The Round Complexity of Two-Party Random Selection*

Saurabh Sanghvi[†] Salil Vadhan[‡]
Division of Engineering & Applied Sciences
Harvard University, Cambridge, MA
ssanghvi@post.harvard.edu,salil@eecs.harvard.edu

October 2, 2005

Abstract

We study the round complexity of two-party protocols for generating a random n-bit string such that the output is guaranteed to have bounded bias (according to some measure) even if one of the two parties deviates from the protocol (even using unlimited computational resources). Specifically, we require that the output's statistical difference from the uniform distribution on $\{0,1\}^n$ is bounded by a constant less than 1.

We present a protocol for the above problem that has $2\log^* n + O(1)$ rounds, improving a previous $2n$-round protocol of Goldreich, Goldwasser, and Linial (FOCS '91). Like the GGL protocol, our protocol actually provides a stronger guarantee, ensuring that the output lands in any set $T \subseteq \{0,1\}^n$ of density μ with probability at most $O(\sqrt{\mu + \delta})$, where δ is an arbitarily small constant.

We then prove a matching lower bound, showing that any protocol guaranteeing bounded statistical difference requires at least $\log^* n - \log^* \log^* n - O(1)$ rounds. As far as we know, this is the first nontrivial lower bound on the round complexity of random selection protocols (of any type) that does not impose additional constraints (e.g. on communication or "simulatability").

We also prove several results for the case when the output's bias is measured by the maximum *multiplicative* factor by which a party can increase the probability of a set $T \subseteq \{0,1\}^n$.

Keywords: cryptography, distributed computing, coin-flipping

*Preliminary versions of this work have appeared in the first author's undergraduate thesis [San04], and in the conference paper [SV05].

[†]Supported in part by a grant from the Harvard College Research Program and NSF grant CCR-0133096.

[‡]Work done in part while a Fellow at the Radcliffe Institute for Advanced Study at Harvard University. Also supported by NSF grants CNS-0430336, ONR grant N00014-04-1-0478, and a Sloan Research Fellowship.

1 Introduction

One of the most basic protocol problems in cryptography and distributed computing is that of *random selection*, in which several mutually distrusting parties aim to generate an n-bit random string jointly. The goal is to design a protocol so that even if a party cheats, the outcome will still not be too "biased". (There are many different choices for how to measure the "bias" of the output; the one we use will be specified later.) Random selection protocols can dramatically simplify the design of protocols for other tasks via the following common methodology: first design a protocol in a model where truly random strings are provided by a trusted third party (generally a much easier task), and then use the random selection protocol to eliminate the trusted third party. For this reason, there is a wide literature on random selection protocols, both in the computational setting, where cheating parties are restricted to polynomial time (starting with Blum's "coin flipping by telephone" [Blu82]), and the information-theoretic setting, where security is provided even against computationally unbounded adversaries.

We will focus on *two-party* protocols in the *information-theoretic setting* (also known as the "full information model"). In addition to its stronger security guarantees, the information-theoretic setting has the advantage that protocols typically do not require complexity-theoretic assumptions (such as the existence of one-way functions). Various such random selection protocols have been used to construct perfectly hiding bit-commitment schemes [NOVY98], to convert honest-verifier zero-knowledge proofs into general zero-knowledge proofs [Dam94, DGW94, GSV98], to construct oblivious transfer protocols in the bounded storage model [CCM98, DHRS04], and to perform general fault-tolerant computation [GGL98]. There has also been substantial work in the k-party case for $k \geq 3$, where the goal is to tolerate coalitions of a minority of cheating players. This body of work includes the well-studied "collective coin-flipping" problem e.g., [BL89, Sak89, AN90, ORV94, RZ98, Fei99] (closely related to the "leader election" problem), and again the use of random selection as a tool for general fault-tolerant computation [GGL98].

In most of the lines of work mentioned above (computational and information-theoretic, two-party and k-party), the *round complexity* has been a major parameter of interest. For some forms of random selection and their applications, constant-round protocols have been found (e.g. [DGW94, GSV98] improving [Dam94], [DHRS04] improving [CCM98], and [Lin01, KO04] improving [Blu82, Yao86]), but for others the best known protocols have a nonconstant number of rounds, e.g. [NOVY98, GGL98, RZ98]. Lower bounds on round complexity, however, have proven much more difficult to obtain, and we only know of examples that impose additional constraints on the protocol (beyond the basic security guarantee of bounded bias). For example, in the computational setting, it has been recently shown that 5 rounds are necessary and sufficient for random selection protocols satisfying a certain "black-box simulation" condition [KO04]. In the information-theoretic setting, a long line of work on the collective coin-flipping problem has culminated in the $(\log^* n + O(1))$-round protocol[1] of Russell and Zuckerman [RZ98] (see also Feige [Fei99]), but the only known lower

[1]As in other work [RZ98], for the purposes of this paper we will define $\log_b^{(k)} n$ to be k base b iterated logarithms of n:

$$\log_b^{(k)} n = \begin{cases} 1 & : \text{ if } \log_b^{(k-1)} n < b \\ \log_b\left(\log_b^{(k-1)} n\right) & : \text{ otherwise} \end{cases}$$

with $\log_b^{(0)} n = n$. Moreover, for $n \geq 1$, we define $\log_b^* n$ to be the least natural number k such that bound (of $\Omega(\log^* n)$ rounds), due to Russell, Saks, and Zuckerman [RSZ99], is restricted to protocols where each party can only communicate a small number of bits per round. Without this restriction, it is not even known how to prove that 1 round is impossible.

The problem and main results. As mentioned above, previous works on random selection have considered a number of different measures of the bias of the output, typically motivated by particular applications. Here we focus on what we consider to be the most natural measure — the statistical difference (i.e., variation distance) of the output from the uniform distribution.[2] Specifically, we seek a two-party protocol (A, B) that produces an output in $\{0,1\}^n$, such that even if one party deviates arbitrarily from the specified protocol, the statistical difference of the output from uniform is bounded by a constant less than 1. Equivalently, we want to satisfy the following criterion.

Statistical Criterion: There are fixed constants $\mu > 0$ and $\epsilon > 0$ such that for every n and every subset $T \subseteq \{0,1\}^n$ of density at most μ, the probability that the output lands in T is at most $1 - \epsilon$, even if one party deviates arbitrarily from the specified protocol.

In addition to being a natural choice, this criterion is closely related to others considered in the literature. In particular, the standard criterion for the "collective coin-flipping" problem is that the output bit $B \in \{0,1\}$ satisfies $\max\{\Pr[B=0], \Pr[B=1]\} < p$, where p is a constant less than 1; this is equivalent to B's statistical difference from uniform being bounded away from 1. (Here we see that the problem we consider is in some sense "dual" to collective coin-flipping — we restrict to two players but the output comes from a large set, whereas in collective coin-flipping there are many players but the output has only two possibilities.)

Of course, the first question is whether or not the Statistical Criterion can be met at all, regardless of round complexity. Indeed, being able to tolerate computationally unbounded cheating strategies is a strong requirement. In fact, when $n = 1$ (i.e. the output is a single bit), it turns out that one of the two parties can always force the outcome to be constant. This implies that the Statistical Criterion is impossible to meet for $\mu = 1/2$. Surprisingly, the criterion is achievable, however, for some smaller constant $\mu > 0$. This is implied by the following result of Goldreich, Goldwasser, and Linial [GGL98].

Theorem 1.1 ([GGL98]) *For every n, there is a two-party protocol producing output in $\{0,1\}^n$ such that, as long as one party plays honestly, the probability that the output lands in any set $T \subseteq \{0,1\}^n$ of density μ is at most $p = O(\sqrt{\mu})$. The protocol has $2n$ rounds.*

Notice that for sufficiently small μ, the probability p is indeed a constant less than 1. This implies that the Statistical Criterion is achievable with a *linear* number of rounds. Our goal is to determine the minimal round complexity of this problem.

First, we give a protocol achieving the Statistical Criterion with substantially fewer rounds than the above.

Theorem 1.2 *For every constant $\delta > 0$, there is a two-party protocol producing output in $\{0,1\}^n$ with $2\log^* n + O(1)$ rounds such that, as long as one party plays honestly, the probability that the output lands in any set T of density μ is at most $p = O(\sqrt{\mu + \delta})$.*

$\log_b^{(k)} n \leq 1$.

[2] The *statistical difference* between two random variables X and Y taking values in a universe $\mathcal{U}$ is defined to be $\Delta(X,Y) = \max_{S \subseteq \mathcal{U}} |\Pr[X \in \mathcal{U}] - \Pr[Y \in \mathcal{U}]|$.

Our protocol is inspired by the $\log^* n$-round protocols for leader election [RZ98, Fei99] and Lautemann's proof that **BPP** is contained in the polynomial hierarchy [Lau83]. Specifically, we exhibit a 2-round protocol that reduces the universe of size $N = 2^n$ to a universe of size $\text{polylog}(N)$, while approximately preserving the density of the set T with high probability. Repeating this protocol $\log^* n$ times reduces the universe size to a constant, after which point we apply the GGL protocol.

Second, we prove a lower bound that matches the above up to a factor of $2 + o(1)$.

Theorem 1.3 *Any two-party protocol producing output in* $\{0,1\}^n$ *that satisfies the Statistical Criterion must have at least* $\log^* n - \log^* \log^* n - O(1)$ *rounds.*

Our proof of this theorem is a technically intricate induction on the game tree of the protocol. Roughly speaking, we associate to each node z of the game tree, a collection $\mathcal{S}_z$ of very small sets such that if the protocol is started at z and R is a random subset of the universe of density $o(1)$, one of the players X can force the outcome of the protocol to land in $R \cup S$ with probability $1-o(1)$, for any $S \in \mathcal{S}_z$. The challenge is to keep the size of the sets in the collections $\mathcal{S}_z$ small as we induct up the game tree (so that they remain of density $o(1)$ when z is the root, which yields the desired lower bound). In particular, a node can have an arbitrary number of children, so we cannot afford to take unions of sets S occurring across all children. The key idea that allows us to keep the sets small is the following. We consider two cases: If we have a collection of sets that contains a large disjoint subcollection, then the random set R will contain one of the sets with high probability and so we do not need to carry the set through the recursion. On the other hand, if the collection of sets has no large disjoint subcollection, then we show how we can use this fact to construct a successful strategy for the *other* player (based on how we inductively construct the collections $\mathcal{S}_z$).

We stress that our lower bound does not impose any additional constraint on the protocol, such as the number of bits sent per round. Thus, we hope that our techniques can help in establishing unrestricted lower bounds on round complexity for other problems, in particular for the collective coin-flipping (and leader election) problem.

Results on multiplicative guarantees. A different measure of the quality of random selection protocol is a *multiplicative guarantee*, whereby we require that, even if one player cheats, the probability that the outcome lands in any set T of density μ is at most $\rho \cdot \mu$, for some parameter $\rho \geq 1$. The goal, naturally, is for ρ to be as small as possible (ideally a constant independent of n). Previous protocols, e.g. [DGW94], have given a multiplicative guarantee to one player while the other has a statistical guarantee (i.e. a bound on the output's statistical difference from uniform if the other cheats). Our observations and results on multiplicative guarantees are the following:

- If both parties have multiplicative guarantees ρ_A and ρ_B, then an argument of [GGL98] implies $\rho_A \cdot \rho_B \geq 2^n$, regardless of the number of rounds.
- There is a simple two-round protocol achieving $\rho_A \cdot \rho_B \leq 2^n$, for any desired ρ_A.
- If one party has a multiplicative guarantee ρ and the other has a statistical guarantee ε, then $\varepsilon \geq 1/\rho - 1/2^n$. This explains inverse relationships in existing protocols of

[DGW94] (where $\varepsilon = 1/\mathrm{poly}(n)$ and $\rho = \mathrm{poly}(n)$) and [GSV98] (where $\varepsilon = \mathrm{poly}(n) \cdot 2^{-k}$ and $\rho = 2^k$ for any k).[3]

- There is a protocol with $2\log^* n + O(1)$ rounds that provides a constant statistical guarantee to one player and a (arbitrarily small) constant multiplicative guarantee to the other. Theorem 1.3 implies that this round complexity is tight up to a constant factor, because a constant multiplicative guarantee implies a constant statistical guarantee.

2 Defining Random Selection Protocols

We can formally characterize a random selection protocol as follows:

Definition 2.1 *A* random selection protocol $\Pi = (A, B, f)$ *over a universe* $\mathcal{U}$[4] *consists of a pair of functions* A *and* B *and a function* f *such that:*

- *Both* A *(Alice) and* B *(Bob) alternately output strings ("messages")* m_i *of arbitrary length that are a function of the conversation thus far and their sequences of random coin tosses* r_A *and* r_B*, respectively. That is,* $m_1 = A(r_A)$*,* $m_2 = B(r_B, m_1)$*,* $m_3 = A(r_A, m_1 m_2)$*, etc.*

- *The* conversation *between Alice and Bob is the transcript* $(A, B) = m_1 m_2 \ldots m_r$*, where* r *is a parameter defining the number of messages*[5] *of the protocol.*

- *The* output *of the protocol is* $f(m_1 m_2 \ldots m_r)$*, which is some element of* $\mathcal{U}$*.*

We are interested in the behavior of the protocol when one of these programs is replaced with an arbitrary "cheating" program A^* or B^*, which may send its messages as an arbitrary function of the conversation and input length.

Although the formulation we have provided assumes a protocol operates over a single fixed universe, in general we will be interested in studying asymptotic behavior of protocols as the universe size increases. Thus, we define a *random selection protocol ensemble* to be a sequence $(\Pi^{(1)}, \Pi^{(2)}, \ldots)$ where each $\Pi^{(N)}$ is a protocol over $\mathcal{U} = \{1, \ldots, N\}$.

