Fault-tolerant Computation in the Full Information Model*

Oded Goldreich† Shafi Goldwasser‡ Nathan Linial§

October 18, 1995

Abstract

We initiate an investigation of general fault-tolerant distributed computation in the *full-information* model. In the full information model no restrictions are made on the computational power of the faulty parties or the information available to them. (Namely, the faulty players may be infinitely powerful and there are no private channels connecting pairs of honest players).

Previous work, in this model, has concentrated on the particular problem of simulating a single bounded-bias global coin flip (e.g. Ben-Or and Linial [4] and Alon and Naor [1]). We widen the scope of investigation to the general question of how well arbitrary fault-tolerant computations can be performed in this model. The results we obtain should be considered as first steps in this direction.

We present efficient two-party protocols for fault-tolerant computation of any bivariate function. We prove that the advantage of dishonest player in these protocols is the minimum one possible (up to polylogarithmic factors).

We also present efficient m-party fault-tolerant protocols for sampling a general distribution ($m \geq 2$). Such an algorithm seems an important building block towards the design of efficient multi-party protocols for fault-tolerant computation of multivariate functions.

*Supported by grants No. 89-00312, 89-00126 and 92-00226 from the United States - Israel Binational Science Foundation (BSF), Jerusalem, Israel. An extended abstract of this work has appeared in the proceedings of the *32nd FOCS*, pp. 447–457, 1991.

†Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel.

‡Laboratory for Computer Science, MIT, Cambridge, USA.

§Institute of Computer Science, Hebrew University, Jerusalem, Israel.

1 Introduction

The problem of how to perform general distributed computation in an unreliable environment has been extensively addressed. Two types of models have been considered. The first model assumes that one-way functions exist and considers adversaries (faults) which are computationally restricted to probabilistic polynomial time [24, 13, 25, 14, 11, 2]. The second model postulates that private channels exist between every pair of players [3, 7, 8, 17, 15]. Hence, in both models fault-tolerance is achieved at the cost of restricting the type of faults.

We want to avoid any such assumption and examine the problem of fault-tolerant distributed computation where the faults are computationally unrestricted, and no private channels are available. Clearly the assumption that one-way functions exist is of no use here. The situation here corresponds to games of complete information.

The general problem can be described informally as follows: m players are interested in globally computing $v = f(x_1, ..., x_m)$ where f is a predetermined m-variate function and x_i is an input given to party i (and initially known only to it). The input x_i is assumed to have been drawn from probability distribution D_i (which without loss of generality can be assumed to be uniform). A coalition F of faulty players may favor a particular value v for f and play any strategy to maximize the probability of such an outcome. We want to bound, for each value v in the range of f, the probability (under the best strategy for the faults) that the outcome of the protocol used to distributively compute f is v. How good can this bound be?

Regardless of the protocol under consideration, there is always one avenue that is open for the faulty players, namely, alter their input values to ones under which the value v is most likely. This is always possible, since players' inputs are not visible to others. That is,

$$q_v := \max_{x_i, i \in F} \{\text{Prob}(f(\vec{x}) = v \text{ where } x_j \in_R D_j, j \notin F)\}$$

is a lower bound on the influence of coalition F towards value v, no matter what protocol is used.

Consider the simple procedure in which each player announces its x_i, and the global output is taken to be $f(x_1, ..., x_m)$. If all players (including the faulty ones) act simultaneously, then for every v, the probability of v being the outcome is indeed at most q_v. Unfortunately, in a distributed network simultaneity cannot be guaranteed, and a delayed action by the faults can result in much better performance for them (e.g., for $f = \sum_{i=1}^{m} x_i \bmod N$ with $x_i \in \{0, 1, ..., N-1\}$, $q_0 = \frac{1}{N}$, but a single faulty player acting last has complete control of the outcome).

In both of the previously studied models (private channels or computationally bounded faults) protocols were developed where for all values v and all *minority* coalitions F, the probability of outcome v is as close to q_v as desired. The key to these protocols is the notion of *simultaneous commitment*. At the outset of these protocols, each player P_i commits to its input x_i. It should be stressed that a faulty party may alter its input in this "committing phase" but not later and that a party's commitment is "independent" of the inputs of the other honest parties.

Obviously, in the full-information model such a qualitative notion of commitment cannot be implemented (even if the faulty parties are in minority). Instead, we need to look for *quantitative* results. Faulty players can and will be able to "alter their inputs" throughout the execution of the protocol in order to influence the outcome. Yet, we can bound the advantage gained by their improper behavior.

Results Concerning the Two-Party Case

The main focus of this paper is on the two-player case of this problem. Even this restricted case provides interesting problems and challenges. We resolve the main problems in this case, showing:

1. A lower bound: for every bivariate function f, for any protocol to compute f and every value v in the range of f, there is a strategy for one of the players, so that if the other player plays honestly, then the probability for the outcome $f = v$ is at least $\max(q_v, \sqrt{p_v})$, where $p_v = \text{Prob}(f(\vec{x}) = v | x_i \in_R D_i))$.

2. More interestingly, we show a matching (up to polylogarithmic factor) constructive upper bound. We describe a probabilistic polynomial time protocol that computes f, given a single oracle access to f, such that for all v,

$$Pr(f \text{ evaluates to } v) = O(\text{poly}\log(1/p_v) \cdot \max(q_v, \sqrt{p_v}))$$

   In the special case where $q_v = p_v$, this protocol is shown to match the lower bound up to a constant factor. Namely,

$$Pr(f \text{ evaluates to } v) = O(\sqrt{p_v})$$

The spirit of our protocol is best illustrated by the following example.
Example: Define $id(x, y) = 1$ if $x = y$ and 0 otherwise. Suppose that the local inputs x, y are chosen uniformly in $\{0,1\}^n$. Clearly, $p_1 = \frac{1}{N}$, and $p_0 = 1 - \frac{1}{N}$, where $N = 2^n$. A protocol in which the first player declares x and then the second player declares y allows the second player complete control on the value of id. A protocol in which the two players alternatingly exchange bits in the description of their inputs is no better if these bits are exchanged in the same order (i.e., both parties send their respective i^{th} bit in round i). A much better idea is for the two players to alternate in describing the bits of their inputs but do so from opposite directions (i.e., in round i the first party sends its i^{th} bit whereas the second party sends its $(n - i + 1)^{\text{st}}$ bit). Clearly, whichever player is faulty, the probability that the outcome of this protocol is "1" is bounded by $\frac{1}{\sqrt{N}}$. In light of the lower bound, this is the best result possible. This idea of gradually revealing appropriately chosen "bits of information" is the key to the general problem of two-party computation.

Results Concerning the Multi-Party Case

The problem of m-party computations, where a subset of $t < m$ faults may exist, is more involved than the two-party case (even for $m = 3$); see discussion in Section 5. Here, we only consider the problem of collectively sampling a given distribution. Without loss of generality, it suffices to consider the uniform distribution (say, on strings in $\{0,1\}^l$). We provide a probabilistic polynomial time sampling protocol such that for every $S \subset \{0,1\}^l$, for every t faults,

$$\Pr(\text{sample } \in S) < \left(\frac{|S|}{2^l}\right)^{1 - c \cdot \frac{t}{m}}$$

for some constant $c > 0$. This result is the best possible (up to the constant c), and is superior to the bound obtained by the trivial protocol which consists of l repeated applications of "collective

coin flipping"; consider, for example, the set S consisting of all strings having at least $(\frac{1}{2}+\frac{t}{m})\cdot l$ ones – under the trivial protocol, t faulty parties can influence the output to almost always hit S, whereas our result guarantees that this set S which forms a negligible fraction of $\{0,1\}^l$ is hit with negligible probability (for, say, $t < m/2c$).[1]
The above sampling protocol can be used to present a (generic probabilistic polynomial-time) protocol that works well for computing *almost all* functions (see our technical report [12]).

Previous Work in the Full Information Model

Collective coin flipping, i.e., common bounded-biased sampling in $\{0,1\}$ has been considered in this full-information model before [4, 5, 16, 1]. Matching lower and (constructive) upper bounds of $\frac{1}{2}+\theta(\frac{t}{m})$ have been shown (by Ben-Or and Linial [4] and Alon and Naor [1],[2] respectively). Our work can be viewed as an extension of these investigations which were concerned with the influences of players on *Boolean* functions (i.e., $Range(f)=\{0,1\}$). The general case, considered in this paper, gives rise to additional difficulties. Let us stress that even the problem of sampling in arbitrary sets is more difficult than collective coin flipping. As mentioned above, the obvious approach to the sampling problem fails; namely, a sampling protocol that consists of repeatedly applying a given coin tossing protocol can be easily influenced to almost always output strings in a subset of negligible size.[3].

However, fault tolerant computation (of arbitrary functions) is more complex than sampling which can be viewed as fault-tolerant computation of a function specially designed for this purpose.

Relation to Work on Slightly-Random Sources

In this paper we present a multi-party protocol for sampling a set of strings $\{0,1\}^l$. In "sampling" we mean producing a single string in $\{0,1\}^l$ so that, for every subset $S\subset\{0,1\}^l$, the probability that the sample hits S is related to the density of S. Our protocol uses the collective coin flipping of [1] as a subroutine. In fact, our sampling protocol can be viewed as a deterministic reduction to the problem of collective coin tossing. The collective coin can be viewed as a slightly random source in the sense of Santha and Vazirani [22], i.e., an *SV-source*[4]. Hence, our result can be interpreted as presenting a sampling algorithm which uses a SV-source (with a parameter $\gamma<\frac{1}{\sqrt{2}}$). Our sampling algorithm performs much better than the obvious algorithm which uses

[1] Using the above choice of parameters, we have a set S of density $\rho\approx\exp\{-(t/m)^2\cdot l\}$ which our protocol hits with probability at most $\sqrt{\rho}$, as long as at most t players are faulty. On the other hand, when repeated collective-coin-flippings are used, t faulty players can influence the outcome to be in S with probability at least $1-\rho$, by biasing each coin-flip towards 1.

[2] Furthermore, the upper bound can be met by protocols of logarithmic round-complexity [9, 19].

[3] An alternative method which also fails is to try to generalize the work of Alon and Naor [1] as follows: the method of [1] consists of randomly selecting one of the players who is appointed to flip a fair coin. Letting this player select a random string is a natural idea, but it is obvious that this approach performs very poorly for a sample space of non-constant size. Specifically, each set $S\subset\{0,1\}^l$ can be hit with probability at least $\frac{1}{2}+\frac{t}{m}$, independently of S and l.

[4] An SV-source with parameter γ is a sequence of binary random variables $X_1, X_2, \ldots$, so that for every n, $\alpha\in\{0,1\}^n$ and $\sigma\in\{0,1\}$, $\mathrm{Prob}(X_{n+1}=\sigma|X_1\cdots X_n=\alpha)\leq\gamma$.

as a sample a sequence of coins produced by the source. (The situation is analogous to the discussion of the multi-party sampling protocols above).

Our sampling algorithm provides an alternative way of recognizing languages in BPP by polynomial-time algorithms which use a SV-source with a parameter $\gamma < \frac{1}{\sqrt{2}}$. First, reduce the error probability in the BPP-algorithm so that it is negligible (i.e. smaller than any polynomial fraction). Next, use our sampling algorithm to produce a sequence of coin tosses for a *single run* of the new BPP-algorithm. Since the "bad runs" form a negligible fraction of all possibles runs of the BPP-algorithm, it follows that the probability that we will sample a bad run (when using a SV-source with parameter $\gamma < \frac{1}{\sqrt{2}}$) is also negligible. This simulation method is different from the original method of Vazirani and Vazirani [23] (adopted also in [6]) where the BPP-algorithm is invoked *many* times, each time with a different sequence of coin tosses.

