [n]$, we let $U_{i,S} : \{+1,-1\}^{n\times(k+1)} \to \{+1,-1\}$ be defined by:

$$U_{i,S}(x, y^1, \ldots, y^k) = U_{i,S}(y^i) = \prod_{j\in S} y^i_j.$$

Since the k coordinates are independent, it follows that if $S \neq S'$ then for all i, i' it holds that

$$\langle U_{i,S}, U_{i',S'}\rangle = 0. \tag{4}$$

Moreover, if $i \neq i'$ then

$$\langle U_{i,S}, U_{i',S}\rangle = \mathbf{E}[\prod_{j\in S} y^i_j y^{i'}_j] = \prod_{j\in S} \mathbf{E}[y^i_j y^{i'}_j] = (1-2\epsilon)^{2|S|}, \tag{5}$$

and the functions $U_{i,S}$ all have norm 1, so $\langle U_{i,S}, U_{i,S}\rangle = 1$.

Lemma 2.2 *Let $i \neq j$. Then*

$$\max_{f_i, f_j \in \mathcal{B}_n} \langle f_i(y^i), f_j(y^j)\rangle = (1-2\epsilon)^2,$$

and the maximum is obtained when $f_i = f_j = f$ and $f(x) = \pm x_r$ for some $1 \leq r \leq n$.

Proof: Express f_i and f_j in terms of their Fourier expansion, $f_i = \sum_{S\subseteq[n]} \hat{f}_i(S)U_{i,S}$ and similarly for f_j. Since, both f_i and f_j are balanced, and $\mathbf{E}[U_{i,S}] = \mathbf{E}[U_{j,S}] = 0$ for nonempty S, it follows that $\hat{f}_i(\emptyset) = \hat{f}_j(\emptyset) = 0$. Now by (4) and (5) it follows that

$$\langle f_i, f_j\rangle = \sum_{\emptyset\neq S\subseteq[n]} \hat{f}_i(S)\hat{f}_j(S)(1-2\epsilon)^{2|S|}. \tag{6}$$

Hence we have:

$$\begin{aligned}
\langle f_i, f_j\rangle &= \sum_{\emptyset\neq S\subseteq[n]} (1-2\epsilon)^{|S|}\hat{f}_i(S)(1-2\epsilon)^{|S|}\hat{f}_j(S) \\
&\leq \sqrt{\sum_{\emptyset\neq S\subseteq[n]} (1-2\epsilon)^{2|S|}\hat{f}_i(S)^2}\sqrt{\sum_{\emptyset\neq S\subseteq[n]} (1-2\epsilon)^{2|S|}\hat{f}_j(S)^2} \qquad \text{(Cauchy-Schwarz)} \\
&\leq \sqrt{\sum_{\emptyset\neq S\subseteq[n]} (1-2\epsilon)^{2}\hat{f}_i(S)^2}\sqrt{\sum_{\emptyset\neq S\subseteq[n]} (1-2\epsilon)^{2}\hat{f}_j(S)^2} \\
&= (1-2\epsilon)^2,
\end{aligned}$$

as $\sum \hat{f}_i(S)^2 = \sum \hat{f}_j(S)^2 = 1$. The second inequality is tight only if f_i and f_j have Fourier degree 1. Note that if $f(x) = \sum_{|S|=1} \hat{f}(S)u_S(x)$ is a function which is ± 1 valued, then for all S of size 1, it holds that $2\hat{f}(S) = f(x) - f(x\oplus e_S) \in \{-2, 0, 2\}$. It follows that $f(x) = \pm x_r$ for some r.

In this case, the first inequality is tight only if f_i and f_j are the same one-bit function. Hence, as claimed, $f_i = f_j = f$ where $f(x) = \pm x_r$ constitutes the only maximizing solution. □

We can now prove Theorem 1.2.

Proof: [of Theorem 1.2] By Lemma 2.1 it follows that maximizing $\mathbf{E}[\#(i,j) : f_i(y^i) = f_j(y^j)]$ is the same as maximizing,

$$\sum_{i,j=1}^{k} \langle f_i, f_j\rangle = k + \sum_{i\neq j}\langle f_i, f_j\rangle.$$

By Lemma 2.2, the above sum is maximized when $f_1 = \cdots = f_k = f$, and $f(x) = x_1$ up to π_S. We thus obtain the first assertion of the theorem.

For the second assertion, note that when $k = 2$,

$$\mathbf{E}[\#(i,j) : f_i(y^i) = f_j(y^j)] = 2 + 2\mathbf{P}[f_1(y^1) = f_2(y^2)],$$

while when $k = 3$,

$$\mathbf{E}[\#(i,j) : f_i(y^i) = f_j(y^j)] = 5 + 4\mathbf{P}[f_1(y^1) = f_2(y^2) = f_3(y^3)],$$

so the second assertion follows. □

3 Convexity and symmetrization

We now show that to maximize $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$, it suffices to look at restricted sets of functions. The methods in the section are related to convexity in general and the Steiner symmetrization in particular, see e.g. [T93] for background.

We begin by using convexity to show that all parties should use the same function:

Proposition 3.1 *Fix k, n, and ϵ. Let $\mathcal{C}$ be any class of boolean functions on n bits. Subject to the restriction that $f_1, \ldots, f_k \in \mathcal{C}$, every protocol which maximizes $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$ has $f_1 = \cdots = f_k$.*

Proof: Let $\mathcal{C} = \{f_1, f_2, \ldots, f_M\}$, and assume $M > 1$, else the proposition is trivial. Suppose that among the k parties, exactly t_j use the function f_j. Then clearly,

$$t_j \geq 0, \qquad \sum_{j=1}^{M} t_j = k, \qquad t_j \in \mathbb{Z}. \tag{7}$$

By the first part of Lemma A.2, the probability that all parties agree is:

$$\mathcal{P} = \sum_{x \in \{0,1\}^n} 2^{-n} \left(\prod_{j=1}^{M} \left(T_\epsilon(f_j)(x)\right)^{t_j} + \prod_{j=1}^{M} \left(1 - T_\epsilon(f_j)(x)\right)^{t_j} \right). \tag{8}$$

Note that each $T_\epsilon(f_j)(x) \in [0,1]$ and that for any $c \in [0,1]$, the function $g(t) = c^t$, is *log-convex* (since $\log c^t = t \log c$ is linear).[1] Therefore the function $g_1 \cdots g_M : \mathbb{R}^M \to \mathbb{R}$ given by $(t_1, \ldots, t_M) \mapsto \prod_{j=1}^{M} g_j(t_j)$ is a log convex function, and therefore a convex function.