¿From now on, we blur the distinction between random selection protocols over a fixed universe and random selection protocol ensembles. Results depending on asymptotics will hold for random selection protocol ensembles, and other results will hold for any fixed-universe random selection protocol–in particular, every protocol in the ensemble.

Two desirable properties of random selection protocols are (a) the output is uniformly distributed in $\mathcal{U}$ assuming honest players, and (b) in a protocol ensemble, honest strategies can be computed in time polynomial in the output length, $\log N$. Our protocols will have these properties, but our lower bounds will apply even to protocols without them.

We now introduce a formalism that will be invaluable in the proofs of this paper.

Definition 2.2 *Given a protocol* Π *over universe* $\mathcal{U}$*, define the* game tree T *as follows:*

[3] Actually, the protocol of [GSV98] does not provide a multiplicative guarantee of 2^k, but rather ensures that the probability that the output lands in any set T of density μ is at most $2^k \cdot \mu + o(1)$. Our lower bound also applies to this more general type of guarantee.

[4] Although we introduced the problem for a universe $\{0,1\}^n$, for the rest of the paper we assume we have an arbitrary universe $\mathcal{U}$.

[5] We use "messages" interchangeably with "rounds" and "turns."

- *A set of nodes V, each representing a partial transcript of messages, $(m_1, \ldots, m_i)$.*
- *A set of edges E, defined by $(u, v) \in E$ if and only if $u = (m_1, \ldots, m_i)$ and (abusing notation) $v = (u, m_{i+1})$, for any message m_{i+1}. That is, u connects to v if v is a potential protocol state after one message from u.*
- *For each node z, a distribution $\mathcal{D}_z$ over the children z_i whereby A or B chooses the next message.*
- *For every leaf $z = (m_1, \ldots, m_r)$, a label equal to $f((m_1, \ldots, m_r))$, the output of the protocol ending at node z.*

One can verify that this formalism produces an equivalent specification as Definition 2.1 of a random selection protocol.

Just as any node of a tree is the root of another, any node of a protocol's game tree induces its own random selection protocol starting from that state. We simply fix the messages leading to that node, and have the players choose the remaining messages as in the original protocol. This observation is one of the main reasons that the abstraction of a random selection protocol as a tree will prove useful.

Evaluating a Random Selection Protocol. We evaluate random selection protocols with metrics measuring how "close" the output is to the uniform distribution on $\mathcal{U}$. The primary metric we use is the following.

Definition 2.3 *The* statistical difference *(a.k.a. variation distance) of a distribution X over universe $\mathcal{U}$ from uniform is defined to be*

$$\max_T \left| \Pr_{x \leftarrow X}[x \in T] - \mu(T) \right| = \max_T \left(\Pr_{x \leftarrow X}[x \in T] - \mu(T) \right)$$

where $T \subseteq \mathcal{U}$ and $\mu(T)$ is the density of T in $\mathcal{U}$ (i.e., $|T|/|\mathcal{U}|$).

Statistical difference finds the subset of the universe that is hit with probability most different from uniform. It can be verified that this distance is in the interval $[0, 1-1/N]$, where N is the size of the universe $\mathcal{U}$. A statistical difference of 0 implies that X is uniform, and $1 - 1/N$ implies X is concentrated on a single point.

We will want to avoid distributions X whose statistical difference from uniform is very close to 1. The following lemma demonstrates this (undesirable) property is equivalent to X landing in a small set with high probability.

Lemma 2.4 *If X has statistical difference at least $1-\epsilon$ from uniform, then there exists a set T such that $\mu(T) \leq \epsilon$ and $\Pr_{x \leftarrow X}[x \in T] \geq 1 - \epsilon$. Conversely, if there exists such a set T, then X has statistical difference at least $1 - 2\epsilon$ from uniform.*

Proof: If X has statistical difference at least $1-\epsilon$ from uniform, then there exists a set T such that $\Pr[x \in T] - \mu(T) \geq 1 - \epsilon$. Since $\Pr[x \in T] \leq 1$, we can conclude $\mu(T) \leq \epsilon$, and since $\mu(T) \geq 0$, we can conclude $\Pr[x \in T] \geq 1 - \epsilon$.

The reverse direction follows directly from the definition of statistical difference. ■

We also consider multiplicative difference:

Definition 2.5 *The* multiplicative difference *of a distribution X is*

$$\max_T \Pr_{x \leftarrow X}[x \in T]/\mu(T)$$

where T ranges over all subsets of $\mathcal{U}$.

We defer all of our results regarding multiplicative difference to Section 5. Given these metrics, we can define:

Definition 2.6 *The statistical guarantee for Alice playing honest strategy A in a protocol Π (denoted ϵ_A) is the maximum over all B^* of the statistical difference between the distribution of $f((A, B^*))$ and the uniform distribution over $\mathcal{U}$. The guarantees for Bob are defined analogously.*

Intuitively, the guarantee of a protocol for a player bounds the damage that the opponent can effect on the distribution by deviating from the protocol. Unfortunately, the terminology here is a bit counterintuitive—the lower the number, the better the guarantee. We will try to avoid confusion by saying a guarantee is "at best x", rather than "at least x."

Armed with this notion of a guarantee, we can state the following important equivalence, following directly from Lemma 2.4:

Proposition 2.7 *The Statistical Criterion is equivalent to both of the statistical guarantees of a protocol being bounded away from 1.*

3 The Iterated Random Shift Protocol

In this section we describe the main protocol of this paper, the Iterated Random Shift Protocol, and prove its main properties. That is, we show that for any constant δ, Iterated Random Shift is a $2\log^* N + O(1)$-round protocol where the probability the output falls in a set of density μ is at most $O(\sqrt{\mu + \delta})$. It follows that the protocol satisfies the Statistical Criterion given above.

Our protocol is inspired by the $\log^* n$-round protocols for leader election [RZ98, Fei99] and Lautemann's proof that **BPP** is contained in the polynomial hierarchy [Lau83]. It is built by iteration of the following 2-round protocol, which we will call the Random Shift Protocol:

The Random Shift Protocol $\Pi(A, B)$: Given a universe $\mathcal{U}$ of size N and $m \in \mathbb{N}$,

1. Alice randomly selects a sequence of strings $a_1, \ldots, a_m \in \mathcal{U}$.
2. Bob randomly selects a sequence of strings $b_1, \ldots, b_m \in \mathcal{U}$.
3. Output the sequence $(a_i + b_j : 1 \leq i \leq m)$, where $+$ is a group operation over $\mathcal{U}$.

Note that the Random Shift Protocol is not, strictly speaking, a random selection protocol over $\mathcal{U}$: its output is a *sequence* of strings from the universe. In using it, we will typically choose the parameter m so that the number of output strings, m^2, is much smaller than N (e.g. $m = \text{polylog}(N)$), and recursively use our random selection protocol to select one of the m^2 output strings. To show that this approach yields a protocol with bounded statistical guarantees, we argue that even if one of the players cheats, any subset T of the universe is unlikely to appear in much more than a $\mu(T)$ fraction of the outputs of the Random Shift Protocol. This is formalized by the following lemma.

Lemma 3.1 *Let T be an arbitrary subset of $\mathcal{U}$. Let $\mu(T) = |T|/N$, and let $\mu'(T)$ denote the density of T in the sequence output by the Random Shift Protocol: $\mu'(T) = \#\{(i,j) : a_i + b_j \in T\}/m^2$. Then as long as one player plays honestly (i.e., chooses strings uniformly at random) and $m \geq (1/2\epsilon^2) \cdot \log(N/\delta)$, we have*

$$\Pr[\mu'(T) \geq \mu(T) + \epsilon] \leq \delta.$$

That is, when one player is honest, the sequence $a_i + b_j$ will be sufficiently random so that it is very unlikely that the density of T in the output sequence will increase substantially.

Proof: Suppose Alice plays honestly and chooses her strings $a_1, \ldots a_m$ at random from $\mathcal{U}$. The lemma certainly holds *a fortiori* for an honest Bob, as a cheating Bob can see what strings Alice has selected.

Fix an arbitrary string $b \in \mathcal{U}$. Define random variables

$$X_i^{(b)} = \begin{cases} 1 & : \text{ if } (a_i + b) \in T \\ 0 & : \text{ otherwise} \end{cases}$$

Then define $X^{(b)} = (1/m) \sum_{i=1}^{m} X_i^{(b)}$. Notice that $E[X_i^{(b)}] = \mu(T)$ and that these random variables are independent.

By a Chernoff bound, we may conclude the following for any ϵ:

$$\Pr[X^{(b)} \geq \mu(T) + \epsilon] \leq e^{-2\epsilon^2 m} \leq \frac{\delta}{N}$$

Using a union bound, we conclude:

$$\Pr[\exists b \in \mathcal{U} \text{ such that } X^{(b)} \geq \mu(T) + \epsilon] \leq N \cdot \frac{\delta}{N} = \delta$$

But if for all b, we have $X^{(b)} < \mu(T) + \epsilon$, then no matter what strings $b_1, \ldots, b_m$ Bob chooses, we have

$$\mu'(T) = (1/m) \sum_{j=1}^{m} X^{(b_j)} < \mu(T) + \epsilon$$

It follows that $\Pr[\mu'(T) \geq \mu(T) + \epsilon] \leq \delta$ as desired. ■

Remark 3.2 We note that the number of strings sent by Bob need only be $(1/2\epsilon^2) \cdot \log(1/\delta)$ (i.e. the $\log N$ factor can be eliminated), since there is no need to do a union bound as in the above proof when proving Bob's guarantee. However, the symmetry of the protocol as presented above has the advantage that it can actually be implemented in *one* round in a model of simultaneous communication (where honest parties can send messages at the same time, but a cheating party may wait to see the other party's message before sending its own message), as is typically used in many-party protocols (e.g. leader election and collective coin-flipping). This reduces the round complexity of our Iterated Random Shift Protocol below to $\log^* N + O(1)$ in the simultaneous communication model. It is interesting to know whether our lower bound of $\log^* N - \log^* \log^* N - O(1)$ rounds (in Section 4.2) can be extended to the simultaneous communication model (without paying the factor of 2 in the trivial reduction to our non-simultaneous model), since we would then have a lower bound in that model that is tight up to a factor of $1 + o(1)$.

Figure 1: **The Iterated Random Shift Protocol**

Let $M \in \mathbb{N}$ be a "cutoff" parameter that is a power of 2, let $\mathcal{U}$ be a universe of size N. We assume that $N \geq M^2$, else we use the protocol below to select from $\mathcal{U}' = \mathcal{U} \times [M^2]$ and take the first component of the output.

Recursive Protocol: If $N > M^2$, let $m = \max\{M, \lceil \log^3 N \rceil\}$, and let $+$ be a fixed group operation over universe $\mathcal{U}$ (e.g., addition mod N).

Alice **Bob**

$$\xrightarrow{\quad a_1,\ \ldots,\ a_m \in_R \mathcal{U} \quad}$$

$$\xleftarrow{\quad b_1,\ \ldots,\ b_m \in_R \mathcal{U} \quad}$$

$$\boxed{\begin{array}{c}\text{Recursive execution on}\\ \mathcal{U} = [m] \times [m]\end{array}}$$

$\Downarrow$ Output (i, j)

Output $x = a_i + b_j$

If $N = M^2$,

Alice **Bob**

$$\xrightarrow{\quad \text{Choose } x \text{ using GGL [GGL98]} \quad}$$

We now describe our Iterated Random Shift Protocol satisfying Theorem 1.2, which consists of recursively applying the Random Shift Protocol until the universe size is small (say, less than a fixed constant), after which we apply the Goldreich–Goldwasser–Linial [GGL98] Protocol. We define the Iterated Random Shift Protocol in Figure 1.

Theorem 3.3 *If $M \geq 1/\epsilon^3$, then for any set $T \subseteq \mathcal{U}$, the probability that the output of the Iterated Random Shift Protocol lands in T is $O(\sqrt{\mu + \epsilon})$, assuming at least one player plays honestly.*

Corollary 3.4 *For a sufficiently large constant M, the Iterated Random Shift Protocol satisfies the Statistical Criterion. Equivalently, there exists a constant $\epsilon > 0$ such that such an Iterated Random Shift Protocol achieves* $\max\{\epsilon_A, \epsilon_B\} \leq 1 - \epsilon$.

Observe that Theorem 3.3 is much stronger than what we need to show Corollary 3.4. Using Theorem 3.3, we know that when one player is honest, for any "small" set T, the probability the output falls in T is close to zero. The Statistical Criterion requires only that this probability is not arbitrarily close to 1.

Proof of Theorem 3.3: The key idea is that in the ith application of the Random Shift Protocol, we can bound the increase in density of any particular set T by at most some small ϵ_i (with high probability) and these ϵ_i's can be chosen so that $\sum_i \epsilon_i \leq \epsilon$. The Iterated Random Shift Protocol concludes by applying the GGL Protocol to this small universe, and then Theorem 1.1 gives us the result.

We first note that the modification of the protocol in case $N < M^2$, of selecting from $\mathcal{U} \times [M^2]$ and taking the first component, does not affect the property claimed in the theorem (because the density of $T \times [M^2]$ in $\mathcal{U} \times [M^2]$ equals the density of T in $\mathcal{U}$). Thus we assume that $N \geq M^2$, and let $N_0, N_1, \ldots, N_{k^*}$ be the universe sizes in an execution of the Iterated Random Shift Protocol, where k^* is the first value of k such that $N_k = M^2$. That is,

$$\begin{aligned} N_0 &= N \\ N_k &= m_k^2, \text{ for } m_k = \max\{M, \lceil \log^3 N_{k-1} \rceil\} \end{aligned}$$

Note that for sufficiently large M, the sequence of N_i's is strictly decreasing and there exists a finite k^* such that $N_{k^*} = M^2$.