Other Related Work

We also present efficient sampling protocols for the two-party case. The basic sampling protocol guarantees, for every set $S \subseteq \{0,1\}^l$, that as long as one party is honest the output hits S with probability at most $O(\sqrt[4]{|S|/2^l})$. (The basic sampling protocol is essential for efficiently implementing our generic two-party function-computation protocol. Interestingly, it the basic sampling protocol is also used as a building block for a better sampling protocol, which is optimal up to a constant factor.)

Our basic two-party sampling protocol is very similar to a protocol, called interactive hashing, which was discovered independently by Ostrovsky et. al. [20]. Interactive hashing has found many applications in cryptography (cf. [20, 18, 21, 10]). For details see Remark 2.

2 Preliminaries

2.1 Bivariate Functions

Throughout the paper we represent the bivariate function $f:\{0,1\}^n\times\{0,1\}^n \mapsto \{0,1\}^*$ as an N-by-N matrix, where $N \stackrel{\text{def}}{=} 2^n$. An entry, (x,y), in the matrix which has value v (i.e., $f(x,y)=v$) is called a *v-entry*. The following quantities, related to the function f and a value v in its range, are central to our analysis.

Notation: The *density of* v, denoted p_v, is the fraction of v-entries in the matrix of f (i.e., $p_v = |\{(x,y):f(x,y)=v\}|/2^{2n}$). The *maximum row density of* v, denoted r_v, is the maximum, taken over all rows, of the fraction of v-entries in a row of f (i.e., $r_v = \max_{x\in\{0,1\}^n}\{|\{y:f(x,y)=v\}|/2^n\}$). The *maximum column density of* v is denoted $c_v = \max_{y\in\{0,1\}^n}\{|\{x:f(x,y)=v\}|/2^n\}$, and q_v is defined as $\max\{r_v, c_v\}$.

Throughout the paper, we consider the case of uniform input distribution. Namely, we assume that each input is selected uniformly from $\{0,1\}^n$ and independently of the other input(s). The more general case, where each input is selected from an arbitrary distribution (yet independently of the other inputs) can be reduced to the uniform case as follows. Suppose that the probability for each input can be expressed as $\frac{q}{2^{\mathrm{poly}(n)}}$, where q is an integer (for some polynomial poly). Then we can replace this input, say z, by q inputs, denoted $(z,1),(z,2),...,(z,q)$, and consider the function $F((x,i),(y,j)) \stackrel{\text{def}}{=} f(x,y)$ $(1 \le i \le \phi(x)2^{\mathrm{poly}(n)}$ and $1 \le j \le \psi(y)2^{\mathrm{poly}(n)}$, where $\phi(x)$ is the probability of the row-input x and $\psi(y)$ is the probability of the column-input y). Protocols for computing F (under the uniform distribution) translate easily to protocols for computing f (under the distribution (ϕ, ψ)) and vice versa. To efficiently transform protocols for computing F into protocols for computing f, an efficient algorithm is needed for computing the original density functions (i.e., ϕ and ψ).

2.2 Protocols

The communication model consists of a single broadcast channel. Each party can, at any time, place a message on this channel which arrives immediately (bearing the identity of its originator) to all other parties. It is not possible to impose "simultaneity" on the channel; namely, the protocols may not contain a mechanism ensuring simultaneous transmission of messages by different parties. Thus, it is best to think of the model as being asynchronous and of the protocols as being message-driven. However, asynchronicity is not a major issue here as all parties share the unique communication medium and thus have the same view.

The output of an execution of a protocol is defined as the last message sent during the execution. We consider the output of the protocol when the inputs are selected uniformly.

We call a player *honest* if it follows the protocol. *Dishonest* players may deviate arbitrarily from the protocol. In discussing our protocols we assume, without loss of generality, that dishonest players do not deviate from the protocol in a manner which may be detected. This assumption can be easily removed by augmenting our protocols with simple detection and recovery procedures (which determine the output of the protocol in case deviation from the protocol is detected). For example, the protocol may be restarted with the input of the cheating party fixed to some predetermined value and all its actions being simulated by the other parties.

All our protocols are *generic*: Players are instructed to take steps that depend only on their inputs, but not on the function f. When the inputs are finally revealed, f is evaluated once, and the protocol terminates.

2.3 Influences

Unlike previous work, we use the term "influence" in a colloquial manner. Typically, by talking "the influence of a party towards a value" we mean the probability that this party can make this value appear as output of the protocol. When discussing the computation of functions, we treat only the influence towards a single value; the influence towards a set of values can be treated by defining a corresponding indicator function.

2.4 Sampling

We also consider the problem of designing two-party and multi-party protocols for sampling in a universe $\{0,1\}^l$. The objective here is to provide upper bounds for the probability that the output falls in some subset $S \subset \{0,1\}^l$. We note that the problem of designing a two-party protocol for sampling $\{0,1\}^l$ can be reduced to the problem of designing a protocol for computing any function $f : \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}^l$ for which all values have the same density and this density equals the maximum row/column densities (i.e., $q_v = p_v = 2^{-l}$ for every $v \in \{0,1\}^l$). An analogous reduction holds also in the multi-party case.

3 Lower Bounds

Theorem 1 : *Let* $f : D_1 \times D_2 \times \ldots \times D_m \mapsto R$ *be a function of* m *variables,* Π *an* m*-party protocol for computing* f*, and* $v \in R$ *a value in the range of* f*. Consider performing* Π *where players in the set* S *are dishonest, while all other players are honest. Let* ϕ_S *be the maximum, over all strategies of coalition* S *of the probability of the outcome being* v*. Then, for any* $1 \le t \le m$ *there is a coalition* Q *of* t *players with* $\phi_Q \ge p_v^{1-\frac{t}{m}}$*.*

In particular,

Corollary 2 : *Let* f *be any bivariate function,* Π *any two-party protocol for computing* f*, and* v *a value in the range of* f*. Then at least one of the players can, by playing (possibly) dishonestly force the outcome to be* v *with probability at least* $\max\{q_v, \sqrt{p_v}\}$ *(the other party plays honestly).*

Proof of Theorem 1: The proof is very similar to that of Theorem 5 in [4], though some changes are required. One observes first that if the time complexity of the protocol is no issue, and the only consideration is to keep influences down, then nothing is lost if all actions are taken sequentially and not in parallel. Therefore, Π can be encoded by a tree T as follows: Leaves of T are marked with values in the range of f, and each internal node of T is marked with a name of a player. The run of Π starts at the root of T. Whenever an internal node is reached, player P_i, whose name marks that node is to take the next step. For each input value in D_i, the protocol Π specifies a probability distribution according to which the next node, a child of the present one, is selected (assuming P_i is honest). The key observation, beyond the technique of [4], is that these distributions (together with the input distribution over D_i) induce a single distribution for the next move of (honest) player i, conditioned on the execution having reached the present node. The outcome of this process is determined by the leaf it reaches (i.e., $f = u$, where u is the mark of the leaf that is reached).

For the analysis, let z be an internal node of T, and consider the same process as above, performed on the subtree of T rooted at z. Suppose that coalition S plays its best strategy to make the outcome $f = v$ most likely, and let $\phi_S^{<z>}$ be that maximum probability (clearly, when z is taken to be the root of T, then $\phi_S^{<z>} = \phi_S$). The key step in the proof is to establish the following inequality for every internal z:

$$\prod_{|R|=t} \phi_R^{<z>} \ge p_{v,z}^{\binom{m-1}{t}} \tag{1}$$

where $p_{v,z}$ is the probability of reaching a v-marked leaf on that subtree, when all players are honest. Extracting the $\binom{m}{t}$-th root of the above inequality, we get $\max_R \phi_R^{<z>} \ge p_{v,z}^{(m-t)/m}$. Taking z to be the root of T the theorem follows.

Inequality (1) is proven by induction on the distance from the leaves in T. In the induction step, we assume that the inequality holds for the children of an internal node z and derive the inequality for node z. Let I denote the set of edges emanating from z and let $\{z_i : i \in I\}$ denote the corresponding children. Suppose, without loss of generality that node z is marked by player 1. The protocol Π and the probability distributions on the sets D_i determine the probabilities, $\{\lambda_i > 0 : i \in I\}$, governing the player's next move provided that the player is honest and conditioned on the execution having reached node z. (This distribution may not be easy to determine, but we only need to know that it exists.) Now, clearly $p_{v,z} = \sum_{i\in I} \lambda_i p_{v,z_i}$ and $\phi_R^{<z>} = \sum_{i\in I} \lambda_i \phi_R^{<z_i>}$, for every coalition R that does not contain player 1. On the other hand, for every coalition R which does contain player 1, we have $\phi_R^{<z>} = \max_{i\in I} \phi_R^{<z_i>}$, Now, denoting $\phi_R^{<z_i>}$ by $a_{i,R}$ (where $R \subseteq [m]$, $|R| = t$) and p_{v,z_i} by b_i, the inductive step reduces to proving the following numerical lemma, which in turn is a generalization of Lemma 5.3 in [4].

Lemma 3 *Let I be a finite set, let $\{a_{i,R} : i \in I, R \subseteq [m], |R| = t\}$, $\{b_i : i \in I\}$ be nonnegative reals, let $\{\lambda_i : i \in I\}$ be positive with $\sum_{i\in I} \lambda_i = 1$, and assume that for every $i \in I$,*

$$\prod_{R\subseteq[m],|R|=t} a_{i,R} \geq b_i^{\binom{m-1}{t}}.$$

Furthermore, let α_R equal $\max_{i\in I} a_{i,R}$ if $1 \in R$ and $\sum_{i\in I} \lambda_i a_{i,R}$ otherwise. Also, let $\beta = \sum_I \lambda_i b_i$. Then,

$$\prod_{R\subseteq[m],|R|=t} \alpha_R \geq \beta^{\binom{m-1}{t}}.$$

Lemma 5.3 in [4] is a special case of Lemma 3 (in which $|I| = 2$ and $\lambda_1 = \lambda_2 = \frac{1}{2}$). However, the ideas presented in the proof of Lemma 5.3 in [4] suffice for proving the general case. In fact, we further generalize Lemma 3 –

Lemma 4 *Let J, K and I be disjoint finite sets, let $\{a_{i,j} | i \in I, j \in J \cup K\}$, $\{b_i | i \in I\}$ be nonnegative reals, let $\{\lambda_i | i \in I\}$ be positive, with $\sum_{i\in I} \lambda_i = 1$, and assume that for every $i \in I$,*

$$\prod_{j\in J\cup K} a_{i,j} \geq b_i^{|K|}$$

For every $j \in J$, let α_j equal $\max_{i\in I} a_{i,j}$ and for every $k \in K$, let $\alpha_k = \sum_{i\in I} \lambda_i a_{i,k}$. Also $\beta = \sum_I \lambda_i b_i$. Then,

$$\prod_{j\in J\cup K} \alpha_j \geq \beta^{|K|}$$

Lemma 3 follows from Lemma 4 by letting J be the set of all t-subsets of $[m]$ which contain the element 1 and K be the set of all t-subsets which do not contain 1.