Since the sum of convex functions is also convex, $\mathcal{P}$ is a convex function of the t_j's. We wish to maximize $\mathcal{P}$ subject to the restrictions (7).

If we relax the assumption $t_j \in \mathbb{Z}$ to $t_j \in \mathbb{R}$, we are simply maximizing a convex function over a convex bounded polytope. The vertices of the polytope are simply the points of the form $(0, \ldots, 0, k, 0, \ldots, 0)$. By convexity theory, the maximum must occur at a vertex, and so it follows that there is at least one maximizing protocol in which all players use the same function.

We defer the proof that $\mathcal{P}$ attains its maximum *only* at vertices to Appendix A. □

Next we use Steiner symmetrization principle in order to obtain more information about functions which optimize $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$. Recall that for $x, y \in \{0,1\}^n$, we write $x \preceq y$ if for all $i \in [n]$ it holds that $x_i \leq y_i$, and we say that f is monotone if $f(x) \leq f(y)$ whenever $x \preceq y$. As a generalization, let $T \subseteq [n]$, and write $x \preceq_T y$ if $x_i \leq y_i$ for $i \in T$, and $y_i \leq x_i$ for $i \notin T$. If a function is monotone with respect to $\preceq_T$ for some T, we call it *unate*.

Proposition 3.2 *Let $\mathcal{C}$ stand for either $\mathcal{B}_n$ or A_n. For any k, n, ϵ, if f is restricted to be in $\mathcal{C}$, then the maximum of $\mathcal{P}_k(f; \epsilon)$ occurs at a monotone function f, and* every *maximum occurs at a unate function f.*

[1] In fact, this doesn't make sense for $c = 0$, but the conclusion — that $\mathcal{P}$ is convex — still clearly holds.

Proof: See Appendix B. □

Recall that for $x, y \in \{0,1\}^n$, we write $x \preceq_L y$ if $\sum_{i=1}^m x_i \leq \sum_{i=1}^m y_i$ for every $m = 1 \ldots n$, and that we call $f : \{0,1\}^n \to \{0,1\}$ left-monotone, if $f(x) \leq f(y)$ whenever $x \preceq_L y$.

Proposition 3.3 *Let $\mathcal{C}$ stand for either $\mathcal{B}_n$ or A_n. For any k, n, and ϵ, if f is restricted to be in $\mathcal{C}$, then any function which maximizes $\mathcal{P}_k(f;\epsilon)$ must be left-monotone, up to π_S.*

Proof: See Appendix C. □

Proof: [of Theorem 1.3] The proof follows from Propositions 3.1, 3.2, and 3.3. □

4 Majorities

In this section we study majority functions and show that these function are optimal for some limiting values of k and ϵ.

4.1 Fixed ϵ, n; $k \to \infty$

We start by proving Theorem 1.4. Given a function $f : \{0,1\}^n \to \{0,1\}$, let $p_1(f, x, \epsilon)$ denote the probability that $f(y) = 1$, given that y is an ϵ-corrupted version of the string $x \in \{0,1\}^n$. Let p_0 be defined similarly for the probability that $f(y) = 0$.

Proposition 4.1 *Fix ϵ, and let f be monotone. Then as a function of x, $p_1(f, x, \epsilon)$ is maximized at $x = \vec{1} = 1 \cdots 1$, and $p_0(f, x, \epsilon)$ is maximized at $x = \vec{0} = 0 \cdots 0$.*

Proof: We prove the claim for p_1, the proof for p_0 being the same. Note that flipping each bit of a string with probability ϵ is the same as *updating* each bit with probability 2ϵ, where an update consists of replacing the bit with a random choice from $\{0,1\}$.

Let $x \in \{0,1\}^n$ be any sequence. Let x' be an ϵ corrupted version of x, and $\vec{1}'$ be an ϵ corrupted version of $\vec{1}$. We claim that we can couple the random variable x' and $\vec{1}'$ in such a way that $x' \preceq \vec{1}'$.

The coupling is achieved in the following simple way: update the same bits of x and $\vec{1}$ with the same values. Clearly, we have $x' \preceq \vec{1}'$. Hence by monotonicity if $f(x') = 1$, then $f(\vec{1}') = 1$. The result follows, as

$$p_1(f, x, \epsilon) = \mathbf{P}[f(x') = 1] \leq \mathbf{P}[f(\vec{1}') = 1] = p_1(f, \vec{1}, \epsilon).$$

□

Proposition 4.2 *For fixed ϵ, $p_1(f, \vec{1}, \epsilon)$ and $p_0(f, \vec{0}, \epsilon)$ are maximized among $f \in \mathcal{B}_n$ by any function which is 0 on all strings with fewer than $n/2$ 1's. In particular, if n is odd, $f = \mathsf{MAJ}_n$ is the unique maximizing function.*

Proof: We prove the assertion about p_1.

$$p_1(f, \vec{1}, \epsilon) = \sum_{x \in f^{-1}(1)} (1-\epsilon)^{n - \Delta(x,\vec{1})} \epsilon^{\Delta(x,\vec{1})},$$

where Δ denotes Hamming distance. The quantity being summed is a strictly decreasing function of $\Delta(x, \vec{1})$. The result follows. $\Box$

Proof: [of Theorem 1.4] We prove the theorem for n odd. The proof for n even is essentially the same. By Theorem 1.3, we may assume without loss of generality that all parties use the same monotone function $f \in \mathcal{B}_n$. Now:

$$\begin{aligned} \mathcal{P}(\mathsf{MAJ}_n, k, \epsilon) &= 2^{-n} \sum_{x \in \{0,1\}^n} \left(p_1(\mathsf{MAJ}_n, x, \epsilon)^k + (p_0(\mathsf{MAJ}_n, x, \epsilon))^k \right) \\ &\geq 2^{-n} \left(p_1(\mathsf{MAJ}_n, \vec{1}, \epsilon)^k + p_0(\mathsf{MAJ}_n, \vec{0}, \epsilon)^k \right). \end{aligned} \tag{9}$$