Now, given a subset $T \subseteq \mathcal{U}$, we track how T evolves through an execution of the Iterated Random Shift protocol by the following for $k = 0, \ldots, k^*$:

$$\begin{array}{ll} \mathcal{U}_0 = \mathcal{U} & \mathcal{U}_k = [m_k] \times [m_k] \\ T_0 = T & T_k = \{(i,j) \in \mathcal{U}_k : (a_i + b_j) \in T_{k-1}, 1 \leq i, j \leq m_k\} \\ \mu(T_k) = |T_k|/|\mathcal{U}_k| & \end{array}$$

where in the definition of T_k, (a_i) and (b_j) are the sequences of elements of $\mathcal{U}_{k-1}$ chosen by Alice and Bob in the kth application of the Random Shift Protocol, and $+$ is the group operation over $\mathcal{U}_{k-1}$ used in the protocol.

Intuitively, $\mathcal{U}_k$ is the remaining universe (of size N_k) after k iterations and T_k represents the portion of the remaining universe such that choosing $(i,j) \in T_k$ will lead to an element of T being the output of the whole protocol. We call $\mu(T_k)$ the "effective density" of T in the kth iteration.

Claim 3.5 *There is a constant c such that for all N and M, we have*

$$\Pr\left[\mu(T_{k^*}) \geq \mu(T) + c \cdot 2^{-M^{1/3}}\right] \leq c \cdot M^{-1/3},$$

provided at least one party plays honestly.

Proof of Claim: Recall that in the k'th iteration, we are applying the Random Shift Protocol with parameter $m = m_k = \max\{M, \lceil \log^3 N_{k-1} \rceil\}$. Define $\delta_k = 2^{-m_k^{1/3}}$, and $\epsilon_k = 1/m_k^{1/3}$. Notice that $m_k \geq (1/2\epsilon_k^2) \cdot \log(N_{k-1}/\delta_k)$.

Inducting on Lemma 3.1 and using a union bound, we have that for any k,

$$\Pr\left[\mu(T_k) \geq \mu(T) + \sum_{i=1}^{k} \epsilon_i\right] \leq \sum_{i=1}^{k} \delta_i$$

Since the N_k's are decreasing extremely fast, we have

$$\begin{aligned} \sum_{i=0}^{k^*} \epsilon_i &= O(\epsilon_{k^*}) \\ &= O(1/m_{k^*}^{1/3})) \\ &= O(1/M^{1/3}). \end{aligned}$$

Similarly,

$$\sum_{i=1}^{k^*} \delta_i = O(2^{-m_{k^*}^{1/3}}) = O(2^{-M^{1/3}}).$$

This completes the proof. □

Applying Claim 3.5 and using Theorem 1.1, we deduce that the probability that the output lands in T is at most

$$O\left(\sqrt{\mu(T) + \frac{c}{M^{1/3}}}\right) + \frac{c}{2^{M^{1/3}}} = O\left(\sqrt{\mu(T) + \frac{1}{M^{1/3}}}\right) = O\left(\sqrt{\mu(T) + \epsilon}\right),$$

provided $M \geq 1/\epsilon^3$. Theorem 3.3 is proven. ■

It finally remains to verify the round complexity of the Iterated Random Shift Protocol:

Proposition 3.6 *For all sufficiently large M and all N, the Iterated Random Shift Protocol over a universe $\mathcal{U}$ of size N with parameter M takes $2\log^* N + O(\log M)$ rounds.*

Proof: Each application of the Random Shift Protocol (except for the last) reduces the universe size from N to $\lceil \log^3 N \rceil^2 < \log^7 N$ for sufficiently large N, and takes two rounds. A lemma proven in [RZ98] states that if $f(n) \leq \log^a n$ for some constant a, then $f^{(\log^* n)}(n) \leq b$ for some constant b, where $f^{(k)}$ represents $k = \log^* n$ repeated applications of f. This implies that, if M is sufficiently large and the initial universe size is $N \geq M^2$, the Random Shift Protocol is applied at most $\log^* N$ times. (If $N < M^2$, then we apply the Random Shift Protocol at most $\log^*(NM^2) = \log^* N + O(\log M)$ times.) By Theorem 1.1, the GGL protocol on a universe of size at most M^2 takes at most $4 \log M$ rounds. ■

Thus, taking M to be a sufficiently large constant, we obtain a protocol with $2\log^* N + O(1)$ rounds satisfying the Statistical Criterion, thereby proving Theorem 1.2. More generally, we obtain a protocol of $2\log^* N + O(\log(1/\epsilon))$ rounds such that the output lands in a sets of density μ with probability at most $O(\sqrt{\mu + \epsilon})$. Note that we can take ϵ to be a slowly vanishing function of N and still have $O(\log^* N)$ rounds.

In the next section, we will prove that the Iterated Random Shift Protocol has optimal round complexity, up to a factor of $2 + o(1)$, among protocols achieving the Statistical Criterion.

4 Lower Bounds on Statistical Guarantees

4.1 Tradeoffs between Statistical Guarantees

As a warmup to our main lower bound, in this section, we present a tradeoff between the statistical guarantees ϵ_A and ϵ_B of Alice and Bob, resp.:

Proposition 4.1 *In any random selection protocol* Π *over universe* $\mathcal{U}$ *achieving statistical guarantees* ϵ_A *and* ϵ_B, $\epsilon_A + \epsilon_B \geq 1 - 1/N$, *where* $N = |\mathcal{U}|$.

Corollary 4.2 *In any random selection protocol* Π, $\max\{\epsilon_A, \epsilon_B\} \geq 1/2 - 1/(2N)$.

Proof:

Suppose we have a protocol where $\epsilon_A + \epsilon_B < 1 - 1/N$. Then we can partition the universe into two sets, S and $\mathcal{U} - S$, where $|S| > \epsilon_A N$ and $|\mathcal{U} - S| > \epsilon_B N$.

View the protocol as a game where Alice wins if the output lands in S and Bob wins if the output lands in $\mathcal{U} - S$. A well-known result in game theory is Zermelo's theorem: that, in such a game, one of the players will have a winning strategy (one that wins regardless of how the other player plays). The basic reasoning is *backwards induction* on the game tree: every leaf node can be labelled A-WIN or B-WIN, and then we inductively label the remaining nodes depending on whether there exists a winning child for the current player to select. If there is, the current player will choose that child and will thus have a winning strategy from the current node. If there is not, then the opposing player will certainly win from the current node, as all children of the node lead to nodes from which he or she has a winning strategy.

This result implies one of the following:

- There exists strategy A^* where $\Pr[f((A^*, B^*)) \in S] = 1$, for any B^*. Taking $B^* = B$ (Bob's honest strategy), this contradicts the guarantee of ϵ_B given to Bob, since $|\mathcal{U} - S| > \epsilon_B N$.
- There exists strategy B^* where $\Pr[f((A^*, B^*)) \in \mathcal{U} - S] = 1$, for any A^*. This similarly contradicts guarantee ϵ_A.

■

The main intuition behind the proof is that, at every stage, either there exists a move that is good for the current player or all moves are good for the other player. In either case, the result is good for one of the two players. All that is needed is a way to make sure that every node on the bottom level can be defined as "winning" for someone, and that this notion can propagate up the tree. As we will see, this type of reasoning will figure strongly in the proof of our main lower bound. There, the primary challenge will be to handle the cases when some nodes do not appear to be "winning" for either player.

4.2 The Main Lower Bound

In this section, we prove Theorem 1.3, giving a lower bound on round complexity matching the Iterated Random Shift Protocol up to a factor of $2 + o(1)$.

Theorem 4.3 (Thm. 1.3, restated) *For any $\epsilon, \mu > 0$, there exists constant c such that any random selection protocol on a universe of size N satisfying the Statistical Criterion with parameters ϵ and μ requires at least $\log^* N - \log^* \log^* N - c$ rounds.*

Corollary 4.4 *If there exists constant δ such that a protocol Π achieves ϵ_A, $\epsilon_B \leq 1 - \delta$, then there exists constant c such that Π has at least $\log^* N - \log^* \log^* N - c$ rounds.*

To prove this theorem, we must show that in a protocol with "few" rounds, one of the two players will be able to find a set of small size that will contain the output with high probability. We will refer to such a set (that the cheating player is trying to make the output fall in) as the *cheating set.* The proof will rely to some degree on the probabilistic method: we will show the existence of such a cheating set by assuming it is chosen, at least in part, *randomly.* Specifically, we will prove:

Theorem 4.5 *There exists a function f such that for any $\mu, \epsilon > 0, r \in \mathbb{N}$ and protocol Π with r rounds, one of the following three cases holds:*

1. *When R is a randomly chosen set of density μ, and Alice plays a strategy maximizing the probability that the output of the protocol falls in R assuming that Bob plays honestly, she will succeed with probability $1 - \epsilon$, on average over all possible R. That is, we say that $E_R \left[\max_{A^*} \left\{\Pr_B \left[\Pi(A^*, B) \in R\right]\right\}\right] \geq 1 - \epsilon$.*
2. *$E_R[\max_{B^*}\{\Pr_A[\Pi(A, B^*) \in R]\}] \geq 1 - \epsilon$.*
3. *When R is a randomly chosen set of density μ, both Alice and Bob can force the output into R plus an additional $o(N)$ elements with high probability. That is, the following two conditions hold:*
   - *(a) $\exists T$, $|T| \leq f(r, \epsilon, \mu)$, such that*

     $$E_R[\max_{A^*}\{\Pr_B[\Pi(A^*, B) \in R \cup T]\}] \geq 1 - \epsilon$$

   - *(b) $\exists S$, $|S| \leq f(r, \epsilon, \mu)$, such that*

     $$E_R[\max_{B^*}\{\Pr_A[\Pi(A, B^*) \in R \cup T]\}] \geq 1 - \epsilon$$

Moreover, f does not grow too fast in r. Specifically, there exists a function $\zeta(\epsilon, \mu)$ such that $f(r, \epsilon, \mu)$ is $o(N)$ for all $r \leq \log^* N - \log^* \log^* N - \zeta(\epsilon, \mu)$.

Putting the three conditions together, this theorem says that either one player can make the output fall into a random set of certain density with high probability, or *both* players can make the output fall into a set consisting of a randomly chosen set of certain density and a certain bounded number of (non-random) elements. We call a protocol in Case 1 a WIN protocol for Alice, Case 2 a WIN protocol for Bob, Case 3 a TIE protocol for both.

To prove Theorem 4.3 using Theorem 4.5, suppose the protocol satisfied the Statistical Criterion with parameters μ' and ϵ'. Then we can set μ slightly less than μ', ϵ slightly less than ϵ', and Cases 1 and 2 would violate the Statistical Criterion. (By averaging, there exists a fixed set R of density μ such that one of the players can force the output into R with probability at least $1-\epsilon$.) Case 3 would also violate it for sufficiently large N, if $f(r,\epsilon,\mu)$ is $o(N)$, which holds unless $r \geq \log^* N - \log^* \log^* N - O(1)$.

Proving Theorem 4.5 will require an intricate analysis of the game tree using backwards induction. Like the proof of Proposition 4.1, we will show how to "label" the nodes of the game tree, where each label corresponds to a power of a player to force a particular outcome. To build intuition for the full result, we begin by proving why the Statistical Criterion cannot be achieved by any protocol of at most r rounds for $r = 1, 2, 3$, in the process sketching the key ideas of Theorem 4.5.

4.2.1 Proof Ideas

We stress that the informal discussion in this section is only meant to convey the main ideas, and the reader who prefers a more precise treatment right away can skip to Section 4.2.2. For a visual depiction of the ideas presented in this section, the reader is directed to the conference talk [San05].

$\mathbf{r = 1.}$ In a 1-round protocol, Alice sends a message that determines the output of the protocol. Certainly the Statistical Criterion cannot be achieved here: Alice can fix the output and so the output will fall with probability 1 in a set of density $1/N$, which will be less than any μ for sufficiently large N. (Note that this is not sufficient to establish Theorem 4.5 for the case $r = 1$, but this will be done by our proof below that 2-round protocols cannot meet the Statistical Criterion.)

$\mathbf{r = 2.}$ In a 2-round protocol, the output is a function of an initial message β from Bob and then a message α from Alice. Suppose such a protocol satisfies the Statistical Criterion with parameters μ, ϵ.

Note that Bob's message β defines a distribution $\mathcal{D}_\beta$ whereby Alice chooses the output. We divide the analysis into cases depending on the size of the support $S_\beta = \mathrm{Support}(\mathcal{D}_\beta)$:

$\mathbf{r = 2,}$ **Case I**. There exists a Bob message β such that $|S_\beta| \leq s(\mu,\epsilon)$, where $s(\mu,\epsilon)$ is a sufficiently large constant to be defined later. In this case, Bob can force the output into the set S_β with probability 1 by sending β as his message. This certainly violates the Statistical Criterion, since $s(\mu,\epsilon)/N < \mu$ for sufficiently large N.

$\mathbf{r = 2,}$ **Case II**. For every Bob message β, $|S_\beta| > s(\mu,\epsilon)$. Then the key observation is that, for an appropriate choice of the function $s(\mu,\epsilon)$, if Alice chooses a set R of density μ at random, then $R \cap S_\beta \neq \emptyset$ with probability $1-\epsilon$ over her choice R and Bob's choice β, in which case Alice will be able to select an output in R. This corresponds to a WIN for Alice in Theorem 4.5.