Proof of Lemma 4: There is, of course, no loss in assuming

$$b_i = (\prod_{j\in J\cup K} a_{i,j})^{1/|K|}$$

for every $i \in I$. Fix all $a_{i,j}$ (over all $i \in I, j \in J$) as well as all $a_{i,k}$ (all $i \in I, k \in K \setminus \{k_1, k_2\}$). Now consider the minimum of $(\sum_{i\in I} \lambda_i a_{i,k_1})(\sum_{i\in I} \lambda_i a_{i,k_2})$ subject to the condition that $a_{i,k_1} \cdot a_{i,k_2}$ are fixed, for all i. A simple calculation with Lagrange multipliers shows that the vectors $(a_{i,k_1} | i \in I)$ and $(a_{i,k_2} | i \in I)$ are proportionate. In other words, there is a nonnegative vector $(u_i | i \in I)$ and nonnegative constants $\rho_k (k \in K)$ such that $a_{i,k} = \rho_k \cdot u_i$ for every $i \in I, k \in K$. Multiply by λ_i and sum over $i \in I$ to conclude that for any $k \in K$, $\alpha_k = \rho_k \sum_I \lambda_i u_i$. We can write now, for every $i \in I$:

$$(\prod_{j\in J} \alpha_j)^{1/|K|} = (\prod_{j\in J} (\max_{i\in I} a_{i,j}))^{1/|K|} \geq (\prod_{j\in J} a_{i,j})^{1/|K|}$$

and,
$$(\prod_{k\in K}\rho_k)^{1/|K|}u_i = (\prod_{k\in K} a_{i,k})^{1/|K|}$$
So, for every $i \in I$,
$$(\prod_{j\in J}\alpha_j)^{1/|K|}(\prod_{k\in K}\rho_k)^{1/|K|}u_i \geq (\prod_{j\in J\cup K} a_{i,j})^{1/|K|} = b_i \tag{2}$$
Multiply Eq. (2) by $\rho_t\lambda_i$, sum over $i \in I$ and use $\alpha_t = \rho_t \sum_{i\in I}\lambda_i u_i$ and $\beta = \sum_{i\in I}\lambda_i b_i$, to conclude that for every $t \in K$,
$$(\prod_{j\in J}\alpha_j)^{1/|K|}(\prod_{k\in K}\rho_k)^{1/|K|}\alpha_t \geq \rho_t \cdot \beta.$$
Now multiply over all $t \in K$ to get the desired conclusion. ∎

4 Two-Party Protocols

In this section we present protocols which meet the lower bounds presented in section 3, up to a polylogarithmic factor. We first present a general framework for the construction of such protocols (subsection 4.1), argue that this framework does indeed yield protocols meeting the lower bound (subsection 4.2), and finally use the framework to present *efficient* protocols meeting the lower bound (subsection 4.3).

Without loss of generality, we assume throughout that every value v in the range of f, appears in each row and column in the matrix of f at least $\frac{p_v}{4} \cdot 2^n$ times. If some row or column has too few occurrences of v, we'd like to add them, without a significant increase in q_v. This can be done as follows: Let $(A_1, \ldots, A_k)$ be a partition of $\{1, \ldots, 2^n\}$, where each A_i has cardinality between $\frac{p_v}{4} \cdot 2^n$ and $\frac{p_v}{2} \cdot 2^n$. It is easy to see that by changing some elements within the $A_i \times A_i$ minors of the matrix to v, it is possible to guarantee that v-values have density $\geq \frac{p_v}{4}$ in every row and column without increasing the largest density in any row or column beyond $q_v + \frac{p_v}{4} = O(q_v)$.

Also, without loss of generality, we assume $p_v \leq 1/2$ (otherwise, the claims hold vacuously).

4.1 Framework for Protocols Meeting the Lower Bounds

The goal of the protocol is to enable the parties to gradually reveal their inputs to each other, without granting any party a substantial influence on the value of f.

The protocol proceeds in rounds, each consisting of two steps. In each step one party sends one bit of information about its input to the other party. In the next step the other party sends such a bit. The bits sent by each party specify in which side, of a bipartition of the residual input-space, its actual input lies. These partitions must satisfy some "value-balance" properties to be discussed below. Following is the code of the generic protocol.

Inputs: $x \in X_0 \stackrel{\text{def}}{=} \{0,1\}^n$ for the row player, $y \in Y_0 \stackrel{\text{def}}{=} \{0,1\}^n$ for the column player.

Round i: Let (X^0_{i-1}, X^1_{i-1}) be a partition of X_{i-1}, and (Y^0_{i-1}, Y^1_{i-1}) a partition of Y_{i-1}.
The row player sends $\sigma \in \{0,1\}$ such that $x \in X^\sigma_{i-1}$. Let $X_i \stackrel{\text{def}}{=} X^\sigma_{i-1}$.
The column player sends $\sigma \in \{0,1\}$ such that $y \in Y^\sigma_{i-1}$. Let $Y_i \stackrel{\text{def}}{=} Y^\sigma_{i-1}$.

Output: When both residual sets become singletons (i.e., $|X_t| = |Y_t| = 1$ after round t) the protocol terminates and the output is defined as $f(x,y)$, where $X_t = \{x\}$ and $Y_t = \{y\}$).

The reader may think of the partitions as splitting the current set evenly and in fact this is almost the case as asserted in Property (P0). In such a case, the protocol terminates after n rounds. For the protocol to achieve its goal (of minimizing the advantage of each party), it employs bipartitions satisfying various (additional) *value-balance* properties. There will be several different types of value-balance properties all sharing the following features. These properties apply both to row-partitions and column-partition. A typical row-partition property (resp., column-partition property) requires that a subset of the rows (resp., columns), specified by some pattern of v-entries, is split almost evenly between the two sides of the partition. For example, Property (P1) below (regarding column-partitions) requires that, for each row, the set of columns containing a v-entry in this row is split almost evenly.

We will introduce the various properties in an ad-hoc manner, each property being introduced just where it becomes essential for analyzing the generic protocol. Thus, at the end of this subsection, we will have a set of properties and a proof that if the protocol utilizes only partitions having these properties, then the advantage of both parties is bounded as claimed in the introduction. The question of whether such partitions exist will be ignored altogether in the current subsection but will be the focus of the next subsection, whereas the third subsection shows how to efficiently generate "pseudorandom" partitions which satisfy these properties.

Motivation to the analysis of the protocol

In analyzing the influence of a dishonest party we consider, without loss of generality, the probability that the row player (following an arbitrary adversarial strategy) succeeds in having the protocol yield a particular value v (in the range of f). For simplicity, we consider first the special case where $q_v = p_v$. In this case there are exactly $K \stackrel{\text{def}}{=} p_v \cdot N$ entries of value v in each row of the matrix. The analysis proceeds in three stages:

stage 1: Consider the first $\log_2 K$ rounds. If every column (resp. row) partition employed halves the number of v-entries in each row (resp. column), then at the end of this stage the residual $\frac{1}{p_v}$-by-$\frac{1}{p_v}$ matrix contains a single v-entry in each row (resp. column), thus preserving the density of v-entries in each row and column. Using a v-balance property of the partitions called (P1), we show that this is roughly the situation (see Corollary 7).

stage 2: Consider the next $\frac{1}{2} \cdot \log_2(1/p_v)$ rounds. If each row (resp. column) partition employed halves the number of v-entries in the residual matrix, then at the end of this stage the residual $\frac{1}{\sqrt{p_v}}$-by-$\frac{1}{\sqrt{p_v}}$ matrix contains a single v-entry, thus preserving the density of v-entries. Using a v-balance property of the partitions called (P2), we show that this is roughly the situation (see Lemma 8).

stage 3: At the last $\frac{1}{2} \cdot \log_2(1/p_v)$ rounds the row player can force the outcome to be v only if the input of the column player is a column containing a v-entry. The probability that the input column of the column player contains a v-entry does not exceed $\Delta \cdot \sqrt{p_v}$, where Δ is the number of v-entries at the outset of this stage.

Preliminaries

All value-balance properties are geared to guarantee an "almost even split" of certain sets. This is quantified in the following definition with bounds that depend on the size of the set to be split. The size-ranges are parameterized by b. For sets smaller than b we require nothing. For sets larger than b^4 we require sublinear discrepancy/bias, and in the midrange we require a small-but-linear discrepancy.

Definition 1 (almost unbiased partitions): *Let $S \subseteq U$ be finite sets and $b > 1$. A partition (U^0, U^1) of U is* at most b-biased with respect to S *if:*
1) *If $|S| \geq b^4$ then $\left| |U^0 \cap S| - \frac{|S|}{2} \right| < |S|^{3/4}$.*
2) *If $b < |S| < b^4$ then $\left| |U^0 \cap S| - \frac{|S|}{2} \right| < \frac{|S|}{20}$.*

In our analysis of the protocol, we assume that it utilizes partitions which are at most $\delta \cdot \log_2(1/p_v))$-biased with respect to specific sets, where δ is a constant to be determined as a function of other constants which appear in the analysis (see subsections 4.2 and 4.3). We stress that p_v denotes the density of v-entries in the original matrix corresponding to the function f (and not the density in any residual submatrices defined by the protocol). We denote $\Delta_v \stackrel{\text{def}}{=} \delta \log_2(1/p_v)$. Whenever obvious from the context, we abbreviate Δ_v by Δ.

In addition to value-balance properties, we use the following more elementary property asserting that the partitions are into almost equal sizes. The parameter of approximation is determined by the frequency of the value being discussed in the context.

Definition 2 (balance property P0): *A partition (U^0, U^1) of U is said to have* Property (P0) (with respect to a parameter Δ) *if the partition is at most Δ-biased with respect to U. When $|U| \geq 2$ it is also required that the partition be* non-trivial*; namely $|U^0|, |U^1| \geq 1$.*

The additional condition guarantees that if the generic protocol uses only partitions with Property (P0) then it terminates. The main condition in Property (P0) implies termination in at most $n + \Delta$ rounds (see Claim 5 and the proof of Lemma 6).

We consider executions of the generic protocol under various strategies of the row player, typically assuming that the column player plays honestly. The *residual submatrix* after i rounds is the submatrix corresponding to $X_i \times Y_i$. We denote by $\#_v(X, Y)$ the number of v-entries in the submatrix induced by $X \times Y$. When X is a singleton, $X = \{x\}$, we abbreviate and write $\#_v(x, Y)$ instead of $\#_v(X, Y)$. For example, for $x \in X_i$, the number of v-entries in the residual x-row after i rounds (resulting in the residual submatrix $X_i \times Y_i$) is denoted $\#_v(x, Y_i)$.

Analysis of the Protocol: The Special Case of $q_v = p_v$

For the analysis of this special case, we need two types of "value-balance" properties. The definition is phrased for column partition. An analogous definition holds for row partitions.

Definition 3 (value-balance properties P1 and P2): *Let X_i and Y_i be residual sets of rows and columns and let (Y_i^0, Y_i^1) be a* (column) *partition of Y_i, and v be a value in the range of f. We consider the following two properties:*

Property (P1): *The partition is v-*balanced with respect to individual rows *if the following holds. For every* (remaining) *row $x \in X_i$, the partition is at most Δ_v-biased with respect to set of columns having v-entries in row x* (i.e., w.r.t. the sets $\{y \in Y_i : f(x,y) = v\}$, for each $x \in X_i$).

Property (P2): *Either $|Y_i| \geq 2/p_v$ or the partition is v-*balanced with respect to the standard coloring *in the following sense. Consider a standard* minimum *coloring, ξ, of the v-entries in $X_i \times Y_i$ where no two v-entries in the same column or row are assigned the same color. For every color α, the partition is at most Δ_v-biased with respect to the set of columns containing a v-entry of color α* (i.e., w.r.t. the sets $\{y \in Y_i : \exists x \in X_i \text{ s.t. } f(x,y) = v \text{ and } \xi(x,y) = \alpha\}$, over $\alpha \in \text{Range}(\xi)$).

Here is an elementary technical claim, which we use extensively in the analysis:

Claim 5 : *Let $\alpha < 1$. Suppose that $z_{i+1} < \frac{z_i}{2} + (z_i)^\alpha$, for every $i = 0, \ldots, T$. Then, there exists a constant c_α, so that $z_t < \frac{z_0}{2^{t-1}}$, for every $t < \min\{T, (\log_2 z_0) - c_\alpha\}$. Likewise, if $z_{i+1} > \frac{z_i}{2} - (z_i)^\alpha$, for every $0 \leq i \leq T$, then $z_t > \frac{z_0}{2^{t+1}}$, for every $t < \min\{T, (\log_2 z_0) - c_\alpha\}$.*

proof: By successively applying the inequality t times, we get $z_t < \frac{z_0}{2^t} + \sum_{i=1}^{t} \frac{z_{t-i}^\alpha}{2^{i-1}}$. Using induction on t, we get

$$\begin{aligned} z_t &< \frac{z_0}{2^t} + \sum_{i=1}^{t} \frac{(z_0/2^{t-i-1})^\alpha}{2^{i-1}} \\ &= \frac{z_0}{2^t} + 2\left(\frac{2z_0}{2^t}\right)^\alpha \cdot \sum_{i=1}^{t} \left(\frac{1}{2^{1-\alpha}}\right)^i \\ &< \frac{z_0}{2^t} + 2^{1+\alpha} \cdot \left(\frac{z_0}{2^t}\right)^\alpha \cdot \frac{1}{2^{1-\alpha} - 1} \end{aligned}$$

which is bounded by $\frac{z_0}{2^{t-1}}$, provided that $\frac{z_0}{2^t} > 2^{c_\alpha}$ where $c_\alpha \stackrel{\text{def}}{=} \frac{1}{1-\alpha} \cdot \log_2(2^{1+\alpha}/(2^{1-\alpha} - 1))$. ■

We start by showing that the density of v-entries in individual rows and columns hardly changes as long as each such row/column contains enough v-entries and the partitions split them almost evenly. This assertion corresponds to stage (1) in the motivating discussion.