By Proposition 4.1, if f is monotone, then

$$\begin{aligned} \mathcal{P}(f, k, \epsilon) &= 2^{-n} \sum_{x \in \{0,1\}^n} \left(p_1(f, x, \epsilon)^k + (p_0(f, x, \epsilon))^k \right) \\ &\leq 2^{-n} \sum_{x \in \{0,1\}^n} \left(p_1(f, \vec{1}, \epsilon)^k + p_0(f, \vec{0}, \epsilon)^k \right) = p_1(f, \vec{1}, \epsilon)^k + p_0(f, \vec{0}, \epsilon)^k. \end{aligned} \tag{10}$$

By Proposition 4.2, if $f \in \mathcal{B}_n$ is monotone, and $f \neq \mathsf{MAJ}_n$, then $p_1(f, \vec{1}, \epsilon) < p_1(\mathsf{MAJ}_n, \vec{1}, \epsilon)$ and $p_0(f, \vec{0}, \epsilon) < p_0(\mathsf{MAJ}_n, \vec{0}, \epsilon)$. Hence for sufficiently large k, we will have

$$2^{-n} p_1(\mathsf{MAJ}_n, \vec{1}, \epsilon)^k > p_1(f, \vec{1}, \epsilon)^k, \quad 2^{-n} p_0(\mathsf{MAJ}_n, \vec{0}, \epsilon)^k > p_0(f, \vec{0}, \epsilon)^k. \tag{11}$$

Combining (9), (10) and (11) we obtain that $\mathcal{P}_k(\mathsf{MAJ}_n; \epsilon) \geq \mathcal{P}_k(f; \epsilon)$ for all monotone $f \in \mathcal{B}_n$ as needed.

Bound (2) follows from (9), (10), and (11) once we note that

$$p_0(\mathsf{MAJ}_n, \vec{0}, \epsilon) = p_1(\mathsf{MAJ}_n, \vec{1}, \epsilon) = 1 - \mathbf{P}[Bin(n, \epsilon) > n/2].$$

$\Box$

4.2 Fixed k, n; $\epsilon \to 0$ or $1/2$

Proposition 4.3 *For all k and n, there exists $\epsilon'(k, n) > 0$, such that for all $0 < \epsilon < \epsilon'(k, n)$, the unique best protocol up to π_S for maximizing $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$ is given by $f_1 = \cdots = f_k = f$, where $f(x) = x_1$.*

Proof: See Appendix D. $\Box$

Proposition 4.4 *For all k and n, there exists $\epsilon'(k, n) < 1/2$, such that for all $eps'(k, n) < \epsilon < 1/2$, the unique best protocol up to π_S for maximizing $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$ is given by $f_1 = \cdots = f_k = f$, where $f(x) = x_1$.*

Proof: See Appendix E. $\Box$

4.3 Fixed ϵ; $k \to \infty$ with n unbounded

We prove Theorem 1.6 in Appendix F. This theorem gives a relatively good lower bound on $\lim_{n\to\infty} \mathcal{P}_k(\mathsf{MAJ}_n;\epsilon)$. In fact, as $n \to \infty$, all distributions involved in the calculation of $\mathcal{P}_k(\mathsf{MAJ}_n;\epsilon)$ become normal, and it is possible to get a couple of more or less closed forms for the limit:

Proposition 4.5

$$\lim_{\substack{n\to\infty \\ n \text{ odd}}} \mathcal{P}_k(\mathsf{MAJ}_n;\epsilon) = 2\,\mathbf{E}\left[\Phi\left(\frac{X}{\sqrt{c(\epsilon)}}\right)^k\right] \tag{12}$$

$$= \frac{2\sqrt{c(\epsilon)}}{(2\pi)^{\frac{1}{2}(c(\epsilon)-1)}} \int_0^1 x^k I(x)^{c(\epsilon)-1}\,\mathrm{d}x, \tag{13}$$

where $c(\epsilon) := \frac{4\epsilon(1-\epsilon)}{(1-2\epsilon)^2} \in (0,\infty)$, *$X$ is a standard normal random variable, and* $I(x) := \phi(\Phi^{-1}(x))$.

(We thank Nati Srebro for his help in calculating (12).) The second formula (13) can be used to get tighter bounds than in Theorem 1.6. For example, from (13) we get that $\lim_{n\to\infty} \mathcal{P}_k(\mathsf{MAJ}_n; 1/2 - \sqrt{2}/4) = 2/(k+1)$.

5 Computer-assisted results and open problems

The problem well avails itself to analysis by computer. In Appendix G we give a nearly complete computer-assisted analysis of the case $n = 5$.

We end with two open problems that computer-assisted analysis has led us to consider:

Open Problem 1 *Prove or disprove: For fixed n, ϵ, and $2 \le k \le 9$, the best antisymmetric protocol is for all parties to use* MAJ_1.

Open Problem 2 *Prove or disprove: There is a universal constant $C < \infty$ such that for every k, ϵ,*

$$\mathcal{P}_k(\mathsf{MAJ}_{n^*};\epsilon) \le C \lim_{\substack{n\to\infty \\ n \text{ odd}}} \mathcal{P}(\mathsf{MAJ}_n, k, \epsilon),$$

where n^* is any odd number (presumably maximizing $\mathcal{P}_k(\mathsf{MAJ}_{n^*};\epsilon)$). I.e., the limiting value of $\mathcal{P}_k(\mathsf{MAJ}_n;\epsilon)$ is no worse than the success probability of the best majority, up to a constant factor.

The worst constant C we know to be necessary in Open Problem 2 is $\pi/2$, from the case $k = 2$, $\epsilon \to 1/2$.

If Conjecture O and Open Problem 2 are both verified, then we can get tight (up to a constant) upper and lower bounds for the optimal value of $\mathcal{P}(f_1, \ldots, f_k; \epsilon)$ under antisymmetric protocols, for any k, ϵ, and unrestricted n, by using Proposition 4.5.

Acknowledgments: We are most grateful to Oded Schramm for many helpful suggestions, ideas and proofs. We are also grateful to Nati Srebro for helpful simulations and stimulating discussions.

References

[Be75] W. Beckner, Inequalities in Fourier analysis, *Annals of Math.* **102** (1975), 159–182.

[BKS99] I. Benjamini, G. Kalai, O. Schramm. Noise sensitivity of boolean functions and applications to percolation. *Institute des Haute Études Scientifique*, vol 90, pp 5–43, 1999.

[Bo70] A. Bonami, Étude des coefficients Fourier des fonctiones de $L^p(G)$, *Ann. Inst. Fourier* **20** (1970), 335–402.