Basically what we have done is proven a simple case of Theorem 4.5 for the 1-round protocol induced by Bob's message β, where Case I corresponds to Case (3b) of the Theorem (taking $S = S_\beta$), and Case II corresponds to Case (1) of the Theorem (WIN for Alice).

$\mathbf{r = 3}$ Assume Alice goes first, sending a message γ, after which Bob sends a message β, and Alice sends a message α. We will again denote by $S_{\gamma,\beta}$ the set of possible outputs when the messages γ and β are fixed and α varies. Fix μ and ϵ that purportedly satisfy the Statistical Criterion.

First, inductively we observe that no "child" (i.e., 2-round protocol based on Alice's first message γ) can be a WIN for Alice (i.e., such that for all β, $|S_{\gamma,\beta}| > s(\mu, \epsilon)$), because then by choosing this child Alice can contradict the Statistical Criterion by the analysis in Case II of the proof for $r = 2$. It follows that for every child γ, Bob can choose a message β such that $|S_{\gamma,\beta}| \leq s(\mu, \epsilon)$. The basic issue now is that although Bob knows he will have the ability to choose a small support, he doesn't know *which* small support he will be able to choose, as this is a function of Alice's first message.

$r = 3$**, Case I**. For every Alice message γ, there exists a collection of $s'(\mu, \epsilon, s(\mu, \epsilon))$ choices for β that yield *disjoint* sets $S_{\gamma,\beta}$. Then, generalizing the probabilistic argument from above, we observe that for an appropriate choice of the function s', if Bob chooses a cheating set R of density μ at random, then with probability greater than $1 - \epsilon$ (over the choice of Alice's message γ and Bob's choice of R), there will exist a β such that $S_{\gamma,\beta} \subseteq R$. Bob can subsequently send the message β, forcing the output to fall in R. But the output falling into R, $\mu(R) \leq \mu$, with probability greater than $1 - \epsilon$ contradicts the Statistical Criterion. This protocol is a WIN for Bob in Theorem 4.5.

$r = 3$**, Case II**. There exists an Alice message γ such that there do not exist $s'(\mu, \epsilon, s(\mu, \epsilon))$ disjoint small sets $S_{\gamma,\beta}$. The key here is the following fact (proven in Lemma 4.14): If a collection $\mathcal{S}$ of nonempty sets, each of size at most s, does not have a disjoint subcollection of size t, then there exists a set X of size at most $t \cdot s$ such that $X \cap S \neq \emptyset$ for all $S \in \mathcal{S}$.

The reason: let $\mathcal{S}'$ be the largest subcollection $\mathcal{S}' \subseteq \mathcal{S}$ where $S_i \cap S_j = \emptyset$ for all $S_i, S_j \in \mathcal{S}'$. Then $X = \bigcup \mathcal{S}'$ will certainly have $|X| \leq t \cdot s$ and moreover, $X \cap S \neq \emptyset$ for all $S \in \mathcal{S}$ because otherwise we contradict the assumption $\mathcal{S}'$ is the largest disjoint subcollection.

We conclude that when fewer than $s'(\mu, \epsilon, s(\mu, \epsilon))$ disjoint small supports exist, we can produce a set X, $|X| \leq s'(\mu, \epsilon, s(\mu, \epsilon)) \cdot s(\mu, \epsilon)$ intersecting every small support. Now, Alice can set her cheating set to be $X \cup R$, where R is randomly chosen of density $\mu' < \mu$. She can send γ as her first message. Then, if Bob chooses a message β leading to a small support $S_{\gamma,\beta}$, by design $X \cap S_{\gamma,\beta} \neq \emptyset$ and Alice can make the output fall in her set $X \cup R$. Otherwise, Bob will choose a large support $S_{\gamma,\beta}$, and so $R \cap S \neq \emptyset$ with probability greater than $1 - \epsilon$. Either way, since $\mu(X \cup R) < \mu$ for sufficiently large N, we will contradict the Statistical Criterion.

$\mathbf{r = 4.}$ We do not present this case in detail, but rather just outline its high-level structure, which reflects the structure of the full induction required to prove Theorem 4.5. Suppose Bob goes first, sending a message δ and assume the Statistical Criterion holds for ϵ' and μ'.

Certainly, if there exists a choice of δ producing an induced subprotocol that is a WIN for Bob, (i.e., a choice whereby for every Alice message γ, there are more than $s'(\mu, \epsilon, s(\mu, \epsilon))$ disjoint sets $S_{\delta,\gamma,\beta}$), then this protocol is a WIN for him and he can violate the Statistical Criterion by choosing that message δ and then applying the strategy from the $r = 3$ analysis. Similarly, if for all messages δ Bob can send, the induced subprotocol is a WIN for Alice, then this protocol is a WIN for Alice too (this would correspond to the case where, for every message Bob can send, there exists a message Alice can send wherein Bob would be forced to send a large support to Alice).

Otherwise, some messages δ lead to a protocol corresponding to the $r = 3$, Case II (where the induced subprotocol is a TIE, as in Case 3a of Thm. 4.5): Alice can pick a message γ, so that there is a set X of size $s'' = s'(\mu, \epsilon, s(\mu, \epsilon)) \cdot s(\mu, \epsilon)$ intersecting every small set $S_{\delta,\gamma,\beta}$. Just as before, the problem for Alice is that the set X to use depends on Bob's first message δ. As above, we partition the analysis into two cases, depending on whether or not there are many disjoint possibilities for the set X. If yes, then a random set will encompass such a set X with high probability, and it is a WIN for Alice. If not, there is a small set Y intersecting all these choices for X. Here, however, the use of this fact to construct an effective strategy for Bob is more subtle than in the $r = 3$ case.

Many of the technical ideas used in the full proof of Theorem 4.5 already occur in the cases above. However, setting up a claim suitable for proof by induction is somewhat delicate, and is done via the lengthy statements of Definition 4.8 and Lemma 4.10 in the next section. Jumping ahead, the reason why the induction will stop at $\log^* n$ rounds is that the sizes of the "small" sets (e.g. the functions s, s', s'' in the intuition above) grow like a tower with the number of rounds.

4.2.2 Proof of Theorem 4.5

We proceed by backwards induction on the game tree of the protocol.

Definition 4.6 *Given a protocol Π with r rounds and constants ϵ and μ, let $f(r, \epsilon, \mu) = g(r, r, \epsilon, \mu)$, where*

$$\begin{aligned} g(0, r, \epsilon, \mu) &= 1 \\ g(k, r, \epsilon, \mu) &= (r/\epsilon) \cdot (re/\mu)^{g(k-1,r,\epsilon,\mu)} \cdot g(k-1, r, \epsilon, \mu) \end{aligned}$$

For clarity we write s_k for $g(k, r, \epsilon, \mu)$, as r, ϵ, and μ will remain fixed throughout the proof.

A concept that will prove helpful is that of a maximal disjoint subcollection.

Definition 4.7 *Let $\mathcal{S}$ be a collection of sets over a given universe $\mathcal{U}$. Then a* maximal disjoint subcollection *$\mathcal{P}$ of $\mathcal{S}$ is a collection $\mathcal{P} \subseteq \mathcal{S}$ satisfying $S \cap T = \emptyset$ for all $S, T \in \mathcal{P}$, and for every $T \in \mathcal{S} \setminus \mathcal{P}$, there exists $S \in \mathcal{P}$ such that $S \cap T \neq \emptyset$.*

Such a subcollection always exists, so we will refer to the *canonical* maximal disjoint subcollection to be one chosen by some fixed but arbitrary method.

Now, fix a protocol Π with r rounds, and consider the game tree T it induces (see Definition 2.2). At each node of this tree, we will associate a certain collection of sets (subsets of the universe $\mathcal{U}$). These sets will correspond to the sets S and T of case (3) of Theorem 4.5. This association will be defined inductively on the game tree.

Specifically, we inductively label the nodes of the tree as either A-WIN, A-LOSE, A-TIE, B-WIN, B-LOSE, or B-TIE. For each of the TIE nodes, we will also associate a collection $\mathcal{S}_z$ of subsets of $\mathcal{U}$, as defined below. The 'A' or 'B' just tells us whose turn it is, and as we will see, WIN, LOSE, and TIE will say something about the power of the player whose turn it is at that point.

Definition 4.8 *Fix a protocol Π. Let z be a node on its game tree at level k (where leaves are at level 0). Assume it is Alice's turn at this node.*

If $k = 0$ (i.e., z is a leaf of the tree) then label z as A-TIE. Moreover, let $\mathcal{S}_z = \{\{x\}\}$, where x is the output of the protocol ending at node z.

If $k > 0$, consider the children $z_1, \ldots, z_\ell$ of z. Use the following rules to label the nodes:

1. *If there exists $1 \leq i \leq \ell$ such that z_i is in case* B-LOSE*, then label z as* A-WIN.

2. *If, for all $1 \leq i \leq \ell$, z_i is in case* B-WIN*, then label z as* A-LOSE.

3. *Otherwise, denote $\mathcal{T}_z = \{S : z_i$ is* B-TIE *and $S \in \mathcal{S}_{z_i}\}$. That is, $\mathcal{T}_z$ is the union of the collections of sets associated with all children of z that are labelled* B-TIE*. Now, let $\mathcal{P}$ denote the canonical maximal disjoint subcollection of $\mathcal{T}_z$, as defined above, and let s_k, s_{k-1} be defined as in Definition 4.6.*

   Two cases:

   (a) $|\mathcal{P}| \geq s_k/s_{k-1} \Rightarrow$ *label z as* A-WIN.
   (b) $|\mathcal{P}| < s_k/s_{k-1} \Rightarrow$ *label z as* A-TIE*, and define $\mathcal{S}_z$ to be $\{S \subseteq \mathcal{U} : |S| \leq s_k$ and $S \cap T \neq \emptyset$ for all $T \in \mathcal{T}_z\}$. That is, the sets associated with z consist of all sets that intersect all of the sets associated to the children z_i (which will be in case* B-TIE*, since those are the only nodes to which we associate sets), and have size $\leq s_k$.*

Likewise, label all nodes at which it is Bob's turn, by swapping A *with* B *in the above specification.*

Intuitively, this structure defines the power of the players at various stages of the protocol. The WIN, LOSE, and TIE nodes refer to cases (1), (2), and (3) of Theorem 4.5. Moreover, the collections $\mathcal{S}_z$ correspond to S and T in Case (3) of Theorem 4.5.

We codify this power in Lemma 4.10. Before stating it, it will help to define the following:

Definition 4.9 *Let $\Pi = (A, B, f)$ be a protocol, and let T be its equivalent game tree (see Definition 2.2). For any node $z = (m_1, \ldots, m_{r-k})$ on level k of the tree T, let $\Pi_z = (A_z, B_z, f_z)$ be a protocol of k rounds, where $f_z(m'_1, \ldots, m'_k) = f(m_1, \ldots, m_{r-k}, m'_1, \ldots, m'_k)$, and where A_z and B_z denote the strategies of A and B conditioned on history z (i.e. we choose their coin tosses r_A and r_B uniformly from those consistent with the history).*

Intuitively, Π_z is the protocol induced by starting the protocol at node z (i.e., assuming all messages leading to z are fixed in advance).

Lemma 4.10 *Fix ϵ and μ, and suppose the protocol has r turns. Let z be some node on the tree at level k, at which it is Alice's turn to play. Throughout, let R be a uniformly random subset of $\mathcal{U}$ of density $k\mu/r$.*

1. *If z is in case* A-WIN*, then*

$$E_R\left[\max_{A^*}\left\{\Pr_B\left[\Pi_z(A^*, B) \in R\right]\right\}\right] \geq 1 - k\epsilon/r$$

   where R is a random subset of $\mathcal{U}$ of density $k\mu/r$, and Π_z is the protocol induced by beginning at node z, as defined in Definition 4.9. (We say Alice can "win" from node z).

2. *If z is in case* A-LOSE*, then similarly,*

$$E_R\left[\max_{B^*}\left\{\Pr_A\left[\Pi_z(A, B^*) \in R\right]\right\}\right] \geq 1 - k\epsilon/r$$

   (We say Bob can "win" from node z).

3. *If z is in case* A-TIE*, then:*

   (a) *$\mathcal{S}_z$ and, if $k > 0$, $\mathcal{T}_z$, are nonempty.*

   (b) *When $k > 0$, then for any $T \in \mathcal{T}_z$,*

$$E_R \left[\max_{A^*} \left\{ \Pr_B [\Pi_z(A^*, B) \in R \cup T] \right\} \right] \geq 1 - k\epsilon/r$$

   (c) *For any $S \in \mathcal{S}_z$,*

$$E_R \left[\max_{B^*} \left\{ \Pr_A [\Pi_z(A, B^*) \in R \cup S] \right\} \right] \geq 1 - k\epsilon/r$$

   (We say both Alice and Bob "win" from node z, with "helper sets" T and S, respectively).

Moreover, the same (with "Alice" exchanged for "Bob", and "A" exchanged for "B") holds for all nodes for which it is Bob's turn.

Lemma 4.10 more precisely asserts Theorem 4.5 at each level of the game tree. To use this lemma to prove Theorem 4.5, we simply need to apply it with $k = r$ and z being the root of the game tree. Certainly, if z_r is in Case (1) or (2) of Lemma 4.10, it is in Case (1) or (2) of Theorem 4.5, respectively. If z_r is in Case (3) of Lemma 4.10, then subcases (3.a), (3.b), and (3.c) directly prove subcases (3.a) and (3.b) respectively, where the sets of $\mathcal{S}_z$ and $\mathcal{T}_z$ of (3.a) and (3.b) correspond precisely to the sets T and S we need in those subcases of the theorem. The existence of such sets is guaranteed by subcase (3.a) of the lemma.