Lemma 6 (stage 1): *Let v be a value in the range of f, and suppose that the protocol uses* column *partitions satisfying Property* (P1) w.r.t. the value v. *Let K_x denote the number of v entries in the original row x. Then, regardless of the players' steps, if row x is in the residual matrix after the first $i \stackrel{\text{def}}{=} \log_2 K_x$ rounds, then there are at most Δ_v residual v-entries in row x.* (i.e., $\#_v(x, Y_i) \leq \Delta_v$). *Furthermore, after $t < K_x$ rounds $\#_v(x, Y_t) \leq \Delta_v \cdot 2^{K_x - t}$.*

proof: The analysis uses the fact that the column partitions are v-balanced with respect to each row. Using condition (1) of the almost unbiased property (Def. 1) and Claim 5, we see that after the first $s \stackrel{\text{def}}{=} \log_2 K_x - 4\log_2 \Delta$ rounds the residual row x has at most $\frac{K_x}{2^{s-1}} = 2\Delta^4$ entries of value v. For the remaining $r \stackrel{\text{def}}{=} 4\log_2 \Delta$ rounds we use condition (2) of the almost unbiased property, to show that the number of v-entries in the row is at most Δ. This follows by considering r iterations of condition (2), namely

$$\begin{aligned}
2\Delta^4 \cdot \left(\frac{1}{2} + \frac{1}{20}\right)^{4\log_2 \Delta} &= 2 \cdot \left(1 + \frac{1}{10}\right)^{4\log_2 \Delta} \\
&= 2 \cdot \Delta^{4\log_2(1+\frac{1}{10})} \\
&< 2 \cdot \Delta^{2/3} \\
&\leq \Delta
\end{aligned}$$

where in the last inequality we use $\delta \geq 8$ (and $p_v \leq 1/2$). The lemma follows. ■

As an immediate corollary, we get –

Corollary 7 (stage 1 for $q_v = p_v$): *Let $v \in Range(f)$, and suppose that $q_v = p_v$. Suppose that the protocol uses* column (resp., row) *partitions satisfying Property* (P1) w.r.t. the value v. *Then after the first $n - \log_2(1/p_v)$ rounds, the number of v-entries in each residual row* (resp., column) *is at most Δ_v $(= \delta \cdot \log 1/p_v)$. This statement holds regardless of the steps taken by the players.*

proof: Observe that $q_v = p_v$ implies that each (original) row has $p_v \cdot 2^n$ entries of value v, and apply Lemma 6. ■

When the number of v-entries in individual rows and columns is small, but not too small, we'd like to assert something in the spirit of stage (2) of the motivating discussion. Namely, that the density of v-entries in the *entire* matrix is preserved as long as their total number is not too small and the partitions behave nicely w.r.t the existing v-entries.

Lemma 8 (stage 2): *Let $M < 2/p_v$. Consider an M-by-M matrix where no row or column contains more than B v-entries. Suppose that the protocol is applied to this matrix, using* column and row *partitions that satisfy Property* (P2) w.r.t. the value v. *Then, after the first $\frac{1}{2}\log_2 M$ rounds, the number of v-entries in the residual submatrix is at most $(2B+1)\cdot\Delta_v$. This statement holds regardless of the steps taken by the players.*

proof: The analysis uses only the fact that the row and column partitions are v-balanced with respect to the standard coloring. (The upper bound on M implies that this is the only way to satisfy Property (P2).) Note that the standard coloring, being a minimum coloring, uses at most $2B+1$ colors since the underlying graph has maximum degree $\leq 2B$. Let α be a color. In each row and column there is at most one v-entry of color α, hence each row/column partition affects the number of remaining v-entries of color α exactly as it affects the number of v-entries in a specific row/column. Hence, using the same arguments as in Lemma 6, we see that after $\frac{1}{2}\log_2 M$ rounds the residual matrix contains at most Δ_v v-entries of color α. The lemma follows. ■

Finally, when the total number of v-entries in the residual matrix is small we observe that v may be the output only if the input of the column player corresponds to a residual column containing a v-entry. This corresponds to stage (3) in the motivating discussion. Thus, using Corollary 7 and Lemma 8, we get

Corollary 9 (advantage in case $q_v = p_v$): *Let $q_v = p_v$ for $v \in Range(f)$. Suppose that the protocol uses only partitions that satisfy Properties* (P0), (P1) *and* (P2) *w.r.t. v. Then the protocol outputs v with probability at most $O(\Delta_v^2 \sqrt{p_v})$ $(= O((\delta \log 1/p_v)^2 \sqrt{p_v}))$, regardless of the row player's steps.*

proof: Corollary 7 and Lemma 8 imply that after the first $\log_2(p_v N) + \frac{1}{2}\log_2(1/p_v)$ rounds, the number of v-entries in the residual matrix is at most $O(\Delta^2)$. If in all partitions the two parts have equal size, then the residual matrix has dimension $\sqrt{1/p_v}$-by-$\sqrt{1/p_v}$. Property (P0) is applied to show that the residual submatrix has size at least $\frac{1}{2}\sqrt{1/p_v}$-by-$\frac{1}{2}\sqrt{1/p_v}$. To this end we use Claim 5 and the observation that $\sqrt{1/p_v} > \Delta_v^4 = (\delta \log_2(1/p_v))^4$, provided that p_v is bounded above by some constant. Such a bound on p_v may be assumed, possibly increasing some constants in the O-terms. Finally, we observe that the output of the protocol is v only if the input of the column player specifies a column containing a v-entry in the residual submatrix. The corollary follows. ■

Using "sufficiently random" partitions, the above bound can be improved to $O(\sqrt{p_v})$. For details see Theorem 22.

Analysis of the Protocol: The General Case – Row Classes

The analysis of the general case (where q_v may exceed p_v) is more cumbersome. To facilitate the understanding we precede each technical step by a motivating discussion. As before, we analyze the advantage of the row player towards some value v. Throughout the analysis we introduce additional value-balance properties that the partitions used in the protocol should satisfy for the analysis to go through. Later in the paper we discuss how to find such partitions, and show that "slightly random" partitions do have these properties.

We classify the rows by density and apply the analysis separately to each class. Let $\rho_v(x)$ denote the density of v-entries in row x of the original matrix; that is,

$$\rho_v(x) \stackrel{\text{def}}{=} \frac{|\{y \in Y_0 : f(x,y) = v\}|}{|Y_0|} = \frac{\#_v(x, Y_0)}{|Y_0|} \tag{3}$$

By our assumption, $\frac{p_v}{4} < \rho_v(x) \leq q_v$, for every $x \in X_0$, and the average of ρ_v, over all $x \in X_0$, equals p_v. For $0 \leq j \leq \log_2(1/p_v) + 1$, define R^j as the class of all rows with v-entry density between 2^{-j} and 2^{-j-1}; that is,

$$R^j \stackrel{\text{def}}{=} \{x \in X_0 : \lfloor \log_2(1/\rho_v(x)) \rfloor = j\} \tag{4}$$

Note that the last class, $R^{\log_2(1/p_v)+1}$, contains all rows with v-entry density smaller than $p_v/2$.

The following simple observation is stated separately only for future reference.

Clearly, the influence of the row player towards value v is bounded by the sum of its influences (towards v) when restricting itself to inputs/rows of a certain class. Recall that the row player behavior is restricted (by our hypothesis that it is not detected cheating) to sending a single bit in each round. The assumption that the row player restricts itself to inputs/rows in a particular set means that its answers must be consistent with some input in the set (i.e., in round i he may send σ only if X_i^σ intersects the restricted set). The above is summarized and generalized in the following claim

Claim 10 *For $Z \subseteq X$ a set of rows, we let θ_Z be the probability for an outcome of v, assuming that the row player must restrict its final choice to a row in Z, but is otherwise free to choose any adversarial strategy. If $(Z_1, \ldots, Z_r)$ is a partition of the set of rows, then the probability for the protocol to have outcome v does not exceed $\sum_i \theta_{Z_i}$.*

proof: The claim follows by conditioning on the row selected by the row player and applying a union bound. ■

We now partition the row classes into two categories: heavy rows with density above $\sqrt{p_v}$ and rows below this density. First, we bound the advantage of the row player when it restricts itself to heavy inputs/rows. A simple counting argument implies that there are at most $\sqrt{p_v}N$ heavy rows. Observe the situation after $\log_2(\sqrt{p_v}N)$ rounds of the protocol. Using an additional v-balance property, denoted **P3**, which asserts that the row partitions split almost evenly the set of heavy rows, we will show that after $\log_2(\sqrt{p_v}N)$ rounds at most Δ of the heavy rows remain in the residual matrix and furthermore that each such row maintains its original v-density, up to a multiplicative factor of Δ. Loosely speaking, the row player can now choose only between Δ possible inputs/rows with probabilities of success that equal the density of the residual row. Thus, the advantage of the row player (towards v) when restricting itself to heavy rows is bounded by $\Delta^2 \cdot q_v = O((\log_2(1/p_v))^2 q_v)$.

Definition 4 (value-balance property P3): *Let X_i and Y_i be residual sets of rows and columns and and let $v \in Range(f)$. A row partition has* Property (P3) (is said to be v-balanced with respect to heavy rows) *if it is at most Δ_v-biased with respect to the set of the* (remaining) *heavy rows* (i.e., w.r.t. the set $\{x \in X_i : \rho_v(x) \geq \sqrt{p_v}\}$).