[BL90] M. Ben-Or, N. Linial. Collective coin flipping. In *Randomness and Computation*, S. Micali ed. Academic Press, New York, 1990.

[CT91] T. M. Cover, J. A. Thomas. Elements of Information Theory. Published by John Wiley and Sons, 1991.

[D00] Y. Dodis. Fault-tolerant leader election and Collective coin-flipping in the full information model — survey. In preparation.

[DR01] Y. Ding, M. Rabin. Hyper-Encryption, and Everlasting Security. To appear; see also Ph.D. thesis of Y. Ding.

[F68] W. Feller. An Introduction to Probability Theory and Its Applications, vols 1 and 2. Published by John Wiley and Sons, 3rd edition, 1968.

[KKL88] J. Kahn, G. Kalai, N. Linial. The influence of variables on boolean functions. *Foundations of Computer Science*, 1988.

[M97] U. Maurer. Information-theoretically secure secret-key agreement by NOT authenticated public discussion. In *Advances in Cryptology — EUROCRYPT '97*, LCS vol 1233, pp 209–225, 1997.

[MO02] E. Mossel, R. O'Donnell. On the noise sensitivity of monotone functions. *Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, 2002.

[T93] G. Talenti. The standard isoperimetric theorem. In *Handbook of convex geometry*, 73–123, North-Holland, Amsterdam, 1993.

A Uniqueness for Proposition 3.3

We begin with a definition and a simple Fourier Lemma.

Definition A.1 *For $f : \{0,1\}^n \to \mathbb{R}$ given by $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) u_S(x)$, let*

$$T_\epsilon(f)(x) = \sum_{S \subseteq [n]} \hat{f}(S)(1-2\epsilon)^{|S|} u_S(x).$$

(The operator T_ϵ was first defined in [Be75, Bo70]; see also [KKL88, BKS99].)

Lemma A.2 *Given ϵ and $f : \{0,1\}^n \to \{0,1\}$, $T_\epsilon(f)(x)$ is equal to the probability that a particular party using f outputs 1, given that the source string is x. If f and g are different boolean functions on $\{0,1\}^n$, then for every $0 < \epsilon < 1/2$, there exists some $x \in \{0,1\}^n$ for which $T_\epsilon(f)(x) \neq T_\epsilon(g)(x)$.*

Proof: Note that given a source x and an ϵ-corrupted version of x called y, the expected value of $u_S(y)$ is $(1-2\epsilon)^{|S|} u_S(x)$ (see e.g. [BKS99] for a detailed proof). Therefore, by linearity of expectation, it follows that for all $f : \{0,1\}^n \to \mathbb{R}$, $T_\epsilon(f)(x)$ is the expected value of $f(y)$. In particular, if $f : \{0,1\}^n \to \{0,1\}$, then $T_\epsilon(f)(x) = \mathbf{E}[f(y)] = \mathbf{P}[f(y) = 1]$ and we have proved the first assertion of the lemma.

For the second assertion note that for $0 < \epsilon < 1/2$, it is obvious from T_ϵ's definition that it is a reversible linear transformation on the space of all functions from $\{0,1\}^n \to \mathbb{R}$. □

Now we complete the proof of Proposition 3.1. It remained to show that $\mathcal{P}$ doesn't obtain the maximum at any point which is not a vertex of the polytope. Note that by convexity, if $\mathcal{P}$ has a maximum which is not a vertex of the polytope, then there exists an interval $I = \{\lambda v_1 + (1-\lambda) v_2 : \lambda \in [0,1]\}$, where v_1 and v_2 are vertices of the polytope, such that f is a constant function on I. Therefore if we could show that f is strictly convex on I (as a function of t), then it will follow that the maximum is obtained only at vertices of the polytope.

Note that when restricted to the edge I joining, e.g., $v_1 = (k, 0, \ldots, 0)$ and $v_2 = (0, k, 0, \ldots, 0)$, $\mathcal{P}$ is given by:

$$\mathcal{P} = \sum_{x \in \{0,1\}^n} 2^{-n} \left((T_\epsilon(f_1)(x))^{\lambda k} (T_\epsilon(f_2)(x))^{(1-\lambda)k} + (1 - T_\epsilon(f_1)(x))^{\lambda k} (1 - T_\epsilon(f_2)(x))^{(1-\lambda)k} \right).$$

By Lemma A.2, we can find an x_0 such that $T_\epsilon(f_1)(x_0)$ and $T_\epsilon(f_2)(x_0)$ differ. Therefore the function

$$(T_\epsilon(f_1)(x_0))^{\lambda k} (T_\epsilon(f_2)(x_0))^{(1-\lambda)k} = (T_\epsilon(f_2)(x_0))^k \left(\frac{T_\epsilon(f_1)(x_0)}{T_\epsilon(f_2)(x_0)} \right)^{\lambda k},$$

is strictly convex, and $\mathcal{P}$ is strictly convex on I as needed.

B Proof of Proposition 3.4

Proof: [of Proposition 3.2] Let $f \in \mathcal{C}$ be any function which maximizes $\mathcal{P}_k(f;\epsilon)$ among functions in $\mathcal{C}$. Let f' be obtained from f by "shifting" up in the first coordinate: Given $x \in \{0,1\}^{n-1}$,

- if $f(0x) = f(1x)$, then set $f'(0x) = f'(1x) = f(0x) = f(1x)$;

- if $f(0x) \neq f(1x)$ then set $f'(0x) = 0$, $f'(1x) = 1$.

It is easy to see that in the case $\mathcal{C} = \mathcal{B}_n$, f' remains in $\mathcal{C}$; a little thought reveals that this is again true in the case $\mathcal{C} = \mathcal{A}_n$.

For $y \in \{0,1\}^n$, let $\tilde{y} \in \{0,1\}^{n-1}$ be the last $n-1$ bits of y. We claim that $\mathcal{P}_k(f';\epsilon) \geq \mathcal{P}(f;\epsilon)$. To show this, it suffices to show that for all $z^1, \ldots, z^k \in \{0,1\}^{n-1}$,

$$\mathbf{P}[f'(y^1) = \cdots = f'(y^k) \mid \tilde{y}^1 = z^1, \ldots, \tilde{y}^k = z^k] \geq \mathbf{P}[f(y^1) = \cdots = f(y^k) \mid \tilde{y}^1 = z^1, \ldots, \tilde{y}^k = z^k]. \tag{14}$$

So suppose each y^i's last $n-1$ bits are fixed to be z^i. Given z^i, $f(y^i)$ is a function from $\{0,1\}$ to $\{0,1\}$, and is therefore either the constant function 0, the constant function 1, the identity function id, or the function $x \to \bar{x}$, which we denote by $\overline{id}$.