We prove Lemma 4.10 by induction on the levels of the tree.

Base Cases:
$\mathbf{k = 0}$: So z is a leaf node, and the output of Π_z is just deterministically fixed at, say, x. According to Definition 4.8, $\mathcal{S}_z = \{\{x\}\}$, and we are in case A-TIE. Since the density of R must be zero (it is $k\mu/r$), $R = \emptyset$ and so we need to show that

$$\max_{B^*} \left\{ \Pr_A [\Pi_z(A, B^*) \in \{x\}] \right\} = 1$$

This of course holds because the output is fixed at x.

$\mathbf{k = 1}$: We assume without loss of generality that it is Alice's turn at node z. Notice first that all of the children of z are in case B-TIE, by the reasoning in the $k = 0$ case. Consequently, z must be labelled A-WIN or A-TIE. Which case we're in depends directly on the size of the canonical maximal disjoint subcollection $\mathcal{P}$ of $\mathcal{T}_z$ ($\mathcal{T}_z$, recall, is $\{S : z_i \text{ is B-TIE and } S \in \mathcal{S}_{z_i}\}$). Notice that since $k = 1$, all of the children of z are labelled B-TIE and $\mathcal{T}_z = \{\{x\} : \exists i \text{ such that } x \text{ is the output at } z_i\}$. It follows that $|\mathcal{P}| = |\{x : \exists i \text{ such that } x \text{ is the output at } z_i\}|$. $|\mathcal{P}|$ is precisely the size of the support of the distribution by which Alice chooses the output of the protocol.

k = 1, Case 1: Suppose $|\mathcal{P}| \geq s_1/s_0 = s_1$. Then by Rule 3a of Definition 4.8, z is in case A-WIN. Thus, we must verify that $E_R[\max_{A^*}\{\Pr_B[\Pi_z(A^*, B) \in R]\}] \geq 1 - \epsilon/r$, where R is a random subset of the universe of density μ/r. That is, if at node z Alice plays to maximize the probability that the output falls in a set R of density μ/r, her average probability of success over choices of R will be $1-\epsilon/r$.

We use the following lemma, which intuitively says that if we have a large number of disjoint small sets, then one would expect a randomly chosen set of constant density to contain one of them with high probability.

Lemma 4.11 *Suppose we have a collection of disjoint sets $S_1,\dots S_m$ over a fixed universe $\mathcal{U}$, $|\mathcal{U}|=n$, where for all i, $|S_i|\le s$. Choose a set R randomly of density μ (i.e., $n\cdot\mu$ distinct elements). With probability $\ge 1-(1/m)(e/\mu)^s$, there will exist S_i such that $S_i\subseteq R$.*

The proof is by Chebyshev's inequality, and is deferred to the appendix.

We know $\mathcal{P}$ consists of $m\ge s_1$ disjoint sets. So, by Lemma 4.11, a randomly chosen set of density μ/r will contain an element of $\mathcal{P}$ with probability $\ge 1-(1/m)(re/\mu)$ (recall the sets have size 1), where $m\ge s_1=(r/\epsilon)(re/\mu)$. Thus, it will contain an element of $\mathcal{P}$ with probability $1-\epsilon/r$. (Notice that Lemma 4.11 explains the way we defined the constants s_k in Definition 4.6.)

In such an event, when the random set R contains an $S\in\mathcal{P}$, we claim there exists a strategy A^* for Alice whereby $\Pi_z(A^*,B)\in R$. By definition, $S=\{x\}$, where x is the output at some leaf z_i that is a child of z. To force the output into R, Alice can play the strategy A^* which selects z_i on her turn at node z. Whenever R contains $S\in\mathcal{P}$ this strategy succeeds with probability 1, and since this event occurs for at least a $1-\epsilon/r$ fraction of the choices of R, it follows that $E_R[\max_{A^*}\{\Pr_B[\Pi_z(A^*,B)]\}]\ge 1-\epsilon/r$.

k = 1, Case 2: Suppose $\mathcal{P}$ consists of $m<s_1$ elements, and thus z is in case A-TIE. To prove z satisfies Lemma 4.10 in this case, we must verify the following: (a) $\mathcal{S}_z$ and $\mathcal{T}_z$ are nonempty, (b), for any $T\in\mathcal{T}_z$,

$$E_R\left[\max_{A^*}\left\{\Pr_B\left[\Pi_z(A^*,B)\in R\cup T\right]\right\}\right]\ge 1-\epsilon/r$$

and (c), for any $S\in\mathcal{S}_z$,

$$E_R\left[\max_{B^*}\left\{\Pr_A\left[\Pi_z(A,B^*)\in R\cup S\right]\right\}\right]\ge 1-\epsilon/r$$

$\mathcal{T}_z$ is certainly nonempty; each child z_i of z is a B-TIE node, and $\mathcal{T}_z$ consists of singletons of each's output.

Since $|\mathcal{P}|\le s_1$, and since all of the elements of $\mathcal{T}_z$ are singletons, it follows that fewer than s_1 distinct elements appear in the sets of $\mathcal{T}_z$. A set S consisting of precisely these elements will be a set of size less than s_1 intersecting every set in $\mathcal{T}_z$—thus $S\in\mathcal{S}_z$ by definition, and $\mathcal{S}_z$ is nonempty. This verifies condition (a) above.

To verify condition (b), notice that node z is Alice's turn, and so Bob has no influence on the output of the protocol. Moreover, any $T\in\mathcal{T}_z$ consists of a single element x that is the output of the protocol at a child z_i. Thus, for any such T, Alice can play the strategy A^* which selects the corresponding child z_i. Thus, $\Pr_B[\Pi_z(A^*,B)\in T]=1$, and *a fortiori*, $E_R[\max_{A^*}\{\Pr_B[\Pi_z(A^*,B)\in R\cup T]\}]\ge 1-\epsilon/r$. This will succeed with probability 1, as well.

To verify condition (c), recall that any set $S\in\mathcal{S}_z$ must intersect every set in $\mathcal{T}_z$, which, when $k=1$, implies that it contains the entire support by which Alice will choose the output. Thus, $\Pr_A[\Pi_z(A, B^*) \in S] = 1$ by definition, which is again stronger than what we need.

Inductive Step. Suppose Lemma 4.10 holds for nodes on all levels up to level $k-1$. We will show that it holds for an arbitrary node z on level k. Assume it is Alice's turn at z. There are several possibilities:

Claim 4.12 *If z is in case* A-LOSE*, then*

$$E_R\left[\max_{B^*}\left\{\Pr_A\left[\Pi_z(A, B^*) \in R\right]\right\}\right] \geq 1 - k\epsilon/r$$

where R is a random subset of density $k\mu/r$.

Proof: We will use Definition 4.8 and the inductive hypothesis to show that every child node z_i is "good" for Bob–that is, on average over R, B^* can make the outcome land in R with probability $1-(k-1)\epsilon/r$. Then certainly the same holds for node z, since Alice cannot help but move to such a node.

Formally, we first notice that it suffices to show:

$$E_{R'}\left[\max_{B^*}\left\{\Pr_A\left[\Pi_z(A, B^*) \in R'\right]\right\}\right] \geq 1 - (k-1)\epsilon/r$$

where R' is a ranges over all sets of density $(k-1)\mu/r$.

Now, for z to be labelled A-LOSE, we must have used Rule 2 of Definition 4.8. Thus, all of the children of z are in case B-WIN. By the inductive hypothesis:

$$E_{R'}\left[\max_{B^*}\left\{\Pr_A\left[\Pi_{z_i}(A, B^*) \in R'\right]\right\}\right] \geq 1 - (k-1)\epsilon/r \tag{1}$$

for each child z_i of z, where R' ranges over all sets of density $(k-1)\mu/r$. Since at node z it is Alice's turn, we have

$$\begin{aligned} E_{R'}\left[\max_{B^*}\left\{\Pr_A\left[\Pi_z(A, B^*) \in R'\right]\right\}\right] &= E_{R', z_i \leftarrow D_z}\left[\max_{B^*}\left\{\Pr_A\left[\Pi_{z_i}(A, B^*) \in R'\right]\right\}\right] \\ &\geq 1 - (k-1)\epsilon/r, \end{aligned}$$

where D_z is the distribution according to which Alice chooses child z_i of z when playing honestly, and the last inequality is by (1). ■

Claim 4.13 *If z is in case* A-WIN*, then*

$$E_R\left[\max_{A^*}\left\{\Pr_B\left[\Pi_z(A^*, B) \in R\right]\right\}\right] \geq 1 - k\epsilon/r$$

where R is a random subset of density $k\mu/r$.

Proof: By Definition 4.8, z could have been labelled A-WIN either by Rule 1 or Rule 3a.

In Rule 1, z has a child z_j that is in case B-LOSE. Since it is Alice's turn at node z, if she can choose a node z_j "good" for her then node z will be "good" for her too.

Formally, by the inductive hypothesis applied to z_j, we have that

$$E_R\left[\max_{A^*}\left\{\Pr_B\left[\Pi_{z_j}(A^*,B)\in R\right]\right\}\right]\geq 1-(k-1)\epsilon/r$$

But $\max_{A^*}\{\Pr_B[\Pi_z(A^*,B)\in R]\}$ will always be at least $\max_{A^*}\left\{\Pr_B\left[\Pi_{z_j}(A^*,B)\in R\right]\right\}$, since node z is Alice's turn and she can always at least choose z_j. Taking expectations of both sides, the claim is proven for this case.

The alternative possibility is that z is in A-WIN because of Rule 3a. So among the sets $\mathcal{T}_z$ (for all children z_i in B-TIE), we can find a disjoint subcollection $\mathcal{P}$, where $|\mathcal{P}|\geq s_k/s_{k-1}$.

Intuitively, what is going on here? Since no B-LOSE nodes are available among the children of z, Alice cannot simply choose such a branch as above. However, we know that from the B-TIE nodes, for a high proportion of sets R, Alice can ensure the output lands in $S\cup R$ where $S\in\mathcal{S}_{z_i}$, with high probability. But this is true for many possible sets S—not only at a given child, but also across all the potential children that are in case B-TIE (i.e., any $S\in\mathcal{T}_z$). Thus, we can expect that with enough disjoint sets in $\mathcal{T}_z$, the random set R will encompass $S\in\mathcal{T}_z$ with high probability. The inductive hypothesis will then give the desired result.

Using Lemma 4.11, since $\mathcal{P}\subseteq\mathcal{T}_z$ consists of at least $s_k/s_{k-1}=(r/\epsilon)(re/\mu)^{s_{k-1}}$ (disjoint) sets of size at most s_{k-1}, we can conclude:

$$E_{R_1}\left[\exists S\in\mathcal{T}_z, S\subseteq R_1\right]\geq 1-\epsilon/r \tag{2}$$

where R_1 is a random subset of density μ/r.

For any $S\in\mathcal{T}_z$, we can then assert:

$$E_{R_2}\left[\max_{A^*}\left\{\Pr_B\left[\Pi_z(A^*,B)\in R_2\cup S\right]\right\}\right]\geq 1-(k-1)\epsilon/r \tag{3}$$

where R_2 is a random subset of density $(k-1)\mu/r$. This comes from applying the inductive hypothesis to the child z_j such that $S\in\mathcal{S}_{z_j}$, and since $\max_{A^*}\{\Pr_B[\Pi_z(A^*,B)\in R_2\cup S]\}$ is always at least $\max_{A^*}\left\{\Pr_B\left[\Pi_{z_j}(A^*,B)\in R_2\cup S\right]\right\}$ (because at node z it is Alice's turn).

Now, considering the selection of a random subset R of density $k\mu/r$ to be the random and independent choices of subsets R_1 and R_2 of densities μ/r and $(k-1)\mu/r$ respectively (compensating for any overlap by adding random elements), we can combine (2) and (3) to derive

$$E_R\left[\max_{A^*}\left\{\Pr_B\left[\Pi_z(A^*,B)\in R\right]\right\}\right]\quad\geq\quad(1-\epsilon/r)\cdot(1-(k-1)\epsilon/r)$$

The claim follows. ■

The final possibility is that z is in case A-TIE. Since z is not a leaf, this can only come about by Rule 3b from Definition 4.8. That is, no children of z are in case B-LOSE, and at least some are in B-TIE. Moreover, among $\mathcal{T}_z$ the canonical maximal disjoint subcollection $\mathcal{P}$ has less than s_k/s_{k-1} elements.

We must prove the following: $\mathcal{S}_z$ is nonempty, $\mathcal{T}_z$ is nonempty, Alice can win from this node with a helper set from $\mathcal{T}_z$, and Bob can win from this node with a helper set from $\mathcal{S}_z$ (see Lemma 4.10).

We will require the following combinatorial lemma:

Lemma 4.14 *Let $\mathcal{S}$ be a collection of nonempty sets $S_1, \ldots S_m$ over a finite universe $\mathcal{U}$, $|\mathcal{U}| = N$, where for all i, $|S_i| \leq s$. Suppose that $\mathcal{S}$ has a maximal disjoint subcollection of size t. Then there exists a set X that intersects every $S \in \mathcal{S}$ (i.e., $X \cap S \neq \emptyset$), and $|X| \leq t \cdot s$.*

That is, either a collection of small sets has many disjoint members or it has a small "intersect-set"—a set intersecting each member of the collection. Intuitively, the union of a maximal disjoint subcollection must intersect every set, for otherwise one could add the disjoint set to form a larger disjoint subcollection.