Lemma 11 (advantage via heavy-row strategies): *Suppose that the protocol is performed using* column and row *partitions satisfying Properties* (P0), (P1) and (P3) w.r.t. the value v. *Then, as long as the row player restricts itself to heavy rows and the column player plays honestly, the output equals v with probability at most $2\Delta_v^2 \cdot q_v$.*

proof: Consider the situation after $\log_2(\sqrt{p_v}N)$ rounds of the protocol. Heavy rows have at least $\sqrt{p_v}N$ entries of value v and so we will be able to apply Lemma 6 to these rows. Using Property (P1) and applying Lemma 6 to each heavy row, we conclude that every remaining heavy row x contains at most $\Delta \cdot 2^i$ v-entries, where

$$\begin{aligned} i &\stackrel{\text{def}}{=} \log_2(\rho_v(x)N) - \log_2(\sqrt{p_v}N) \\ &\leq \log_2(q_v N) - \log_2(\sqrt{p_v}N) \\ &= \log_2(q_v/\sqrt{p_v}) \end{aligned}$$

(We are assuming that heavy rows exist, i.e., $q_v \geq \sqrt{p_v}$, whence $i \geq 0$.) Thus, each such heavy row contains at most $\Delta \cdot q_v/\sqrt{p_v}$ v-entries. Also, using Property (P3) and an argument as in the proof of Lemma 6, it follows that the residual matrix has at most Δ heavy rows. Thus, the entire residual matrix contains at most $\Delta^2 \cdot q_v/\sqrt{p_v}$ *v-entries in heavy rows.* Using Property (P0) we know that the residual matrix at this stage contains at least $\frac{1}{2}\sqrt{1/p_v}$ columns. Thus, by an argument as in the proof of Corollary 9, the probability that the protocol terminates with a pair (x,y) so that x is heavy and $f(x,y)=v$ does not exceed

$$\frac{\#_v(H\cap X_i,Y_i)}{|Y_i|}\leq\frac{\Delta^2\cdot q_v/\sqrt{p_v}}{1/(2\sqrt{p_v})}=2\Delta^2 q_v$$

where H is the set of heavy rows and $X_i\times Y_i$ is the residual matrix. The lemma follows. ■

Having analyzed strategies where the row player confines itself to heavy rows, we turn to strategies where it refrains from heavy rows. The analysis is split according to the remaining row-classes; that is, for every $1\leq j\leq\frac{1}{2}\log_2(1/p_v)$, we bound the advantage of the row player assuming that it restricts itself to the class (of rows) $R\stackrel{\rm def}{=}R^{j+\frac{1}{2}\log_2(1/p_v)}$ that have density $\approx\sqrt{p_v}2^{-j}$. By a counting argument –

$$|R|\leq\sqrt{p_v}2^jN \tag{5}$$

Consider the situation after $\log_2(\sqrt{p_v}2^{-j}N)$ rounds. Note that this corresponds to stage (1) in the motivating discussion and thus we can apply Lemma 6 and assert that after these $\log_2(\sqrt{p_v}2^{-j}N)$ rounds no residual row of R has more than Δ v-entries. Using an additional v-balance property, denoted **P4**, which asserts that the row partitions split R almost evenly, we will show that after these $\log_2(\sqrt{p_v}2^{-j}N)$ rounds the residual matrix contains at most $\max\{\Delta,2^{2j+1}\}$ rows of R.

Definition 5 (value-balance property P4): *Let X_i and Y_i be residual sets of rows and columns and $v\in Range(f)$. A row partition has* Property (P4) (is said to be v-balanced with respect to row-density classes) *if, for every j ($\frac{1}{2}\log_2(1/p_v)\leq j\leq 1+\log_2(1/p_v)$), it is at most Δ_v-biased with respect to the set of the* (remaining) *rows in R^j* (i.e., w.r.t. the sets $\{x\in X_i:\lfloor\log_2\rho_v(x)\rfloor=j\}$, for $\frac{1}{2}\log_2(1/p_v)\leq j\leq 1+\log_2(1/p_v)$).

Lemma 12 (strategies restricted to $R=R^{j+\frac{1}{2}\log_2(1/p_v)}$ – the first rounds): *Let $v\in Range(f)$, and suppose that the protocol uses* row and column *partitions satisfying Properties* (P0), (P1) and (P4) w.r.t. the value v. *Then after the first $i\stackrel{\rm def}{=}n-j-\frac{1}{2}\log_2(1/p_v)$ rounds, the resulting $X_i\times Y_i$ submatrix satisfies the following conditions, regardless of the players' steps:*

1. *each remaining row of R contains at most Δ_v entries of value v* (i.e., $\#_v(x,Y_i)\leq\Delta_v$, for every $x\in R\cap X_i$);
2. *at most $\Delta_v\cdot 2^{2j+1}$ rows of R remain* (i.e., $|R\cap X_i|\leq\Delta_v\cdot 2^{2j+1}$);
3. *the number of columns is at least $\frac{1}{2}\cdot\frac{2^j}{\sqrt{p_v}}$* (i.e., $|Y_i|\geq\frac{1}{2}\cdot\frac{2^j}{\sqrt{p_v}}$).

proof: Item (1) follows from Lemma 6 (using Property (P1)). Using Property (P4), we derive Item (2) as in the second part of the proof of Lemma 11. Finally, Item (3) follows using Property (P0). ■

For "small" j's (say, $j\leq\log_2\Delta$) we get into a situation as in the analysis of heavy rows. Namely,

Corollary 13 (advantage via $R=R^{j+\frac{1}{2}\log_2(1/p_v)}$ strategies – simple analysis): *Consider a protocol in which all* column and row *partitions satisfy Properties* (P0), (P1) and (P4) w.r.t. the value v. *Then, as long as the row player restricts itself to rows in R and the column player plays honestly, the output equals v with probability at most $2^{j+2}\Delta_v^2\cdot\sqrt{p_v}$.*

proof: Using Lemma 12 we infer that the residual matrix after i rounds has at most $\Delta^2 \cdot 2^{2j+1}$ v-entries in rows of R and at least $\frac{1}{2} \cdot \frac{2^j}{\sqrt{p_v}}$ columns. Thus, the probability that the column chosen by the column player has a v-entry in a residual row of R does not exceed

$$\frac{\#_v(R \cap X_i, Y_i)}{|Y_i|} \leq \frac{\Delta^2 \cdot 2^{2j+1}}{\frac{1}{2} \cdot \frac{2^j}{\sqrt{p_v}}} = 2^{j+2}\Delta^2\sqrt{p_v}$$

The corollary follows. ■

So far we dealt with heavy rows and the row classes $R^{j+\frac{1}{2}\log_2(1/p_v)}$ for "small" j's, $j \leq \log_2 \Delta_v$. The rest of the analysis concentrates on row classes $R^{j+\frac{1}{2}\log_2(1/p_v)}$ for $j > \log_2 \Delta_v$.

Analysis of the Protocol: The General Case – Column Subclasses

Lemmas 11 and 12 summarize what we can infer by considering only row-classes defined by the density of v-entries. We learnt that after $i = n - j - \frac{1}{2}\log_2(1/p_v)$ rounds the resulting matrix has approximately 2^{2j} rows of the class $R = R^{j+\frac{1}{2}\log_2(1/p_v)}$ with no more than Δ v-entries each. Thus, in total the resulting submatrix has approximately 2^{2j} v-entries in rows of R. Had these v-values been distributed evenly among the columns, then we could apply an argument analogous to Lemma 8 (corresponding to stage (2) in the motivating discussion). At the other extreme, if these v-values are all in one column, then we should have further applied Lemma 6 to this column. In general, however, the distribution of these v-entries may be more complex and in order to proceed we classify columns according to the approximate density of v-entries within each particular row-class. Once this is done, the matrix is split to submatrices such that the density of v-entries in each induced sub-column is about the same. Each such submatrix is easy to analyze and we can combine these analyses to derive the final result.

Notations: Let $\ell \stackrel{\text{def}}{=} \frac{1}{2}\log_2(1/p_v)$. Recall that we are currently dealing with an arbitrary $R = R^{j+\ell}$, where $1 < j \leq \ell + 1$. For $0 \leq k \leq 2j$, let

$$C_j^k \stackrel{\text{def}}{=} \{y \in Y_0 : \lfloor \log_2(1/\mu_v(y, R^{j+\ell})) \rfloor = k\} \tag{6}$$

where $\mu_v(y, R)$ is the density of v-entries in the *portion* of column y restricted to rows R, that is

$$\mu_v(y, R) = \frac{|\{x \in R : f(x, y) = v\}|}{|R|} \tag{7}$$

Columns having lower v-density within R (i.e., $\mu_v < 2^{-2j-1}$) are defined to be in $C_j = C_j^{2j+1}$ and will be treated separately. The advantage of the row player towards v when restricting its input to R is the sum, over all k, of the probabilities of the following $2j+2$ events. For $k = 0, ..., 2j+1$, the k^{th} *event* occurs if the input of the column player happens to be in C_j^k *and* the output of the protocol is v (when the row player restricts its input to be in R). Thus, it suffices to bound the probability of each of these $2j+2$ events. We first observe that, for $j = 0, ..., l+1$,

$$|C_j^k| \leq \frac{\#_v(R^{j+\ell}, Y_0)}{\min_{y \in C_j^k}\{\#_v(R^{j+\ell}, y)\}} \leq \frac{|R^{j+\ell}| \cdot (\sqrt{p_v}2^{-j} \cdot N)}{2^{-k-1} \cdot |R^{j+\ell}|} = 2^{k+1-j}\sqrt{p_v} \cdot N \tag{8}$$

Thus, the probability that the input of the column player is in C_j^k is bounded by $2^{k+1-j}\sqrt{p_v}$. This by itself provides a sufficiently good bound for the case $k \leq j$ and so it is left to consider the case where $j < k \leq 2j$ and to deal with the columns in C_j. We start with the latter. (Warning: the next two paragraphs consist of an imprecise motivating discussion – a rigorous treatment follows.)

Considering the submatrix $R \times C_j$ and using Item (2) of Lemma 12 we know that, after $i = n - j - \ell$ rounds, each residual row in this submatrix contains at most Δ v-entries. Assuming that the row partitions split the v-entries in the sub-column of this submatrix almost evenly (as postulated in an additional value-balance property, denoted **P6**), we conclude that residual sub-columns of the submatrix contain at most Δ v-entries (note that there are at most 2^{2j+1} rows of R and that the v-density of columns in $R \times C_j$ is at most 2^{-2j-1}). Thus, we can apply an analysis analogous to stage (2) in the motivating discussion. It follows that after an additional j rounds, the resulting submatrix contains at most Δ^2 v-entries. At this stage, there are still $\ell = \frac{1}{2}\log_2(1/p_v)$ rounds to go so we conclude that the probability that the column player's input is in C_j and the output is v (when the row player restricts its input to be in R) is at most $\Delta^2\sqrt{p_v}$. This argument will be made precise as a special case of the argument for C_j^k, $k > j$.

We now consider the submatrix $R \times C$, where $C \stackrel{\text{def}}{=} C_j^k$ for $k > j$. Again, by Property (P6) we expect each residual sub-column to contain $2^{-k} \cdot 2^{2j}$ entries of value v. Assuming that the column-partitions spilt C almost evenly, as postulated in yet another value-balance property (**P5** below), and using Eq. (8), we expect the residual submatrix to contain at most 2^{k+1} columns of C_j^k (and, recall, 2^{2j} rows of R). Thus, the next $2j - k < k$ rounds are expected to preserve the density of C columns in the residual matrix as well as the density of v-entries in residual sub-columns of the submatrix $R \times C$, provided that Properties (P5) and (P6) hold. Thus, at this point (after a total of $(n - j - \ell) + (2j - k)$ rounds) each remaining row of R is left with at most Δ entries of value v and each remaining column of C has at most Δ entries of value v in the portion of the rows of R. Furthermore, we expect the residual $R \times C$ to have $2^{2j-(2j-k)} = 2^k$ rows and $2^{k-(2j-k)} = 2^{2k-2j}$ columns. We can now apply an argument analogous to Lemma 8 (corresponding to stage (2) in the motivating discussion). To this end we introduce the last value-balance property, denoted **P7**, which analogously to (P2) asserts that, with respect to each color in a standard minimum coloring of the v-entries in $R \times C$, the row (resp. column) partitions split almost evenly the set of rows (resp. columns) having v-entries colored by this color. Finally, consider the situation after another additional $k - j$ rounds. Using (P7) in an argument analogous to Lemma 8, we show that after these $k - j$ rounds, the residual $R \times C$ submatrix has at most Δ^2 v-entries. Furthermore, this residual submatrix is expected to have $2^{k-(k-j)} = 2^j$ rows and $2^{(2k-2j)-(k-j)} = 2^{k-j}$ columns. Thus, assuming that the column player's input, denoted y, is in C the probability that it falls in one of the residual columns which has a v-entry in the R-portion is at most $\Delta^2/2^{k-j}$. It follows that the probability for the input column to be in C_j^k *and* the output be v (when the row player restricts its input to R) is at most

$$\frac{\Delta^2}{2^{k-j}} \cdot 2^{k-j}\sqrt{p_v} = \Delta^2 \cdot \sqrt{p_v}$$

Thus, the claimed bound follows also in this case.

We now turn to a rigorous analysis of the advantage of the row player in executions where it restricts itself to inputs in $R = R^{j+\ell}$ and the input column happens to fall in $C \stackrel{\text{def}}{=} C_j^k$, for some

$k > j > 0$. (Recall that for $k \le j$, Eq. (8) by itself asserts that input column falls in C_j^k with probability at most $\sqrt{p_v}$.)