If $f(y^i)$ is already determined by z^i, then so is $f'(y^i)$ and the determined value is the same. Otherwise, $f(y^i)$ is a function of the one remaining unknown bit, y^i_1, and is either the function id or $\overline{id}$. In *either* case, $f'(y^i)$ is the identity function on y^i_1.

Assume that given $(z^1, \ldots, z^k)$, there are $a+b$ undetermined functions $f(y^i_1)$, with a of them id, and b of them $\overline{id}$. The probability that all of these functions agree on 0 (or 1) is

$$q = \frac{1}{2}\left((1-\epsilon)^a\epsilon^b + \epsilon^a(1-\epsilon)^b\right),$$

and the probability that all of the undetermined f''s agree on 0 (or 1) is

$$q' = \frac{1}{2}\left((1-\epsilon)^{a+b} + \epsilon^{a+b}\right).$$

There are three cases to consider:

- If some of the determined functions are determined to be 0 and some to be 1, then both terms in (14) are zero.
- If all of the determined functions are determined to be 0 (1), then the left side of (14) is q' and the right side of (14) is q.
- If there are no determined functions, then the left side of (14) is $2q'$ and the right side of (14) is $2q$.

Therefore the claim will follow once we show that $q' \geq q$. But this is:

$$\frac{1}{2}\left((1-\epsilon)^{a+b} + \epsilon^{a+b}\right) \geq \frac{1}{2}\left((1-\epsilon)^a\epsilon^b + \epsilon^a(1-\epsilon)^b\right) \tag{15}$$

$$\Leftrightarrow \quad 1 + \left(\frac{\epsilon}{1-\epsilon}\right)^{a+b} \geq \left(\frac{\epsilon}{1-\epsilon}\right)^b + \left(\frac{\epsilon}{1-\epsilon}\right)^a, \tag{16}$$

which follows by the convexity of the function $t \to \left(\frac{\epsilon}{1-\epsilon}\right)^t$.

Thus we've established $\mathcal{P}_k(f';\epsilon) \geq \mathcal{P}(f;\epsilon)$. We further claim that this inequality is strict unless f was already monotone or anti-monotone on the first coordinate. If f is neither monotone nor anti-monotone on the first coordinate, then there exist z^1 and z^2 such that f, when the last $n-1$ coordinates are restricted to z^1, becomes id, and when the last $n-1$ coordinates are restricted to z^2, becomes $\overline{id}$. Picking $z^3, \ldots, z^k$ so that all the other restricted functions are either id or $\overline{id}$, we obtain $a, b \geq 1$, so (16) is strict inequality and therefore $q' > q$.

Repeating the above argument for all other coordinates, it follows that any maximizing function f must be unate, and that there exists a maximizing function which is monotone. □

C Proof of Proposition 3.5

Proof: [of Proposition 3.3 The proof is similar to the proof of Proposition 3.2, so we will be more brief. By Proposition 3.2, we may assume that f is monotone.

Now apply a new sort of shift to f. Suppose we fix all but two input bits to f. Since f is monotone, there are only 6 possibilities for what the restricted function is; its support may be $\emptyset$, $\{11\}$, $\{11,10\}$, $\{11,01\}$, $\{11,10,01\}$ or $\{11,10,01,00\}$. Define f' to be the same function in all cases except when the support is $\{11,01\}$; in this case, switch it to $\{11,10\}$. This rule preserves balance and asymmetry.

We want to show that $\mathcal{P}_k(f';\epsilon) \geq \mathcal{P}_k(f;\epsilon)$. As before, we condition on all but two bits of each of $y_1, \ldots, y_k$, and show that f' is better. Say that under this conditioning, a of the $f(y^i)$'s restrict to the function with support $\{11,10\}$, and b of the $f(y^i)$'s restrict to the function with support $\{11,01\}$. Since all other possible restricted functions have the same value for 01 as they do for 10, it suffices again to compare the probability with which the $a+b$ functions agree on 1 with the probability that the corresponding shifted functions agree on 1. Further, by symmetry, we need only consider the cases when the two source bits from x are different (otherwise f and f' do equally well).

So considering the two cases — the source bits are 10 or the source bits are 01 — we get that the contribution from the f-restricted functions will be $(1/2)((1-\epsilon)^a\epsilon^b) + \epsilon^b(1-\epsilon)^a)$, and the contribution from their shifted versions will be $(1/2)((1-\epsilon)^{a+b}+\epsilon^{a+b})$. As we saw in Proposition 3.2, this latter quantity is always at least the former quantity. Hence the shift can only improve the probability of agreement.

Hence we indeed have $\mathcal{P}_k(f';\epsilon) \geq \mathcal{P}_k(f;\epsilon)$. If we repeatedly apply this shift to all pairs of coordinates, we end up with a left-monotone function.

Note that if none of the shifting operations strictly increased the probability of agreement for f, then for every pair of coordinates (i,j) which were shifted, either all the balanced restrictions of f to coordinates (i,j) have support $\{11,10\}$, or all the balanced restrictions have support $\{11,01\}$. In either case, all the shifting did was replace the function f by a function $f \circ \pi_\emptyset$, where $\pi_\emptyset$ is the transposition of coordinates (i,j). It thus follows that the original function was left monotone up to some $\pi_\emptyset$, as needed. □

D Proof of Proposition 4.3

Proof: [of Proposition 4.3] From Proposition 3.1, it follows that the maximum can only be obtained if $f_1 = \cdots = f_k = f$. Note that the probability that there is more than one corrupted bit is $O(\epsilon^2)$ (the constant in the $O(\cdot)$ *does* depend on k and n). Suppose that only the ith bit for party j was corrupted. Then all the parties will agree if and only if $f(x) = f(x \oplus e_i)$, where $x \oplus e_i$ is the vector x with the ith bit flipped. We therefore obtain,

$$\mathcal{P}_k(f;\epsilon) = (1-\epsilon)^{kn} + k\epsilon(1-\epsilon)^{nk-1}\sum_{i=1}^{n}\mathbf{P}_x[f(x) = f(x \oplus e_i)] + O(\epsilon^2). \tag{17}$$