Proof: Given a maximal disjoint subcollection $\mathcal{P} = \{S_1, \ldots, S_t\}$, we can define $X = \bigcup_{1 \leq i \leq t} S_i$. That $|X| \leq t \cdot s$ follows from the assumption that all $S \in \mathcal{S}$ have $|S| \leq s$. Since $S \neq \emptyset$ for any $S \in \mathcal{S}$, $X \cap S \neq \emptyset$ for any $S \in \mathcal{P}$. Now, suppose there exists $S \notin \mathcal{P}$ such that $X \cap S = \emptyset$. But then, by the definition of X, $\mathcal{P} \cup \{S\}$ consists of $t+1$ disjoint sets, contradicting the assumption that $\mathcal{P}$ was a maximal disjoint subcollection. So $X \cap S \neq \emptyset$ for all $S \in \mathcal{S}$. ■

Claim 4.15 $\mathcal{S}_z \neq \emptyset$ *and* $\mathcal{T}_z \neq \emptyset$*.*

Proof: By Lemma 4.14, since the canonical maximal disjoint subcollection $\mathcal{P}$ of $\mathcal{T}_z$ has size $\leq s_k/s_{k-1}$ and since all $S \in \mathcal{P}$ have size $\leq s_{k-1}$, there exists a set X of size at most $(s_k/s_{k-1}) \cdot s_{k-1} = s_k$ intersecting every set in $\mathcal{T}_z$. By Definition 4.8, $X \in \mathcal{S}_z$.

We have already established that z has children in case B-TIE (this follows from Definition 4.8 and from our assumption that $z \in$ A-TIE). By the inductive hypothesis on such a child z_i, $\mathcal{S}_{z_i}$, and thus $\mathcal{T}_z$ is nonempty. ■

Claim 4.16 $E_R[\max_{A^*}\{\Pr_B[\Pi_z(A^*, B) \in S \cup R]\}] \geq 1 - k\epsilon/r$ *where R is a random subset of density $k\mu/r$, and S is any element of $\mathcal{T}_z$.*

The proof of this claim is identical to the proof of equation (3) in the proof of Claim 4.13.

The final claim required to prove Lemma 4.10 is the following:

Claim 4.17 $E_R[\max_{B^*}\{\Pr_A[\Pi_z(A, B^*) \in S \cup R]\}] \geq 1 - k\epsilon/r$ *where R is a random subset of density $k\mu/r$, and S is any element of $\mathcal{S}_z$.*

This claim is the heart of the entire proof. All we know now is that there is at least one B-TIE node that is a child of the current node z, and that among the corresponding sets in $\mathcal{T}_z$, the canonical maximal disjoint subcollection $\mathcal{P} \subseteq \mathcal{T}_z$ contains fewer than s_k/s_{k-1} sets. That $\mathcal{P}$ is so small is a limitation on the power of Alice, who would like there to be enough such disjoint sets in $\mathcal{P}$ that she could choose randomly and encompass a set in $\mathcal{P}$ with high probability. The key to this proof is converting this *limitation* on Alice into an *ability* for Bob to cheat.

Proof: Fix a set $S \in \mathcal{S}_z$. Since an honest Alice will choose a child z_i at random, it suffices to prove the following for each child z_i:

$$E_R\left[\max_{B^*}\left\{\Pr_A\left[\Pi_{z_i}(A, B^*) \in S \cup R\right]\right\}\right] \geq 1 - k\epsilon/r \tag{4}$$

where R is a random subset of density $k\mu/r$. So fix an arbitrary child z_i. Looking to Definition 4.8, the only way we could have defined z to be in case A-TIE is if all children z_i are either in case B-WIN or case B-TIE. So z_i is in one of these two cases.

If z_i is in case B-WIN, then we are done by the inductive hypothesis. So suppose z_i is in case B-TIE. Applying the inductive hypothesis to z_i, we know that $\mathcal{T}_{z_i}$ is nonempty. Moreover, for any $T \in \mathcal{T}_{z_i}$:

$$E_{R_1}\left[\max_{B^*}\left\{\Pr_A\left[\Pi_{z_i}(A, B^*) \in T \cup R_1\right]\right\}\right] \geq 1 - (k-1)\epsilon/r \tag{5}$$

where R_1 is a random subset of density $(k-1)\mu/r$.

We divide the proof in two cases: First, suppose there exists $T \in \mathcal{T}_{z_i}$ such that $T \subseteq S$. Then (4) follows immediately from (5). Otherwise, consider the collection of sets $\mathcal{T} = \{T - S : T \in \mathcal{T}_{z_i}\}$. By assumption, $\emptyset \notin \mathcal{T}$.

Claim 4.18 *Let $\mathcal{T}$ be defined as above. Assuming $\emptyset \notin \mathcal{T}$, then there exists a disjoint subcollection $\mathcal{T}' \subseteq \mathcal{T}$, where $|\mathcal{T}'| \geq s_{k-1}/s_{k-2}$.*

Informally, there aren't many disjoint sets in $\mathcal{T}_{z_i}$—if there were, we would have labelled z_i as a case B-WIN node for Bob. That said, by intersecting every (small) set that intersected every set in $\mathcal{T}_{z_i}$, S captures the lack of disjointness of $\mathcal{T}_{z_i}$ in the first place. This claim states that once the elements of S are removed from consideration, the result has a large number of disjoint sets.

Proof: Suppose for the sake of contradiction that $\mathcal{T}$ contains fewer than s_{k-1}/s_{k-2} disjoint sets. Recalling that these sets all have size at most s_{k-2}, and since $\emptyset \notin \mathcal{T}$, by Lemma 4.14 we can produce a set I of size at most $(s_{k-1}/s_{k-2}) \cdot s_{k-2} = s_{k-1}$ intersecting every element of $\mathcal{T}$. Without loss of generality, we can assume $I \cap S = \emptyset$ (since for every $T' \in \mathcal{T}$, $S \cap T' = \emptyset$). Since I intersects every set in $\mathcal{T}$, it follows that I intersects every set of $\mathcal{T}_{z_i}$, and since we know I has size at most s_{k-1}, we conclude that by definition, $I \in \mathcal{S}_{z_i}$. But we defined S to be an arbitrary element of $\mathcal{S}_z$, which means it intersects all elements of $\mathcal{S}_{z_i}$, including I. Contradiction. Thus, $\mathcal{T}$ contains at least s_{k-1}/s_{k-2} disjoint sets. ■

Returning to the proof of Claim 4.17, by Lemma 4.11 we may conclude the following:

$$\Pr_{R_2}\left[\exists T' \in \mathcal{T}, T' \subseteq R_2\right] \geq 1 - \epsilon/r$$

where R_2 is a random subset of density μ/r. By the definition of $\mathcal{T}$, this in turn implies:

$$\Pr_{R_2}\left[\exists T \in \mathcal{T}_{z_i}, T \subseteq S \cup R_2\right] \geq 1 - \epsilon/r$$

Using (5) and choosing R through independent choices of R_1 and R_2 as in the proof of Claim 4.13, we are done. ■

Taking together Claims 4.12, 4.13, 4.15, 4.16, and 4.17, the proof of Lemma 4.10 is complete.

■

To conclude Theorem 4.5, it remains to prove the function f defining the set sizes s_k does not grow too fast in the number of rounds. Intuitively, the reason the lower bound only holds for protocols with fewer than $\log^* n - \log^* \log^* n - O(1)$ rounds is that these "helper sets" must have no more than $o(N)$ elements to be useful, but this function f grows as a tower—where *both* the base and the height of the tower grow with the number of rounds. Our challenge is to lower bound the number of rounds that keep this tower of size $o(N)$.

Lemma 4.19 *Recall the definition $f(r,\epsilon,\mu) = g(r,r,\epsilon,\mu)$, where we define*

$$\begin{aligned} g(0,r,\epsilon,\mu) &= 1 \\ g(k,r,\epsilon,\mu) &= (r/\epsilon)(re/\mu)^{g(k-1,r,\epsilon,\mu)} g(k-1,r,\epsilon,\mu) \end{aligned}$$

There exists a function $\zeta(\epsilon,\mu)$ such that, when $r < \log^* N - \log^* \log^* N - \zeta(\epsilon,\mu)$, we have $f(r,\epsilon,\mu) \leq \log N$.

By applying Lemma 4.10 to the root of the tree and using Lemma 4.19, we prove Theorem 4.5 and thus Theorem 4.3.

5 Multiplicative Guarantees

5.1 Defining Multiplicative Guarantees

Recall Definition 5: the multiplicative difference of a distribution X from uniform is $\max_T \Pr_{x \leftarrow X}[x \in T]/\mu(T)$, where T ranges over all subsets of $\mathcal{U}$.

The multiplicative difference is always a rational in $[1,N]$, where 1 implies the uniform distribution, and N implies one element is chosen with probability 1. The multiplicative difference of a distribution is actually equal to the factor by which a single element's probability of being the protocol's output can be increased from uniform. Formally:

Lemma 5.1 *The multiplicative difference of a distribution X from uniform is equal to*

$$\max_{s \in \mathcal{U}} \left(N \cdot \Pr_{x \leftarrow X}[x = s] \right)$$

Proof: Fix a distribution X. That the multiplicative difference is at least $\max_s N \cdot \Pr_{x \leftarrow X}[x = s]$ follows because we can consider the subset $T = \{s\}$, where s maximizes $\Pr_{x \leftarrow X}[x = s]$.

For the other direction, let T be the set maximizing $\Pr_{x \leftarrow X}[x \in T]/\mu(T)$. We claim there exists an element $s \in T$ such that the following holds:

$$\Pr_{x \leftarrow X}[x \in T]/\mu(T) \leq N \cdot \Pr_{x \leftarrow X}[x = s]$$

To see this, consider an arbitrary set $T = \{x_1, \ldots x_t\}$. Then we have:

$$\Pr_{x \leftarrow X}[x \in T] = \sum_{x_i \in T} \Pr_{x \leftarrow X}[x = x_i] \leq t \cdot \max_i \Pr_{x \leftarrow X}[x = x_i].$$

This implies that there exists an i such that

$$\Pr_{x \leftarrow X}[x = x_i] \geq \frac{\Pr_{x \leftarrow X}[x \in T]}{t} = \frac{\Pr[x \leftarrow X]}{\mu(T) \cdot N}.$$

■

Multiplicative and statistical difference are different, but related. Both work by bounding a function of the probability a distribution falls in a set and the density of that set. Even so, multiplicative difference tends to focus on the concentration of probability into small sets (indeed, by Lemma 5.1, sets of size 1), while statistical difference will prove more useful when considering larger subsets (e.g., a constant fraction of the universe).

This said, we can prove some basic relationships between the two metrics that will prove useful:

Lemma 5.2 *Let X be an arbitrary distribution over universe $\mathcal{U}$, with $N = |\mathcal{U}|$. Denote by ϵ the statistical difference of X from uniform and by ρ the multiplicative difference from uniform. Then:*

1. $\rho \leq N\epsilon + 1$
2. $\epsilon \leq 1 - 1/\rho$.

That a distribution will have statistical difference at most $1-1/\rho$ is especially interesting because the relationship contains no dependence on N. This fact implies that a distribution with a constant multiplicative difference will have a constant statistical difference, though the converse is not necessarily true. Put another way, a strong multiplicative guarantee is harder to achieve than a strong statistical guarantee.

Proof of Lemma 5.2: Suppose a distribution X has statistical difference ϵ from uniform. Let ρ denote the multiplicative difference of X. Then by Lemma 5.1, we have:

$$\rho = N \cdot \max_{s \in \mathcal{U}} \Pr_{x \leftarrow X}[x = s]$$

But by the definition of statistical difference, we have for any $x \in \mathcal{U}$,

$$\Pr_{x \leftarrow X}[x = s] \leq \epsilon + 1/N$$

(Just set $T = \{s\}$.) Part (1) of the lemma follows.

For the reverse direction, suppose X has multiplicative difference ρ. It suffices to show that for all T:

$$\Pr_{x \leftarrow X}[x \in T] - \mu(T) \leq 1 - 1/\rho$$

If $\mu(T) \geq 1/\rho$, then this certainly holds, since $\Pr_{x \leftarrow X}[x \in T] \leq 1$. So suppose $\mu(T) < 1/\rho$. Then we can derive:

$$\Pr_{x \leftarrow X}[x \in T] = \sum_{s \in T} \Pr_{x \leftarrow X}[x = s] \leq |T| \cdot (\rho/N) = \mu(T)\rho$$

Plugging back into the above, we have:

$$\Pr_{x \leftarrow X}[x \in T] - \mu(T) \leq \mu(T)\rho - \mu(T) = (\rho - 1)\mu(T) < (\rho - 1)/\rho$$

■

5.2 Multiplicative Lower Bounds

In this subsection, we concentrate on lower bounds regarding multiplicative guarantees– and indeed show that no protocol exists that provides constant multiplicative guarantees to either player. This is a very strong limitation on the ability of protocols to limit cheating player's power in this regard.

An Initial Lower Bound. Proposition 4.1 can be adapted to provide a quick lower bound for multiplicative guarantees. Specifically:

Proposition 5.3 *In any random selection protocol,* $(\rho_A - 1)/\rho_A + (\rho_B - 1)/\rho_B \geq 1 - 1/N$. *Moreover,* $\epsilon_A + (\rho_B - 1)/\rho_B \geq 1 - 1/N$ *(or equivalently,* $\epsilon_A \geq 1/\rho_B - 1/N$*).*

Proof: The results follow immediately from Theorem 4.1 and from the second part of Lemma 5.2. ■

This lower bound for multiplicative guarantees is not very strong—ρ is a number from 1 to N, but this lower bound is satisfied (for instance) as long as both ρ_A and ρ_B are at least 2. In Theorem 5.4 we will prove that $\rho_A\rho_B \geq N$, which is a substantially stronger result.