Definition 6 (value-balance properties P5, P6 and P7): *Let X_i and Y_i be residual sets of rows and columns. Let (X_i^0, X_i^1) be a* row *partition, (Y_i^0, Y_i^1) be a* column *partition, and $v \in Range(f)$. We consider the following three properties:*

- Property (P5): *The column partition (Y_i^0, Y_i^1) is v-*balanced with respect to column subclasses *if, for every j, k satisfying $0 < j < k \le 2j \le 2\ell + 2$, the partition is at most Δ_v-biased with respect to the set of columns in C_j^k* (i.e., w.r.t. the sets $Y_i \cap C_j^k$, for each j, k s.t. $0 < j < k \le 2j \le 2\ell + 2$).

- Property (P6): *For every j and every $y \in Y_i$, either $\frac{\#_v(X_i \cap R^{j+\ell}, y)}{|Y_i|} \le \frac{p_v}{4\Delta_v}$ or the row partition (X_i^0, X_i^1) is v-*balanced with respect to the j^{th} subcolumn *y in the sense that the partition is at most Δ_v-biased with respect to the set of rows in $R^{j+\ell}$ having v-entries in y* (i.e., w.r.t. $\{x \in X_i \cap R^{j+\ell} : f(x,y) = v\}$, for each $y \in Y_i$ and j s.t. $0 < j \le \ell + 1$).

- Property (P7)*: Either $|Y_i| \ge 4/p_v$ or the partition (Y_i^0, Y_i^1) is v-*balanced with respect to the standard coloring of subclasses *in the following sense. For every j, k as in* (P5)*, consider a standard* minimum *coloring ξ, of the v-entries in $(X_i \cap R^{j+\ell}) \times (Y_i \cap C_j^k)$ so that every two v-entries in the same column or row are colored differently. For every color α, the partition is at most Δ_v-biased with respect to the set of columns containing a v-entry of color α* (i.e., w.r.t. the sets $\{y \in Y_i \cap C_j^k : \exists x \in X_i \cap R^{j+\ell} \text{ s.t. } f(x,y) = v \text{ and } \xi(x,y) = \alpha\}$, for each j, k and α.)

Definition 7 (the (j,k)-event): *Let $0 < j \le \ell + 1$ and $0 \le k \le 2j + 1$. Fix an arbitrary strategy in which the row player restricts its input to rows in $R^{j+\ell}$. The (j,k)-*event (or k^{th} event) *is said to occur if both the input column is in C_j^k and the output is v.*

Lemma 14 (bounding individual events): *Let $0 < j \le \ell + 1$ and $0 \le k \le 2j + 1$. Suppose that the protocol uses partitions which satisfy Properties* (P0), (P1), (P4), (P5), (P6) and (P7)*. Then, for any strategy in which the row player restricts its input to rows in $R^{j+\ell}$, the probability of the (j,k)-event is at most $5\Delta_v^4 \cdot \sqrt{p_v}$.*

We remark that a lower power of Δ_v can be obtained by a more careful analysis.

proof: As observed above, the bound holds in case $k \le j$, since in this case Eq. (8) implies that the column player's input is in C_j^k with probability at most $\sqrt{p_v}$. We thus turn to the case $j < k \le 2j + 1$.

First, we consider the situation after $i \stackrel{\text{def}}{=} (n - j - \ell) + (2j - k) = n + j - k - \ell$ rounds. Note that $j < k \le 2j + 1$ implies $i \ge (n - j - \ell) - 1 \ge n - \log_2(2/p_v)$ and $i < n - \ell$. We first bound the number of v-entries in the residual subrows and subcolumns of $R \times C$.

claim 14.1: Each remaining row of $R \stackrel{\text{def}}{=} R^{j+\ell}$ contains at most Δ v-entries; namely, $\#_v(x, Y_i) \le \Delta$, for every $x \in R \cap X_i$.

proof: Since $i \ge (n - j - \ell) - 1$, we can apply Lemma 12, and the claim follows by Item (1). □

claim 14.2: Each remaining column of $C \stackrel{\rm def}{=} C_j^k$ contains at most Δ entries of value v within its R-portion; namely, $\#_v(R \cap X_i, y) \leq \Delta$, for every $y \in C \cap Y_i$.

proof: We first bound the number of v-entries in the R-portion of each column $y \in C$. By combining the definition of C and Eq. (5), we get

$$\begin{aligned} \#_v(R,y) &\leq 2^{-k} \cdot |R| \\ &\leq \sqrt{p_v} \cdot 2^{j+n-k} \\ &= 2^{j+n-k-\ell} \\ &= 2^i \end{aligned}$$

We now wish to apply Property (P6) and argue that $\#_v(R \cap X_i, y) \leq \Delta \cdot \#_v(R,y) \cdot 2^{-i}$, but we need to be careful since Property (P6) is useful only when $\#_v(R \cap X_t, y) \geq \frac{p_v}{4\Delta} \cdot |Y_t|$. Thus, before applying Property (P6), we consider the simple case in which there are many v-entries in the R-portion of y; namely, $\#_v(R,y) \geq p_v \cdot |Y_0|$. Using Properties (P6) and (P0), we infer inductively that the ratio $\#_v(R \cap X_i, y)/|Y_i|$ is maintained after $r < i$ rounds. In the induction step we assume that the ratio after r rounds is at least $p_v/2$ and applying Proposition (P6) infer the same for $r+1$ rounds, provided $\#_v(R \cap X_r, y) \geq \Delta^4$. In the last ($\approx 4\log_2 \Delta$) rounds we maintain as invariant the assumption that the ratio is at least p_v/Δ_v. We conclude (analogously to Lemma 6) that $\#_v(R \cap X_i, y) \leq \Delta \cdot 2^{i-i} = \Delta$ as claimed. Yet, all the above is valid only in case the initial number of v-entries in the subcolumn is large enough (i.e., $\#_v(R,y) \geq p_v \cdot |Y_0|$) which need not be the case in general. Intuitively, this cannot be a problem since fewer v-entries in the subcolumn can only help. Formally, we proceed as follows. Let $y_0 \stackrel{\rm def}{=} |Y_0|$ and $z_0 \stackrel{\rm def}{=} \#_v(R,y)$. Consider i iterations of the following rule

- If $y_t > \Delta^4$ then set y_{t+1} to be in the interval $[(y_t/2) \pm y_t^{3/4}]$. If $y_t > \Delta$ then set y_{t+1} to be in the interval $[(y_t/2) \pm (y_t/20)]$. Otherwise set y_{t+1} to be in the interval $[0, y_t]$.
- If $z_t > (p_v/\Delta) \cdot y_t$ then set z_{t+1} analogously to the way y_{t+1} is set. Otherwise (i.e., $z_t \leq (p_v/\Delta) \cdot y_t$), set z_{t+1} to be in the interval $[0, z_t]$.

The above process corresponds to the decline (with $t = 0, ..., i$) of $|Y_t|$ (represented by y_t) and $\#_v(R \cap X_t, y)$ (represented by z_t), as governed by Properties (P0) and (P6). In case the initial ratio z_0/y_0 is sufficiently large, say at least p_v/Δ, Claim 5 implies that $z_i \leq \Delta$. As far as the y_t's are concerned, Claim 5 can be applied to yield $y_i \leq \Delta \cdot 2^{\ell+k-j}$, which by $k \leq 2j+1$ and $j \leq \ell+1$ yields $y_i \leq 2\Delta \cdot (1/p_v)$. Thus, it is clear that z_i is bounded by the maximum of the bound obtained in the simple case (i.e., Δ) and $(p_v/\Delta) \cdot y_i \leq 2$. The claim follows. $\Box$

We are now in a situation analogous to the end of stage (1) in the motivating discussion, except that the bounds on v-entries hold with respect to the residual $R \times C$ submatrix (rather than to the entire residual matrix). Our goal is to apply now a process analogous to stage (2) in the motivating discussion. To this end we first consider a minimum coloring of the v-entries in this residual submatrix (i.e., a coloring in which no v-entries in the same row/column are assigned the same color). Using Claims 14.1 and 14.2, we first observe that this coloring requires at most $2\Delta + 1$ colors (since the degrees in the induced graph do not exceed 2Δ). Next we derive an upper bound on the size of independent sets in the graph, whence on individual color classes in this coloring. An independent set in this graph meets every row and column at most once, so its cardinality cannot exceed $\min\{|R \cap X_i|, |C \cap Y_i|\}$.

claim 14.3: $\min\{|R \cap X_i|, |C \cap Y_i|\} \le 2\Delta \cdot 2^{2(k-j)}$.

proof: Using Property (P4) and Eq. (5), we get $|R \cap X_i| \le \Delta \cdot 2^{(n+j-\ell)-i} = \Delta \cdot 2^k$, so the claim holds when $k \ge 2j - 1$. Likewise, using Property (P5) and Eq. (8) and assuming $k \le 2j$, we get $|C_j^k \cap Y_i| \le \Delta \cdot 2^{(n+k+1-j-\ell)-i} = \Delta \cdot 2^{2(k-j)+1}$. This proves the claim for the range $k \le 2j$. □

We now consider an execution of the next $(k - j)$ rounds. As a preparation to applying Property (P6) we first consider the simple case in which the number of v-entries in the R-portion of column y is large; namely, $\#_v(R, y) \ge p_v \cdot |Y_0|$. Using Property (P7), we proceed analogously to Lemma 8. First, we upper bound the size of each residual color class by $\Delta \cdot \frac{2\Delta 2^{2(k-j)}}{2^{2(k-j)}} = 2\Delta^2$ (essentially, its size after i rounds divided by a factor of 2 for each of the $2(k - j)$ steps in the next $k - j$ rounds). Adding up the bounds for all color classes, we obtain a bound on the total number of v-entries in the resulting $R \times C$ submatrix; namely,

$$\#_v(X_{i+k-j} \cap R, Y_{i+k-j} \cap C) \le (2\Delta + 1) \cdot 2\Delta^2 < 5\Delta^3 \tag{9}$$

We are now in a situation analogous to the end of stage (2) in the motivating discussion. We note that till now $i + (k - j) = n - \ell$ rounds were performed. We distinguish two cases.

case 1: If $|C| < \Delta^3\sqrt{p_v} \cdot N$ then the bound on the (j, k)-event is obvious by Eq. (8) (as in case $k \le j$).

case 2 (the interesting case): Suppose $|C| \ge \Delta^3\sqrt{p_v} \cdot N$. In this case we use Property (P5) to infer that $|C \cap Y_{n-\ell}| \ge \frac{1}{\Delta} \cdot \frac{|C|}{\sqrt{p_v}N}$. Thus, using Eq. (9), the probability for the (j, k)-event is at most

$$\begin{aligned} \frac{|C|}{N} \cdot \frac{\#_v(X_{i+k-j} \cap R, Y_{i+k-j} \cap C))}{|C \cap Y_{n-\ell}|} &\le \frac{|C|}{N} \cdot \frac{5\Delta^3}{|C|/(\Delta\sqrt{p_v}N)} \\ &= 5\Delta^4 \cdot \sqrt{p_v} \end{aligned}$$

The lemma follows. ■

Combining Lemmas 11 and 14, we get

Theorem 15 (advantages in the general case): *Let f be an arbitrary bivariate function and suppose the generic protocol is performed with row and column partitions satisfying Properties* (P0) *through* (P7). *Then, for every value v in the range of f, if one party plays honestly then, no matter how the other player plays, the outcome of the protocol is v with probability at most* $O(\log^6(1/p_v) \cdot \max\{q_v, \sqrt{p_v}\})$.

proof: Just sum up the bounds for the probabilities of the ℓ^2 events corresponding to the advantage from "non-heavy" strategies (provided by Lemma 14) and add the bound on the advantage from heavy strategies provided by Lemma 11. (The summation over the strategies is an upper bound, whereas summation over the events corresponding to different column subclasses is exact.) ■