Writing A for the set $\{x : f(x) = 1\}$, and $\partial_E(A)$ for the *edge-boundary* of the set A,

$$\partial_E(A) = \cup_{i=1}^{n}\{(x, x \oplus e_i) : x \in A, x \oplus e_i \notin A\},$$

we see that

$$\sum_{i=1}^{n}\mathbf{P}_x[f(x) = f(x \oplus e_i))] = n - 2^{-n+1}\partial_E(A).$$

So for small ϵ (compared to k and n), in order to maximize $\mathcal{P}_k(f;\epsilon)$, we should minimize $\partial_E(A)$ for sets such that $|A| = 2^{n-1}$. By the isoperimetric inequality for the cube, the sets A which minimize $\partial_E(A)$ among all sets of size 2^{n-1} are exactly the sets $A = \{x : x_i = 0\}$, or $A = \{x : x_i = 1\}$. Thus f must be $f(x) = x_1$ up to π_S, as claimed. □

E Proof of Proposition 4.4

Proof: [of Proposition 4.4] Again, Proposition 3.1 implies that we need only consider the case $f_1 = \cdots = f_k = f$. In this proof it will be helpful again to assume that the bit values are ± 1, so we want to show that the maximizing functions are $f(x) = x_i$, or $f(x) = -x_i$.

It will be useful to work with the "updating representation". Let $X(i,j)$ for $1 \le i \le k$ and $1 \le j \le n$ be a sequence of i.i.d $\{0,1\}$ variables s.t. $\mathbf{P}[X(i,j) = 1] = \delta = 1 - 2\epsilon$. Note that we may produce the y^i's from x in the following manner. If $X(i,j) = 1$, then $y^i_j = x_j$, otherwise y^i_j is chosen uniformly at random from $\{+1,-1\}$ independently from everything else.

Note that if all the $X(i,j)$ are 0, the inputs to the functions are independent, so for all balanced f's,

$$\mathbf{P}[f(y^1) = \cdots = f(y^k) \mid \sum_{i,j} X_{i,j} = 0] = 2^{-k+1}.$$

Similarly for all balanced f,

$$\mathbf{P}[f(y^1) = \cdots = f(y^k) \mid \sum_{i,j} X_{i,j} = 1] = 2^{-k+1},$$

and for all i, i' and $j \neq j'$,

$$\mathbf{P}[f(y^1) = \cdots = f(y^k) \mid X_{i,j} = X_{i',j'} = 1, \sum_{s,t} X_{s,t} = 2] = 2^{-k+1}.$$

Moreover,

$$\mathbf{P}\left[\sum_{i,j} X_{i,j} > 2\right] = O(\delta^3)$$

(the constant in the $O(\cdot)$ depending on k and n).

We therefore conclude that

$$\mathcal{P}_k(f;\epsilon) = c_k + \delta^2(1-\delta)^{nk-2} \sum_{i \neq i'} \sum_{j=1}^{n} \mathbf{P}[f(y^1) = \cdots = f(y^k) \mid X_{i,j} = X_{i',j} = 1] + O(\delta^3), \qquad (18)$$

where c_k is independent of f. Writing z for a uniformly chosen element of $\{+1,-1\}^n$, and z' for an element which is chosen by picking $1 \leq i \leq n$ uniformly at random, and then choosing $z' \in \{0,1\}$ uniformly among all z' s.t. $z'_i = z_i$, we obtain,

$$\sum_{i \neq i'} \sum_{j=1}^{n} \mathbf{P}[f(y^1) = \cdots = f(y^k) \mid X_{i,j} = X_{i',j} = 1] = nk(k-1)\mathbf{P}[f(z) = f(z')]. \qquad (19)$$

Therefore, if we could show that $\mathbf{P}[f(z) = f(z')]$ is maximized among all balanced functions f when $f(x) = x_1$ up to π_S, then the proof will follow from (18).

In order to prove this claim, we note that

$$\mathbf{P}[f(z) = f(z')] = (1 + \mathbf{E}[f(z)f(z')])/2.$$

It therefore suffices to maximize $\mathbf{E}[f(z)f(z')]$ over all balanced functions f, i.e., functions with $\hat{f}(\emptyset) = \mathbf{E}[f] = 0$. We pass to the Fourier representation as in the proof of Theorem 1.2. It is easy to see that if $u_S(x) = \prod_{i \in S} x_i$, then

$$\mathbf{E}[u_S(z')|z] = \begin{cases} u_S(z) & \text{if } S = \emptyset, \\ u_S(z)/n & \text{if } |S| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

So,

$$\mathbf{E}[f(z)f(z')] = \frac{1}{n} \sum_{|S|=1} \hat{f}^2(S) \leq 1/n,$$

and equality is achieved iff $f(x) = x_i$, or $f(x) = -x_i$ as needed (see proof of Lemma 2.2). □

Proof: [of Theorem 1.5] Follows immediately from Propositions 4.3 and 4.4. □

F Proof of Theorem 1.6

Proof: [of Theorem 1.6] Fix ϵ and k. We consider $\mathcal{P}_k(\mathsf{MAJ}_n;\epsilon)$ as a function of n, as $n \to \infty$ through the odd numbers. Our proof will go by showing that there is at least a $\Omega(1/k^{2.01/(1-2\epsilon)^2})$ chance that the source string x has significantly more 1's than 0's. Then we show that in this case, the probability any particular party says 1 is at least $1 - 1/k$, and hence the probability that all parties say 1 is at least a constant.

Let X be the random variable given by (# 1's in x) $-$ (# 0's in x). By the Central Limit Theorem, as $n \to \infty$, the distribution of X approaches a normal distribution with mean 0 and variance n. Let $c = \frac{2}{1-2\epsilon}$. The probability that an $N(0,n)$ normal variable exceeds $c\sqrt{\log k}\sqrt{n}$ is:

$$1 - \Phi(c\sqrt{\log k}) \geq \frac{1}{2}\frac{1}{c\sqrt{\log k}}\frac{1}{\sqrt{2\pi}}\exp\left(-c^2 \log k/2\right) \geq \Omega(1/k^{2.01/(1-2\epsilon)^2}).$$

Here Φ denotes the cumulative distribution function of a standard normal variable, and the first inequality follows from the fact that $1 - \Phi(x) \geq (1/x - 1/x^3)\phi(x)$ (see [F68]), where $\phi(x)$ is the density function of a standard normal variable .