On the other hand, when looking at one player getting a statistical guarantee and the other player getting a multiplicative guarantee, Proposition 5.3 does provide some useful information. Specifically, it tells us that (minus a small $1/N$ term) we can always expect the statistical guarantee for one player to be worse than the reciprocal of the multiplicative guarantee to the other player. This explains inverse relationships in existing protocols of [DGW94] (where $\varepsilon = 1/\text{poly}(n)$ and $\rho = \text{poly}(n)$) and [GSV98] (where $\varepsilon = \text{poly}(n) \cdot 2^{-k}$ and $\rho = 2^k$ for any k).[6] Notice that these earlier works focus on the case of nonconstant guarantees ($\varepsilon \to 0$ and $\rho \to \infty$). Later, we show that the Iterated Random Shift Protocol presented earlier achieves simultaneous *constant* statistical and multiplicative guarantees, and prove it has optimal round complexity up to a factor of $2 + o(1)$.

A Tight Lower Bound. The lower bound follows from the work of Goldreich, Goldwasser, and Linial [GGL98].

Theorem 5.4 [GGL98] *For any protocol* Π, $\rho_A \cdot \rho_B \geq N$.

Corollary 5.5 *In any protocol* Π, $\max\{\rho_A, \rho_B\} \geq \sqrt{N}$.

Recalling the multiplicative guarantee ρ_A is the greatest factor by which Bob can improve the probability that a single element is chosen over uniform, we conclude:

In any random selection protocol, at least one of the players can improve the probability that a single element is chosen by a factor *exponential* in the length of the output (which equals $\log N$).

Goldreich et. al. [GGL98] showed a more general result than Theorem 5.4 (for multiparty protocols) using different language and a moderately involved proof. Restricting to the two-party case, as we will see, provides a simple and elegant proof.

Proof of Theorem 5.4: Fix some element v of the universe. Now, consider the game tree T of the protocol (see Definition 2.2). At each node z of the tree, denote the protocol induced by beginning at node z to be Π_z. Then define:

$$\begin{aligned}
\phi_A^z &= \max_{A^*} \Pr_B[\Pi_z((A^*, B)) = v] \\
\phi_B^z &= \max_{B^*} \Pr_A[\Pi_z((A, B^*)) = v] \\
p_z &= \Pr_{A,B}[\Pi_z((A, B)) = v]
\end{aligned}$$

[6] Actually, the protocol of [GSV98] does not provide a multiplicative guarantee of 2^k, but rather ensures that the probability that the output lands in any set T of density μ is at most $2^k \cdot \mu + o(1)$. Our lower bound also applies to this more general type of guarantee.

That is, ϕ_A^z (resp. ϕ_B^z) is the highest probability Alice (resp. Bob) can make the the output to be v, given that the protocol is now at node z and that Bob (resp. Alice) is playing honestly. p_z is the probability that v is chosen starting from z and assuming both player play honestly.

The following lemma is the heart of the proof:

Lemma 5.6 *For every node z on T, $\phi_A^z \cdot \phi_B^z \geq p_z$.*

To prove the theorem from Lemma 5.6, take z to be the root of the tree. Then we have that $\phi_A \cdot \phi_B \geq p$, where ϕ_A (resp. ϕ_B) is the probability that Alice (resp. Bob) can force the output to be v, and p is the probability that v is chosen when both players play honestly. Notice that v was arbitrary, so we certainly can choose v such that $p \geq 1/N$. By definition, $\rho_A \geq \phi_B/(1/N) = \phi_B \cdot N$, and likewise $\rho_B \geq \phi_A \cdot N$. But then we have:

$$\rho_A \cdot \rho_B \geq (\phi_A \cdot N)(\phi_B \cdot N) \geq pN^2 \geq N$$

Proof of Lemma 5.6: We will prove the lemma by backwards induction on the levels of the tree.

When z is a leaf, the protocol is complete. If v is the output of the protocol at leaf z, then $\phi_A^z = \phi_B^z = p_z = 1$. Otherwise, $\phi_A^z = \phi_B^z = p_z = 0$. In either case, the lemma holds.

Now, suppose that the lemma holds for all children of z—denote them $z_1, \ldots z_m$. Thus, we know $\phi_A^{z_i}\phi_B^{z_i} \geq p_{z_i}$ for all children z_i. Suppose also, without loss of generality, that at node z it is Alice's turn.

Suppose an honest Alice chooses child node z_i with probability λ_i. Then $p_z = \sum \lambda_i p_{z_i}$, and $\phi_B^z = \sum \lambda_i \phi_B^{z_i}$. This latter equality holds because the probability v will be chosen when Alice is honest will just be the sum of the probabilities v will be chosen from each child, weighted by the probability of reaching that child. When considering ϕ_A^z, however, Alice will have the option of cheating. She will simply choose the child that affords her the best probability of successfully forcing the output to be v. That is, $\phi_A^z = \max_{z_i} \phi_A^{z_i}$, and so in particular for all i, $\phi_A^z \geq \phi_A^{z_i}$.

Now just compute:

$$\phi_A^z \phi_B^z = \phi_A^z \sum \lambda_i \phi_B^{z_i} \geq \sum \lambda_i \phi_A^{z_i} \phi_B^{z_i} \geq \sum \lambda_i p_{z_i} = p_z$$

This completes the proof of the lemma, and thus the theorem is proven.

■

■

What is the intuition for this result? It is best to try to understand the intuition for the lemma—that $\phi_A^z \cdot \phi_B^z \geq p_z$. From there it is only symbol manipulation to determine that $\rho_A \cdot \rho_B \geq N$.

To see this, suppose that there were only one path down the tree that led to v being chosen as the output. At each node along that path, starting from the root, there is a certain probability that the given player will choose the (unique) next node in the path when it plays honestly. So the probability that v is chosen is the product of these probabilities when both players play honestly. If Alice is cheating, then the probability v will be chosen is the product of the probabilities at nodes where it's Bob's turn. Likewise, the probability that v is chosen when Bob is cheating is the product of the probabilities at nodes where it's Alice's turn. In this case, the lemma holds with equality.

Intuitively, then, the probability v is chosen honestly is the product of the two cheating probabilities because choosing v honestly requires *both* Alice and Bob to happen to choose the right paths (i.e., to the leaf where v is selected), whereas choosing v with one player cheating requires only the honest player to happen to choose correctly (the cheating player will always choose the right path).

It remains to understand why multiple paths yielding v only help the cheaters. This makes sense, because it merely provides more options to the cheating player—if two routes exist that both could yield output v, the cheating player can now choose the more attractive option.

Note that, unlike Theorem 4.1, this result relies centrally on the assumption that, when one player is cheating, the other player is playing *honestly.* This assumption is certainly one of the key drivers of the result, as it then allows us to relate the probability of an element being chosen by both players being honest to the probability it is chosen when one cheats.

This lower bound is tight:

Proposition 5.7 *For all positive integers $N \geq K$, there exists a 2-round protocol for selecting from a universe of size N satisfying $\rho_A = K$ and $\rho_B = N/K$.*

Proof: Consider the following protocol $\Pi(A, B)$ over universe $\mathcal{U}$ of size N, which we call the Random Set Protocol with parameter K :

1. Alice chooses a random subset T of $\mathcal{U}$ of size K.
2. Bob chooses a random element $x \in T$.
3. The output is x.

Claim 5.8 $\rho_B = N/K$.

Proof of Claim: Using Lemma 5.1,

$$\rho_B = \max_{s \in \mathcal{U}} \max_{A^*} \left(N \cdot \Pr_B[\Pi((A^*, B)) = s] \right)$$

But for every $s \in \mathcal{U}$ and fixed A^*, since we assume Bob plays honestly, $\Pr_B[\Pi((A^*, B)) = s] \leq 1/K$. Conversely, considering the strategy A^* that sends any fixed set S to Bob, $\Pr_B[\Pi((A^*, B)) = s] = 1/K$ for any $s \in S$. □

Claim 5.9 $\rho_A = K$.

Proof of Claim: Again, by Lemma 5.1,

$$\rho_A = \max_{s \in \mathcal{U}} \max_{B^*} \left(N \cdot \Pr_A[\Pi((A, B^*)) = s] \right)$$

If Alice plays honestly, then in order for $\Pi((A, B^*)) = s$, she must select it, which occurs with probability K/N. Bob can achieve this probability for any $s \in U$ assuming he always selects s when available. □

■

Note that one negative aspect of the Random Set Protocol is that it is not efficient—sending a description of the random subset requires communication linear in N (rather than polylog(N)). It should be emphasized that this is *not* necessary to achieve $\rho_A \rho_B = N$: other very simple and efficient protocols achieve this tradeoff. Specifically, instead of using all sets of size K, we can use any subcollection such that every element of $[N]$ is contained in the same number of sets. For example, if $N = K \cdot L$ for an integer L, then we can view the universe as $[K] \times [L]$ and use only the sets of the form $S_a = [M] \times \{a\}$ for each $a \in [L]$, so the communication becomes $\log L + \log K = \log N$.

The optimality of such a trivial protocol tells us that, ultimately, multiplicative guarantees are not by themselves very interesting metrics of study for two-party random selection protocols. Optimal protocols are very easy to come by, and protocols that do not *feel* very effective prove to be the best possible. We must look to other metrics that are more capable of separating out protocols that are intuitively "good" from those that are not.

That said, it *is* interesting to consider multiplicative guarantees as one half of the equation: producing a protocol with an optimal tradeoff of statistical guarantee to one player and multiplicative guarantee to the other player is certainly nontrivial. Though we will not be able to match the lower bound of Proposition 5.3, in the next subsection we show that the Iterated Random Shift Protocol presented earlier achieves simultaneous *constant* statistical and multiplicative guarantees, and prove it has optimal round complexity up to a factor of $2 + o(1)$.

5.3 The Multiplicative Guarantees of the Iterated Random Shift Protocol

In this section we discuss the multiplicative guarantees provided by the Iterated Random Shift Protocol. Although we have seen how lower bounds require that one of the players (in this case, Alice) receives a very poor multiplicative guarantee, we can show that Bob receives a very strong guarantee. In this way, we can say something about the ability of a protocol to provide a strong multiplicative guarantee to one player, while providing a strong statistical guarantee to the other.

Theorem 4.3 implies a lower bound on the round complexity of protocols achieving simultaneous statistical and multiplicative guarantees:

Theorem 5.10 *For every two constants $\epsilon_A < 1$ and ρ_B, there exists a constant c such that any protocol Π selecting from a universe of size N and achieving statistical guarantee ϵ_A and multiplicative guarantee ρ_B will have at least $\log^* N - \log^* \log^* N - c$ rounds. Similarly for ρ_A and ϵ_B.*

This theorem follows immediately from Theorem 4.3 and the second part of Lemma 5.2.

We can show that the Iterated Random Shift Protocol achieves this:

Theorem 5.11 *There exist constants $\epsilon < 1$ and ρ such that the Iterated Random Shift Protocol with the cutoff parameter M taken to be a sufficiently large constant achieves guarantees $\rho_B \leq \rho$ and $\epsilon_A \leq \epsilon$.*

This is the first protocol achieving constant statistical and multiplicative guarantees we know of, and according to the lower bound of Theorem 5.10 it has optimal round complexity up to a factor of $2 + o(1)$.

Given Theorem 1.2, it suffices to show the following:

Proposition 5.12 *Let Π be a Iterated Random Shift Protocol defined with constant cutoff parameter M. Then Π provides a constant multiplicative guarantee to Bob: there exists constant ρ such that, as long as Bob plays honestly, the output of the Iterated Random Shift Protocol will fall in a set T with probability at most $M^2 \cdot \mu(T)$, for any set T.*

Proof of Proposition 5.12: Fix an arbitrary set $T \subseteq \mathcal{U}$. Recall the definition of the following variables in the proof of Theorem 1.2: $N_0, N_1, \ldots, N_{k^*}$ are the universe sizes in an execution of the Iterated Random Shift Protocol, k^* is the first value of k such that $N_k = M^2$, $m_k = \max\{M, \lceil \log^3 N_{k-1} \rceil\}$ is the parameter used in the k'th execution of the Random Shift Protocol, and we define:

$$\begin{array}{ll} \mathcal{U}_0 = \mathcal{U} & \mathcal{U}_k = [m_k] \times [m_k] \\ T_0 = T & T_k = \{(i,j) \in \mathcal{U}_k : (a_i + b_j) \in T_{k-1}, 1 \leq i, j \leq m_k\} \end{array}$$

$$\mu(T_k) = |T_k|/|\mathcal{U}_k|.$$

The following is the key lemma:

Lemma 5.13 *Assuming Bob plays honestly, $E[\mu(T_k)] = E[\mu(T_{k-1})]$ for all $k = 1, \ldots, k^*$.*

Proof: Consider the Random Shift Protocol. Let $a_1, \ldots, a_m$ be given. Then if $b_1, \ldots, b_m$ are chosen uniformly at random, it follows that for each i, j, the element $a_i + b_j$ is uniform over $\mathcal{U}$ (since $+$ is a group operation), and thus $\Pr[a_i + b_j \in T] = \mu(T)$. By linearity of expectations, we can conclude that $E[\#(a_i + b_j) \in T] = \mu(T) \cdot m^2$ (where m^2 is the size of the new universe), and thus $E[\mu(T')] = \mu(T)$, where $\mu(T')$ is the residual density of T in the resulting universe.