We stress that some logarithmic factors (but not all) can be eliminated by a more careful analysis.

	stated for	Description of the property: the partition approximately halves the number of	Number of sets for $M \times M$ matrix	Applicable for
P0	col	columns	1	all M
P1	col	columns with v-entries in row x, per row	M	all M
P2	col	columns with v-entries in color ϕ, per color	$\leq 2M+1$	$M \leq 2/p_v$
P3	row	heavy rows	1	all M
P4	row	rows of approximate weight 2^{-j}, per $j = 0, ..., \ell$	$\leq \ell$	all M
P5	col	columns of a weight class inside a row class, per row class and column subclass	$< 2\ell^2$	all M
P6	row	rows with v-entries in subcolumn, per column and row-class (provided residual subcolumn is sufficiently dense)	$M \cdot \ell$	all M
P7	col	columns with v-entries in color ϕ, per color in rectangle	$\leq (2M+1) \cdot \ell^2$	$M \leq 4/p_v$

Figure 1: Value-Balanced Properties $\quad (\ell \stackrel{\text{def}}{=} \frac{1}{2}\log_2(1/p_v))$

Digest of the Value-Balanced Properties

The value-balance properties, referred to in Theorem 15, are tabulated in Figure 1. Property (P2) is a specialization of Property (P7) for the case $q_v = p_v$ and is not used in the proof of Theorem 15 (but rather in the proof of Corollary 9). Properties (P2) and (P7) differ from all other value-balance properties in that their definition depends on a standard coloring of a graph induced by the current residual matrix $X_i \times Y_i$. In particular, the sets relevant to these properties, in different rounds vary in size. In contrast, we stress that sets relevant to the other properties reduce to about a half with every round. This "irregularity" of Properties (P2) and (P7) introduces difficulties in the subsequent subsections. To compensate for these difficulties, these properties were defined to hold vacuously as long as the residual matrix is "large" (i.e., $\Omega(1/p_v)$). As we pointed out, this convention does not affect the analysis, since Properties (P2) and (P7) are applied only to "small" residual matrices. For similar reasons, Property (P6) which refers to many (i.e., $|Y_i|$) sets which may be very small is also defined to hold vacuously in case the number of sets is much larger than the size of these sets. Note that all other properties either apply to fewer (i.e., poly(ℓ)) sets or refer to relatively big sets. Specifically, Properties (P3), (P4) and (P5) apply to poly(ℓ) sets. On the other hand, whenever Properties (P0) and (P1) are applied to many, say M, sets each of these sets has cardinality at least $M/2$ and $(p_v/4) \cdot M$, respectively.

4.2 On the Existence of Value-Balanced Partitions

In this subsection we prove the existence of partitions that have all the value-balance properties used in the previous subsection. We first bound the probability that a random partition is not balanced with respect to a specific set. In the analysis we use an unspecified constant, denoted c_1. The constant δ (in the definition of Δ_v) is determined in terms of c_1 (in fact $\delta = O(c_1)$ will do, $c_1 \geq 2$ suffices for the results of the current subsection and $c_1 \geq 10$ suffices for all the results of this section).

Lemma 16 : *Let $S \subseteq U$ be finite sets, with $|S| = k$. Then, for every $c_1 > 0$ there exists δ, so that a uniformly selected bipartition of U is Δ_v-biased with respect to S with probability $\geq 1-(p_v/k)^{c_1}$.*

proof: We consider two cases corresponding to the two conditions of Definition 1. By Chernoff's Bound, the probability that a uniformly selected partition fails condition (1) in Definition 1 (with respect to a set S with $k \geq \Delta_v^4$) does not exceed

$$2\exp\{-2(k^{-1/4})^2 \cdot k\} = 2\exp\{-2k^{1/2}\} \tag{10}$$

Using $k \geq (\delta \log_2(1/p_v))^4$, we upper bound Eq. (10) by

$$\exp\{-k^{1/2}\} \cdot \exp\{(\delta \cdot \log_2(1/p_v))^2\}$$

which for sufficiently large δ (or $1/p_v$) yields the desired bound (of $(p_v/k)^{c_1}$). Similarly, the probability that condition (2) is not satisfied by a random partition is bounded by

$$2\exp\{-2(1/20)^2 \cdot k\} = 2\exp\{-k/200\} \tag{11}$$

Using $k > \delta \log_2(1/p_v)$ and $\delta \geq 400c_1$, we upper bound Eq. (11) by

$$\exp\{-k/400\} \cdot \exp\{c_1 \log_2(1/p_v)\}$$

which for sufficiently large δ (or $1/p_v$) yields again the desired bound. ■

Proposition 17 (existence of value-balance partitions): *Let the generic protocol run for i rounds, using only partitions which satisfy all value-balance properties* w.r.t. all values in *$Range(f)$. Let $X_i \times Y_i$ be the residual matrix after these i rounds. Then there exist a row partition* (of X_i) *and a column partition* (of Y_i) *that satisfy all value-balance properties* w.r.t. all values. *Furthermore, for every $v \in Range(f)$, all but a $p_v^{c_1-1}$ fraction of the possible partitions satisfy all v-balance properties.*

proof: We consider only row-partitions, the proof for column-partitions being identical. Let $v \in \text{Range}(f)$. For $|X_i| < \Delta_v$ every non-trivial partition will do, so henceforth we assume $|X_i| \geq \Delta_v$. Lemma 16, yields an upper bound on the probability that a uniformly chosen partition of X_i violates one of the v-balance properties. For each property, we multiply the number of sets considered by the probability that a uniformly selected bipartition of X_i is not Δ_v-biased with respect to an individual set. An obvious (lower) bound on the size of an individual set considered is Δ_v, but in some cases better lower bounds hold. For each of the eight properties, we prove an upper bound of $p_v^{c_1-1}/8$ on the probability that a uniformly chosen partition violates the property.

- Property (P0) is violated with probability at most $|X_i| \cdot (p_v/|X_i|)^{c_1}$ which can be bounded by $p_v^{c_1-1}/8$.
- Property (P1) is violated with probability at most $|Y_i| \cdot \max_{y \in Y_i}\{(p_v/\#_v(X_i,y))^{c_1}\}$. In case $|Y_i| < \Delta/p_v$, this probability is easily bounded by $(p_v/\Delta_v)^{c_1-1} < p_v^{c_1-1}/8$. Otherwise, we

argue as follows. Since Property (P1) was satisfied in previous rounds, it follows (as in Lemma 6) that

$$\begin{aligned}\#_v(X_i, y) &\geq 2^{-i-1} \cdot \#_v(X_0, y) \\ &\geq \frac{p_v}{8} \cdot |X_i|\end{aligned}$$

and so Property (P1) is violated with probability at most $|Y_i| \cdot (8/|X_i|)^{c_1}$. Using Property (P0) for the previous rounds we get $|X_i| \geq |Y_i|/4$ and again obtain a bound of $O((p_v/\Delta_v)^{c_1-1}) < p_v^{c_1-1}/8$.

- For Property (P2), we need only consider the case $|X_i| < (2/p_v)$. In this case, Property (P2) is violated with probability at most $(|X_i| + |Y_i| + 1) \cdot (p_v/\Delta_v)^{c_1}$ which is bounded by $O(p_v^{c_1-1}/\Delta_v^{c_1}) < p_v^{c_1-1}/8$. Property (P7) is dealt similarly, but the bound here is $O(p_v^{c_1-1}/\Delta^{c_1-2}) < p_v^{c_1-1}/8$.
- For Property (P6) we need to consider only $j \leq \ell+1$ and $y \in Y_i$ such that $\#_v(R^{j+\ell} \cap X_i, y) \geq \max\{\Delta_v, (p_v/4\Delta_v) \cdot |Y_i|\}$. Let us denote the set of these pairs by P_i. Then, Property (P6) is violated with probability at most

$$\begin{aligned}\sum_{(j,y)\in P_i} \left(\frac{p_v}{\#_v(R^{j+\ell} \cap X_i, y)}\right)^{c_1} &\leq \left(\frac{p_v}{\Delta}\right)^{c_1-1} \cdot \left(|P_i| \cdot \frac{p_v}{(p_v/4\Delta_v) \cdot |Y_i|}\right) \\ &\leq \left(\frac{p_v}{\Delta}\right)^{c_1-1} \cdot \frac{(\ell+1) \cdot |Y_i|}{|Y_i|/4\Delta} \\ &< \frac{p_v^{c_1-1}}{8}\end{aligned}$$

- For the remaining properties (i.e., (P3), (P4) and (P5)) we have a total of $O(\log^2(1/p_v))$ sets and so the bound holds easily.

Thus, the probability that a random partition of X_i violates some property with respect to the value v is at most $p_v^{c_1-1}$. The main claim of the proposition follows by summing the bounds obtained for all possible v's, and using $c_1 \geq 2$. ■

Combining Theorem 15 and Proposition 17, we get

Corollary 18 (existence of a protocol meeting the lower bound): *Let f be as in Theorem 15. Then, there exists a (deterministic) two-party protocol for computing the function f, so that for every $v \in Range(f)$, if one party plays honestly, then the outcome of the protocol is v with probability at most $O(\log^6(1/p_v) \cdot \max\{q_v, \sqrt{p_v}\})$.*

4.3 Efficient Protocols Meeting the Lower Bounds

The protocols guaranteed by Corollary 18 are not efficient. In particular, merely specifying the partitions used by the protocol takes space that is exponential in size of the inputs, not to mention that the proof is nonconstructive and that a naive construction would require double exponential time. An efficient implementation of the protocols is achieved by using partitions which can be specified by polynomially many bits. These partitions will not be hard-wired into the protocol but rather selected online by the two parties. Namely, at the outset of each step, the parties perform a sampling protocol to select a partition for that step. The partition is specified by an $m^{\rm th}$ degree ($m = {\rm poly}(n)$) polynomial over the field $F \stackrel{\rm def}{=} GF(2^n)$ and a fixed partition of the elements of F into two equal parts F^0 and F^1. For example, suppose polynomial P (over F) is chosen to specify a partition of Y_i, then Y_i^σ is defined as the set of all points $y \in Y_i$ satisfying $P(y) \in F^\sigma$. This plan is materialized via a two-party protocol for sampling these partitions and a proof that every partitions selected (for the generic protocol) by the sampling protocol satisfies all v-balance properties with probability at least $1 - p_v$. To this end we first bound the probability that, for an appropriately chosen $m = {\rm poly}(n)$, a random $m^{\rm th}$ degree polynomial induces a partition that fails to satisfy some v-balance properties. Next, we present a two-party protocol for sampling l-bit strings and bound the advantage of each party towards any set as a function of the density of that set.

Terminology: Partitions induced by $(\delta n)^4$-degree polynomials are hereafter called polynomial-partitions. We modify these partitions so that they are never trivial (e.g., by replacing each trivial partition by a fixed non-trivial partition). Recall that Property (P0) forbids trivial partitions, except if the universe is a singleton. The modification is introduced to guarantee this.

Bounding the Probability of Non-Balanced Polynomial-Partitions

We start by bounding the probability for a random polynomial-partition to fail some v-balance property.

Lemma 19 : *For every $c_1 > 0$ there exists δ, so that for every set S of cardinality k, a uniformly selected polynomial-partition is not Δ_v-biased with respect to S with probability at most $(p_v/k)^{c_1}$.*

proof: The modification described in the Terminology (above) can only decrease the probability that a partition is not Δ-biased (w.r.t. any set S). Thus, it suffices to analyze the distribution of unmodified polynomial-partitions.