Given that this happens, pessimistically assume that x contains just $c\sqrt{\log k}\sqrt{n}$ more 1's than 0's; i.e., x contains exactly $n/2 + (c/2)\sqrt{\log k}\sqrt{n}$ 1's. We now show that the probability that a particular party using MAJ_n outputs 1 given x is at least $1 - 1/k$.

Consider the ϵ-corrupted version of x the party sees; call it y. The number of 1's in y is distributed as the sum of n Bernoulli trials, $n/2 + (c/2)\sqrt{\log k}\sqrt{n}$ of which have success probability $1 - \epsilon$, and $n/2 - (c/2)\sqrt{\log k}\sqrt{n}$ of which have success probability ϵ. We can use a single Chernoff bound to upper-bound the probability of getting fewer than $n/2$ 1's in y. The expected number of 1's is $n/2 + (1-2\epsilon)c\sqrt{\log k}\sqrt{n} = n/2 + 2\sqrt{\log k}\sqrt{n}$. Since $n/2 = (1-\delta)(n/2 + 2\sqrt{\log k}\sqrt{n})$ when $\delta = 2\sqrt{\log k}\sqrt{n}/(n/2 + 2\sqrt{\log k}\sqrt{n}) > 2\sqrt{\log k}/\sqrt{n}$, Chernoff tells us that the probability that y has fewer than $n/2$ 1's is at most $\exp(-\frac{4\log k\ (n/2)}{2n}) = 1/k$.

Thus as claimed, given a source string with at least $c\sqrt{\log k}\sqrt{n}$ more 1's than 0's, the probability a particular party outputs 1 is at least $1 - 1/k$. Hence the probability that all parties output 1 is at least $(1-1/k)^k = \Omega(1)$.

Hence by taking n sufficiently large, we can make $\mathcal{P}_k(\mathsf{MAJ}_n;\epsilon) \geq \Omega(1/k^{2.01/(1-2\epsilon)^2})$. By applying the Berry-Esséen bounds on the rate of convergence in the Central Limit Theorem (see [F68]), one can show that it suffices for n to be $O(k^{4.01/(1-2\epsilon)^2})$. □

G Computer-assisted analysis of the case $n = 5$

As stated, the problem well avails itself to analysis by computer. In particular, given any explicit function $f : \{0,1\}^n \to \{0,1\}$, a computer mathematics package can easily calculate $\mathcal{P}_k(f;\epsilon)$ exactly, as a function of k and ϵ. Furthermore, if n is very small, a computer program can enumerate all antisymmetric left-monotone functions on n bits. We determined there are "only" 135 such functions for $n = 7$ and 2470 such functions for $n = 8$. (The number jumps to 319124 for $n = 9$.) Thus for particular small values of n and k, we can completely solve the problem by comparing

an explicit set of polynomials in ϵ on the range $(0, 1/2)$. As an example, we now analyze the case $n = 5$.

There are exactly 7 antisymmetric left-monotone functions on 5 bits; they are MAJ_1, MAJ_3, MAJ_5, and four functions expressible as thresholds: $T_1 = Th(3, 1, 1, 1, 1; 4)$, $T_2 = Th(2, 1, 1, 1, 0; 3)$, $T_3 = Th(3, 2, 2, 1, 1; 5)$, and $T_4 = Th(4, 3, 2, 2, 2; 7)$, where $Th(a_1, \ldots, a_5; \theta)$ is 1 iff $\sum_{i=1}^{5} a_i x_i \geq \theta$. Table G shows $\mathcal{P}_k(f; \epsilon)$ for each of the functions.

For small values of k, we plotted these polynomials for $\epsilon \in (0, 1/2)$. This led to the following facts, which in principle could be proved by elementary analysis:

Fact G.1

- *For $n = 5$, $2 \leq k \leq 9$, and any ϵ, the best antisymmetric protocol is MAJ_1.*
- *For $n = 5$, $k = 10, 11$, there exist $0 < \epsilon'_k < \epsilon''_k < 1/2$ such that MAJ_3 is the best antisymmetric protocol for $\epsilon \in [\epsilon'_k, \epsilon''_k]$, and MAJ_1 is the best antisymmetric protocol for all other ϵ.*
- *For $n = 5$, $k = 12$, there exist $0 < \epsilon'_k < \epsilon''_k < \epsilon'''_k < 1/2$ such that MAJ_5 is the best antisymmetric protocol for $\epsilon \in [\epsilon'_k, \epsilon''_k]$, MAJ_3 is the best antisymmetric protocol for $\epsilon \in [\epsilon''_k, \epsilon'''_k]$, and MAJ_1 is the best antisymmetric protocol for all other ϵ.*

The pattern for $k = 12$ appears to hold for all higher k, with MAJ_5 dominating more and more of the interval, as expected from Theorem 1.4.

At this point we can prove two facts mentioned early on in the paper:

Proposition G.2 *There exist k, ϵ, odd n, and odd $1 < r < n$ such that MAJ_r is a superior protocol to both MAJ_1 and MAJ_n.*

Proof: Substitute $k = 10$, $\epsilon = .26$ into Table G. By explicit calculation, $\mathcal{P}_{10}(\mathsf{MAJ}_1; .26) \leq .0493$, $p_{10}(\mathsf{MAJ}_5; .26) \leq .0488$, $p_{10}(\mathsf{MAJ}_3; .26) \geq .0496$. □

Proposition G.3 *There exist n, k, and $f \in \mathcal{B}_k$ such that the probability all parties agree on 1 differs from the probability all parties agree on 0.*

Proof: With $n = 5$, $k = 3$, and f the left-monotone function with minterms 10010 and 01101, explicit calculation gives $\frac{1}{2} - \frac{39}{16}\epsilon + 9\epsilon^2 - \frac{459}{16}\epsilon^3 + \frac{297}{4}\epsilon^4 - \frac{2331}{16}\epsilon^5 + \frac{3465}{16}\epsilon^6 - 234\epsilon^7 + 171\epsilon^8 - 75\epsilon^9 + 15\epsilon^{10}$ and $\frac{1}{2} - \frac{39}{16}\epsilon + \frac{69}{8}\epsilon^2 - \frac{381}{16}\epsilon^3 + \frac{93}{2}\epsilon^4 - \frac{885}{16}\epsilon^5 + \frac{519}{16}\epsilon^6 + 6\epsilon^7 - 24\epsilon^8 + 15\epsilon^9 - 3\epsilon^{10}$ for the probabilities of agreement on 1 and 0, respectively. □