Applying this logic within the Iterated Random Shift Protocol, the lemma is proven (since for given $\mu(T_{k-1})$, we know $E[\mu(T_k)] = \mu(T_{k-1})$).

■

By induction, we then have that for all k, $E[\mu(T_k)] = \mu(T)$. In particular, this is true for $k = k^*$. We can then derive:

$$\begin{aligned} \mu(T) &= E[\mu(T_{k^*})] \\ &\geq (1/M^2)] \cdot E[|T_{k^*}|] \\ &\geq (1/M^2) \cdot \Pr[|T_{k^*}| > 0] \end{aligned}$$

Since if $|T_{k^*}| = 0$, the protocol's output cannot possibly fall in T, we conclude that the probability the output falls in T is at most $M^2 \cdot \mu(T)$. This proves the proposition, and thus Theorem 5.11. ■

As an aside, notice that by using Lemma 5.2, this logic allows us to conclude one half of Theorem 1.2: the Iterated Random Shift Protocol provides a constant statistical guarantee to Bob.

We can conclude that the Iterated Random Shift Protocol has the following properties:

- It has only $2\log^* N + O(1)$ rounds.
- It provides both Alice and Bob with constant statistical guarantees (equivalently, it satisfies the Statistical Criterion).

- It provides Bob with a constant multiplicative guarantee.

Notice that in the above proof, we never used the multiplicative guarantee properties of the GGL Protocol–we simply relied on the initial recursions of Random Shift to provide the strong guarantee to Bob.

In fact, by changing the protocol used when the universe size becomes of size M^2 in the definition of the Iterated Random Shift Protocol, we can improve even further the multiplicative guarantee given to Bob. The current protocol only implies that Bob gets *some* constant multiplicative guarantee. By using the Random Set Protocol on the universe of size M^2 with parameter $K = M^2/(1+\gamma)$ instead of GGL, however, (see definition in Proposition 5.7), Bob can achieve a multiplicative guarantee $1+\gamma$, while still keeping Alice's statistical guarantee constant (when γ is constant).

Acknowledgements

We thank Oded Goldreich, Grant Schoenebeck, and Alexander Healy for several helpful discussions.

References

[AN90] N. Alon and M. Naor. Coin-flipping games immune against linear-sized coalitions. In *Proc. 31st FOCS*, 1990.

[BL89] M. Ben-Or and N. Linial. Collective coin-flipping.*Randomness and Computation*, S. Micali ed., Academic Press, New York, 1989.

[Blu82] M. Blum. Coin flipping by telephone. In *IEEE Spring COMPCOM*, 1982.

[BN98] R. Boppana and B. Narayanan. Perfect information leader election with optimal resilience. *SIAM J. Computing* 29(4), 1998.

[CCM98] C. Cachin, C. Crepeau, and J. Marcil. Oblivious transfer with a memory-bounded receiver. In *Proc. 39th FOCS*, 1998.

[Dam94] I. Damgard. Interactive hashing can simplify zero-knowledge protocol design. In *Proc. CRYPTO '95*, Springer LNCS 403, 1994.

[DGW94] I. Damgard, O. Goldreich, and A. Wigderson. Hashing functions can simplify zero-knowledge protocol design (too). TR RS-94-39. BRICS, 1994.

[DHRS04] Y. Ding, D. Harnik, A. Rosen, and R. Shaltiel. Constant-round oblivious transfer in the bounded storage model. In *Proc. 1st TCC*, Springer LNCS 2951, 2004.

[Fei99] U. Feige. Noncryptographic selection protocols. In *Proc. 40th FOCS*, 1999.

[GGL98] O. Goldreich and S. Goldwasser and N. Linial. Fault-tolerant computation in the full information model. *SIAM J. Computing* 27(2), 1998.

[GMW91] O. Goldreich, S. Micali, and A. Wigderson. Proofs that yield nothing but their validity or all languages in **NP** have zero-knowledge proof systems. *J. ACM* 38(1), 1991.

[GSV98] O. Goldreich, A. Sahai, and S. Vadhan. Honest-verifier statistical zero-knowledge equals general statistical zero-knowledge. In *Proc. 30th STOC*, 1998.

[KO04] J. Katz and R. Ostrovsky. Round-optimal secure two-party computation. In *Proc. CRYPTO '04*. Springer LNCS 3152, 2004.

[Lau83] C. Lautemann. **BPP** and the polynomial hierarchy. *IPL* 14, 1983.

[Lin01] Y. Lindell. Parallel coin-tossing and constant-round secure two-party computation. In *Proc. CRYPTO '01*. Springer LNCS, 2001.

[NOVY98] M. Naor, R. Ostrovsky, R. Venkatesan, and M. Yung. Perfect zero-knowledge arguments for **NP** can be based on general complexity assumptions. *J. Cryptology* 11, 1998.

[ORV94] R. Ostrovsky, S. Rajagopalan, U. Vazirani. Simple and efficient leader election in the full information model. In *Proc. 26th STOC*, 1994.

[RSZ99] A. Russell, M. Saks, and D. Zuckerman. Lower bounds for leader election and collective coin-flipping in the perfect information model. In *Proc. 31st STOC*, 1999.

[RZ98] A. Russell and D. Zuckerman. Perfect information leader election in $\log^* n + O(1)$ rounds. In *Proc. 39th FOCS*, 1998.

[Sak89] M. Saks. A robust noncryptographic protocol for collective coin-flipping. *SIAM J. Discrete Math*, 1989.

[San04] S. Sanghvi. A study of two-party random selection protocols. Undergraduate Thesis. Harvard University, 2004.

[San05] S. Sanghvi. The round complexity of two-party random selection. PowerPoint presentation given at STOC 2005. Available at http://eecs.harvard.edu/ salil/papers/randsel-abs.html.

[SV05] S. Sanghvi and S. Vadhan. The round complexity of two-party random selection. In *Proc. 37th STOC*, 2005.

[Yao86] A. Yao. How to generate and exchange secrets. In *Proc. 27th FOCS*, 1986.

A Assorted Proofs

We include in this appendix the proofs of two lemmas required for the main lower bound.

A.1 Growth of $f(r, \epsilon, \mu)$

Lemma 4.19 *Recall the definition $f(r, \epsilon, \mu) = g(r, r, \epsilon, \mu)$, where we define*

$$\begin{aligned} g(0, r, \epsilon, \mu) &= 1 \\ g(k, r, \epsilon, \mu) &= (r/\epsilon)(re/\mu)^{g(k-1, r, \epsilon, \mu)} g(k-1, r, \epsilon, \mu) \end{aligned}$$

There exists a function $\zeta(\epsilon, \mu)$ such that, when $r < \log^* N - \log^* \log^* N - \zeta(\epsilon, \mu)$, we have $f(r, \epsilon, \mu) \leq \log N$.

Proof: First, bound r by $\log^* N$. Again, for shorthand we will write s_k for $g(k, r, \epsilon, \mu)$. Thus, we have that

$$s_k = (r/\epsilon)(re/\mu)^{s_{k-1}} s_{k-1}$$

Notice that this is no more than $(r^2 e/(\epsilon\mu))^{s_{k-1}}$. ($xy \leq x^y$ if $x, y \geq 2$.) Letting $d = r^2 e/(\epsilon\mu)$, we can then bound s_k by t_k, where t_k is defined by $t_0 = 1$ and $t_k = d^{t_{k-1}}$.

This means that we can set $k = \log_d^* N - 1$ and still have $s_k \leq t_k \leq \log N$ (recall that by our definition, $\log_b^* N$ is always an integer, for any b or N). It only remains to relate this to a base 2 logarithm:

Claim A.1 *If $d \geq 4$, $\log_d^* N \geq \log^* N - \log^*(2 \log d)$.*

Proof: Recall $\log^{(k)} N$ is k iterated logarithms of N. We claim the following:

Claim A.2 *For $k \leq \log_d^* N$, $d \geq 4$, $\log^{(k)} N \leq (2 \log d) \log_d^{(k)} N$.*

Proof: The base case $k = 0$ is clear. Assume, then, that $\log^{(k-1)} N \leq (2 \log d) \log_d^{(k-1)} N$. Applying log to both sides, we have that:

$$\begin{aligned}
\log^{(k)} N &\leq \log(2 \log d) + \log(\log_d^{(k-1)} N) \\
&= \log(2 \log d) + (\log_d^{(k)} N)(\log d) \\
&\leq (2 \log d)(\log_d^{(k)} N)
\end{aligned}$$

where the last line follows because for $d \geq 4$, $d \geq 2 \log d$ and for $k \leq \log_d^* N$, $\log_d^{(k)} N \geq 1$. ■

Plugging in $k = \log_d^* N$, then we have that $\log^{(\log_d^* N)} N \leq 2 \log d$. Applying $\log^*(2 \log d)$ logarithms to both sides, we have $\log^{(\log_d^* N + \log^*(2 \log d))} N \leq 1$. Since $\log^* N$ is defined to be the least k such that $\log^{(k)} N \leq 1$, it follows that $\log^* N \leq \log_d^* N + \log^*(2 \log d)$. ■

Thus we have that we can set k to be at least $\log_d^* N - 1$ and s_k will be no more than $\log N$. Moreover,

$$\begin{aligned}
\log_d^* N - 1 &\geq \log \log^* N - \log^*(2 \log d) - 1 \\
&= \log^* N - \log^*(2 \log((r^2 e)/\epsilon\mu)) - 1 \\
&\geq \log^* N - \log^* \log^* N - \zeta(\epsilon, \mu)
\end{aligned}$$

for an appropriately chosen constant ζ, since we can bound r by $\log^* N$.

■

A.2 Encompassing Small Sets with a Random Set

Lemma 4.11 *Suppose we have a collection of disjoint sets $S_1, \ldots S_m$ over a fixed universe $\mathcal{U}$, $|\mathcal{U}| = n$, where for all i, $|S_i| \le s$. Choose a set R randomly of density μ (i.e., $n \cdot \mu$ elements). With probability $\ge 1 - (1/m)(e/\mu)^s$, there will exist S_i such that $S_i \subseteq R$.*

Proof: The proof is an application of Chebyshev's inequality.

Define random variable X_i to be 1 if $S_i \subseteq R$, 0 otherwise, and set $X = \sum X_i$. We are interested in upper-bounding the probability that $X = 0$. By Chebyshev's inequality, we have:

$$\begin{aligned}
\Pr[X = 0] &\le \Pr[|X - E[X]| \ge E[X]] \\
&\le \mathrm{Var}[X]/E[X]^2 \\
&= \left(\sum_i \mathrm{Var}[X_i] + \sum_{i \ne j} \mathrm{Cov}[X_i, X_j] \right) /E[X]^2
\end{aligned}$$

Claim A.3 $\forall i \ne j$, *Cov*$[X_i, X_j] \le 0$

Intuitively, the fact that $S_i \subseteq R$ makes it less likely that $S_j \subseteq R$.

Proof of Claim:

More formally, let $|S_i| = p$, $|S_j| = q$, $|R| = r$ and assume without loss of generality that $p \le q$. Since S_i and S_j are disjoint, $E[X_i] = \binom{n}{p}/\binom{r}{p}$ (sim. for X_j), and $E[X_i X_j] = \binom{n}{p+q}/\binom{r}{p+q}$. Then:

$$\begin{aligned}
\frac{E[X_i X_j]}{E[X_i]E[X_j]} &= \frac{\binom{r}{p+q}\binom{n}{p}\binom{n}{q}}{\binom{n}{p+q}\binom{r}{p}\binom{r}{q}} \\
&= \left(\frac{r!}{(p+q)!(r-p-q)!} \frac{n!}{p!(n-p)!} \frac{n!}{q!(n-q)!} \right) \cdot \\
&\quad \left(\frac{(p+q)!(n-p-q)!}{n!} \frac{p!(r-p)!}{r!} \frac{q!(r-q)!}{r!} \right) \\
&= \frac{n!(n-p-q)!(r-p)!(r-q)!}{r!(r-p-q)!(n-p)!(n-q)!} \\
&= \left(\frac{n(n-1)\ldots(n-p+1)}{r(r-1)\ldots(r-p+1)} \right) \cdot \\
&\quad \left(\frac{(r-q)(r-q-1)\ldots(r-p-q+1)}{(n-q)(n-q-1)\ldots(n-p-q+1)} \right) \\
&\le \left(\frac{n-p}{r-p} \right)^p \left(\frac{r-q}{n-q} \right)^p \\
&\le \left(\frac{n-p}{r-p} \right)^p \left(\frac{r-p}{n-p} \right)^p \\
&= 1
\end{aligned}$$

Since $\mathrm{Cov}[X_i, X_j] = E[X_i X_j] - E[X_i]E[X_j]$, the claim follows.

$\square$

We can thus remove the covariance term from the upper bound:

$$\begin{aligned}
\Pr[X=0] &\leq \left(\sum_i \mathrm{Var}[X_i]\right) / E[X]^2 \\
&\leq \left(\sum_i E[X_i]\right) / E[X]^2 \\
&= 1/\left(\sum_i E[X_i]\right) \\
&\leq \frac{1}{m} \cdot \frac{\binom{n}{s}}{\binom{r}{s}} \\
&\leq \frac{1}{m} \cdot \left(\frac{ne}{r}\right)^s \\
&= \frac{1}{m} \cdot \left(\frac{e}{\mu}\right)^s
\end{aligned}$$

■