f	$\mathcal{P}_k(f;\epsilon)$
MAJ_1	$\epsilon^k + (1-\epsilon)^k$
T_1	$1/16\,(-6\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5 + 4\,\epsilon^2)^k + 1/16\,(1 + 6\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5 - 4\,\epsilon^2)^k + 1/16\,(4\,\epsilon - 10\,\epsilon^2 + 10\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5)^k + 1/16\,(1 - 4\,\epsilon + 10\,\epsilon^2 - 10\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5)^k + 1/4\,(\epsilon - \epsilon^2 + 4\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5)^k + 1/4\,(1 - \epsilon + \epsilon^2 - 4\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5)^k + 1/4\,(1 - 2\,\epsilon + 4\,\epsilon^2 - 6\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5)^k + 1/4\,(2\,\epsilon - 4\,\epsilon^2 + 6\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5)^k + 3/8\,(\epsilon + \epsilon^2 - 4\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5)^k + 3/8\,(1 - \epsilon - \epsilon^2 + 4\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5)^k$
T_2	$1/8\,(-2\,\epsilon^3 + 3\,\epsilon^2)^k + 1/8\,(1 + 2\,\epsilon^3 - 3\,\epsilon^2)^k + 1/8\,(3\,\epsilon - 6\,\epsilon^2 + 4\,\epsilon^3)^k + 1/8\,(1 - 3\,\epsilon + 6\,\epsilon^2 - 4\,\epsilon^3)^k + 3/8\,\epsilon^k + 3/8\,(1-\epsilon)^k + 3/8\,(1 - 2\,\epsilon + 3\,\epsilon^2 - 2\,\epsilon^3)^k + 3/8\,(2\,\epsilon - 3\,\epsilon^2 + 2\,\epsilon^3)^k$
MAJ_3	$1/4\,(-2\,\epsilon^3 + 3\,\epsilon^2)^k + 1/4\,(1 + 2\,\epsilon^3 - 3\,\epsilon^2)^k + 3/4\,(2\,\epsilon - 3\,\epsilon^2 + 2\,\epsilon^3)^k + 3/4\,(1 - 2\,\epsilon + 3\,\epsilon^2 - 2\,\epsilon^3)^k$
T_3	$1/8\,(-6\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5 + 4\,\epsilon^2)^k + 1/8\,(1 + 6\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5 - 4\,\epsilon^2)^k + 1/16\,(\epsilon - \epsilon^2 + 4\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5)^k + 1/16\,(1 - \epsilon + \epsilon^2 - 4\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5)^k + 1/4\,(1 - 2\,\epsilon + 4\,\epsilon^2 - 6\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5)^k + 1/8\,(1 - \epsilon - \epsilon^2 + 4\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5)^k + 1/8\,(\epsilon + \epsilon^2 - 4\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5)^k + 3/16\,(8\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5 + 3\,\epsilon - 7\,\epsilon^2)^k + 3/16\,(1 - 8\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5 - 3\,\epsilon + 7\,\epsilon^2)^k + 3/16\,(2\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5 + 2\,\epsilon^2 + 1 - 2\,\epsilon)^k + 3/16\,(-2\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5 - 2\,\epsilon^2 + 2\,\epsilon)^k + 1/16\,(2\,\epsilon^3 - 5\,\epsilon^4 + 2\,\epsilon^5 + 2\,\epsilon^2)^k + 1/16\,(1 - 2\,\epsilon^3 + 5\,\epsilon^4 - 2\,\epsilon^5 - 2\,\epsilon^2)^k$
T_4	$1/16\,(\epsilon^2 + 6\,\epsilon^3 - 10\,\epsilon^4 + 4\,\epsilon^5)^k + 1/16\,(1 - \epsilon^2 - 6\,\epsilon^3 + 10\,\epsilon^4 - 4\,\epsilon^5)^k + 1/8\,(\epsilon + 2\,\epsilon^2 - 8\,\epsilon^3 + 10\,\epsilon^4 - 4\,\epsilon^5)^k + 1/8\,(1 - \epsilon - 2\,\epsilon^2 + 8\,\epsilon^3 - 10\,\epsilon^4 + 4\,\epsilon^5)^k + 1/16\,(1 - 2\,\epsilon + \epsilon^2 + 6\,\epsilon^3 - 10\,\epsilon^4 + 4\,\epsilon^5)^k + 1/16\,(2\,\epsilon - \epsilon^2 - 6\,\epsilon^3 + 10\,\epsilon^4 - 4\,\epsilon^5)^k + 3/16\,(5\,\epsilon^2 - 10\,\epsilon^3 + 10\,\epsilon^4 - 4\,\epsilon^5)^k + 3/16\,(1 - 5\,\epsilon^2 + 10\,\epsilon^3 - 10\,\epsilon^4 + 4\,\epsilon^5)^k + 3/8\,(3\,\epsilon - 8\,\epsilon^2 + 12\,\epsilon^3 - 10\,\epsilon^4 + 4\,\epsilon^5)^k + 3/8\,(1 - 3\,\epsilon + 8\,\epsilon^2 - 12\,\epsilon^3 + 10\,\epsilon^4 - 4\,\epsilon^5)^k + 3/16\,(1 - 2\,\epsilon + 5\,\epsilon^2 - 10\,\epsilon^3 + 10\,\epsilon^4 - 4\,\epsilon^5)^k + 3/16\,(2\,\epsilon - 5\,\epsilon^2 + 10\,\epsilon^3 - 10\,\epsilon^4 + 4\,\epsilon^5)^k$
MAJ_5	$1/16\,(10\,\epsilon^3 - 15\,\epsilon^4 + 6\,\epsilon^5)^k + 1/16\,(1 - 10\,\epsilon^3 + 15\,\epsilon^4 - 6\,\epsilon^5)^k + 5/16\,(6\,\epsilon^2 - 14\,\epsilon^3 + 15\,\epsilon^4 - 6\,\epsilon^5)^k + 5/16\,(1 - 6\,\epsilon^2 + 14\,\epsilon^3 - 15\,\epsilon^4 + 6\,\epsilon^5)^k + 5/8\,(3\,\epsilon - 9\,\epsilon^2 + 16\,\epsilon^3 - 15\,\epsilon^4 + 6\,\epsilon^5)^k + 5/8\,(1 - 3\,\epsilon + 9\,\epsilon^2 - 16\,\epsilon^3 + 15\,\epsilon^4 - 6\,\epsilon^5)^k$

Table G