A $2t^{\rm th}$ moment argument easily shows that if $x_1, x_2, ..., x_k$ are m-wise independent random variables uniformly distributed in $\{0,1\}$ then ${\rm Prob}(|\sum_{i=1}^k x_i - \frac{k}{2}| > B) < (\frac{\sqrt{k}t}{B})^{2t}$, for every $t \leq m/2$. Therefore, the probability for a uniformly chosen polynomial-partition to fail condition (1) in Definition 1 does not exceed

$$\left(\frac{\sqrt{k} \cdot t}{k^{3/4}}\right)^{2t} = \left(\frac{t}{k^{1/4}}\right)^{2t} \tag{12}$$

for any $t \leq (\delta n)^4/2$. We now use Eq. (12) with two different settings for t. First we set $t = \Delta_v/2$ (Since $p_v \geq 2^{-n}$, it follows that $\Delta_v \leq \delta \cdot 2n$ and this t is indeed smaller than $(\delta n)^4/2$) and using $k \geq \Delta^4$, we bound Eq. (12) by

$$\left(\frac{\Delta_v/2}{\Delta_v}\right)^{\Delta_v} = p_v^\delta < p_v^{2c_1}$$

where the last inequality comes from $\delta \geq 2c_1$. Secondly, we set $t = 8c_1$, and bound Eq. (12) by

$$\left(\frac{8c_1}{k^{1/4}}\right)^{16c_1} = \left(\frac{(8c_1)^8}{k^2}\right)^{2c_1} < \frac{1}{4} \cdot k^{-2c_1}$$

where we have used $k \geq \Delta^4 \geq 4 \cdot (8c_1)^8$. Multiplying these two bounds, we bound Eq. (12) by

$$\sqrt{p_v^{2c_1} \cdot \frac{k^{-2c_1}}{4}} = \frac{1}{2} \cdot (p_v/k)^{c_1}$$

as desired. To bound the probability for failure in condition (2), note that for $k \leq \Delta^4$ we have, $k \leq (\delta n)^4$ (as previously observed $\Delta_v \leq \delta n$). Thus, a uniformly selected polynomial-partition splits k elements exactly as a totally random partition and so the bound obtained for this case (i.e., for $k \leq \Delta^4$) in Lemma 16 holds also here. ■

Proposition 20 (polynomial-partition satisfy value-balance properties): *Fix $v \in Range(f)$, and consider an execution of the generic protocol with uniformly selected polynomial-partitions. Let π_i be the probability that the first failure of some v-balance property occurs on the i^{th} round. Then,*

$$\sum_{i \geq 1} \sqrt[4]{\pi_i} \leq O(\Delta_v \cdot p_v)$$

The mysterious choice of the 4^{th} roots will be clarified when we get to prove Theorem 22.
proof: It suffices, of course, to consider only row-partitions. Let $\pi_{i,t}$ be the probability that our first failed row-partition occured in round i and that Property (P_t) was violated (for some $0 \leq t \leq 7$ and $i \geq 1$). Clearly,

$$\begin{aligned} \sum_{i \geq 1} \sqrt[4]{\pi_i} &\leq \sum_{i \geq 1} \sqrt[4]{\sum_{t=0}^{7} \pi_{i,t}} \\ &\leq \sum_{t=0}^{7} \sum_{i \geq 1} \sqrt[4]{\pi_{i,t}} \end{aligned}$$

So it remains to bound, $\sum_{i \geq 1} \sqrt[4]{\pi_{i,t}}$, for each $t = 0, ..., 7$. Analogously to the proof of Proposition 17, we use Lemma 19 to upper bound the probability that a uniformly chosen polynomial-partition violates one of the v-balance properties. For each property, we multiply the number of sets considered by the probability that a uniformly selected polynomial-partition is not Δ_v-biased with respect to an individual set. An obvious (lower) bound on the size of an individual set considered is Δ_v, but in some cases better lower bounds hold. We now assume $c_1 \geq 10$.

- We upper bound the probability that Property (P0) is violated for the first time in the i^{th} round by $|X_{i-1}| \cdot (p_v/|X_{i-1}|)^{c_1}$. Letting $x_j := |X_j|$, we have:

$$\begin{aligned} \pi_{i,0} &\leq x_{i-1} \cdot (p_v/x_{i-1})^{c_1} && (13) \\ &\quad \text{where } x_j \geq \max\{|\Delta, |X_0|/2^{j-1}\} && (14) \end{aligned}$$

where the lower bound on x_j follows, since Property (P0) held in the previous rounds. Furthermore, if Property (P0) held in all first n rounds, then $|X_n| \leq \Delta$ and henceforth every non-trivial partition satisfies all properties vacuously. Therefore,

$$\begin{aligned}
\sum_{i \geq 1} \sqrt[4]{\pi_{i,0}} &= \sum_{i=1}^{n} \sqrt[4]{\pi_{i,0}} \\
&\leq \sum_{i=1}^{n} \sqrt[4]{x_{i-1} \cdot \left(\frac{p_v}{x_{i-1}}\right)^{c_1}} \\
&< \sum_{i=1}^{n} \frac{p_v}{x_{i-1}} \\
&< p_v \cdot \sum_{i=1}^{n} \frac{2^{i+2}}{2^n}
\end{aligned}$$

where the last inequality uses the lower bounds for the x_j's. It follows that $\sum_{i \geq 1} \sqrt[4]{\pi_{i,0}} = O(p_v)$.

- Adopting the analysis in the proof of Proposition 17, we know that the probability that the first failure is with Property (P1) in round i is at most $4|X_{i-1}| \cdot (p_v/|X_{i-1}|)^{c_1}$. Using the same analysis as above, we conclude $\sum_{i \geq 1} \sqrt[4]{\pi_{i,1}} = O(p_v)$.

- For Properties (P2) and (P7), we need only consider rounds i so that $|X_i| < (2/p_v)$. Using the analysis in the proof of Proposition 17, we bound the probability that the partition in such a round violates Property (P2) (resp., (P7)) by $O(p_v^{c_1-1}/\Delta_v^{c_1})$ (resp., $O(p_v^{c_1-1}/\Delta^{c_1-2})$). The bound on $\sum_{i \geq 1} \sqrt[4]{\pi_{i,t}}$, for $t = 2, 7$, follows, since there are at most Δ such rounds.

- Following the analysis in the proof of Proposition 17, we consider for Property (P6) only $j \leq \ell + 1$ and $y \in Y_i$ such that $\#_v(R^{j+\ell} \cap X_i, y) \geq \max\{\Delta_v, (p_v/4\Delta_v) \cdot |Y_i|\}$. Let us denote the set of these pairs by P_i. The probability that our first violation is on round i and Property (P6) is being violated, is at most

$$\begin{aligned}
\sum_{(j,y) \in P_i} \left(\frac{p_v}{\#_v(R^{j+\ell} \cap X_i, y)}\right)^{c_1} &\leq |P_i| \cdot \left(\frac{p_v}{\Delta}\right)^{(c_1+1)/2} \cdot \left(\frac{p_v}{(p_v/4\Delta_v) \cdot |Y_i|}\right)^{(c_1+1)/2} \\
&\leq ((\ell+1) \cdot |Y_i|) \cdot \left(\frac{p_v}{\Delta}\right)^4 \cdot \left(\frac{4\Delta}{|Y_i|}\right)^5 \\
&< \left(\frac{\Delta_v \cdot p_v}{|Y_i|}\right)^4
\end{aligned}$$

Using the same analysis as for Property (P0), we obtain $\sum_{i \geq 1} \sqrt[4]{\pi_{i,6}} < \Delta_v \cdot p_v$.

- For the remaining properties (i.e., (P3), (P4) and (P5)) we have a total of $O(\log^2(1/p_v))$ sets and so we can handle each of these sets separately. Consider, for example, the set $R^{j+\ell}$ from the definition of Property (P4). The row-partition of round $i+1$ violates the balance property on this set with probability at most $\left(\frac{p_v}{|R^{j+\ell} \cap X_i|}\right)^{c_1}$. Setting $x_i \stackrel{\text{def}}{=} |R^{j+\ell} \cap X_i|$, we can apply the same analysis as applied to Eq. (13), except that here we use Property (P4) for the previous rounds. The desired bound for $\sum_{i \geq 1} \sqrt[4]{\pi_{i,t}}$ follows, for $t = 3, 4, 5$.

Having shown that $\sum_{i\geq 1}\sqrt[4]{\pi_i} < \Delta_v \cdot p_v$, for each $t = 0, ..., 7$, the proposition follows. ■

Protocol for String Sampling

We now present a two-party protocol for sampling l-bit strings and bound the advantage of each party towards any set as a function of the set's density. The protocol is a simplification of the protocol for computing a function. The parties proceed in l rounds. In each round one party should select a random polynomial-partition of the residual sample space and the other party should flip a coin to select a side of this partition. In the next round the parties switch roles. All partitions selected by each party must divide the residual space into two sets of equal cardinality. Specifically, the partition is defined by a linear combination of the bits in the representation of the sample point. Following is the code of the protocol (the parties are called P_0 and P_1).

Round i:

- $P_{i \bmod 2}$ uniformly selects an l-dimensional binary vector v_i which is linearly independent of the vectors used in previous rounds, and sends v_i to the other party.
- $P_{(i+1) \bmod 2}$ uniformly selects $\sigma_i \in \{0,1\}$ and sends it to the other party.

Intuition: The residual sample space after round i consists of all l-dimensional binary vectors x so that $< x, v_j > = \sigma_j$ for every $j \leq i$ ($< \cdot, \cdot >$ is mod-2 inner product, and this residual set is an affine subspace).

Proposition 21 (analysis of the two-party sampling protocol): *Let $S \subseteq \{0,1\}^l$ be arbitrary and let $p \stackrel{\text{def}}{=} |S|/2^l$. If one of the parties that participate in the above protocol plays honestly, then the probability for the protocol's outcome to be in S is at most $O(p^{\frac{1}{4}})$.*

proof: Let U_i denote the residual sample space after round i; namely

$$U_i \stackrel{\text{def}}{=} \{x :< x, v_j > = \sigma_j \;\; \forall j \leq i\}$$

Let $S_i \stackrel{\text{def}}{=} S \cap U_i$ denote the residual target set ($U_0 = \{0,1\}^l$ and $S_0 = S$). We want to consider the cardinality of S_i as i grows (i.e., the execution proceeds) and treat differently "small" and "large" S_i. For "small" S_i we bound the probability of hitting S_i as $|S_i|$ times the probability of hitting any specific element. If S_i is "large", then with sufficiently high probability $|S_{i+1}| \approx |S_i|/2$ and hence the density, $|S_i|/|U_i|$, is approximately preserved. Details follow.

The following three claims do not depend on the residual sample space U_i. Thus, S_i (the residual target set after i rounds) can be considered fixed, too.

Claim 21.1: If the $(i+1)$-st partition is chosen by an honest player, then, with probability at least $1 - |S_i|^{-\frac{4}{5}}$:

$$\frac{|S_i|}{2} - |S_i|^{\frac{9}{10}} < |S_{i+1}| < \frac{|S_i|}{2} + |S_i|^{\frac{9}{10}},$$

regardless of the choice of σ_i.

proof: By hypothesis v_{i+1} is uniformly selected among the vectors which are linearly independent of $v_1, ..., v_i$. Instead, let us select v_{i+1} uniformly at random from the entire space Z_2^l. The additional partitions come from v_{i+1} in the linear span of $(v_1, ..., v_i)$, and thus induce a trivial partition on U_i, so the partition is only less likely to be balanced.

We show that with very high probability, even the partition induced by a uniformly chosen vector is quite balanced. For every $\sigma \in \{0,1\}$, we consider random variables ζ_s, $(s \in S_i)$ where $\zeta_s = 1$ if $< s, v_{i+1} >= \sigma$ and 0 otherwise. Since v_{i+1} is selected uniformly, each ζ_s is uniformly distributed in $\{0,1\}$. Furthermore, these random variables are pairwise independent, as long as $|U_i| \geq 2$ (i.e., the protocol did not terminate). Thus, we have

$$\text{Prob}\left(\left|\sum_{s\in S_i} \zeta_s - \frac{|S_i|}{2}\right| \geq |S_i|^{\frac{9}{10}}\right) < \frac{1}{4 \cdot |S_i|^{2\cdot\frac{9}{10}-1}}$$

and the claim follows. □

On the other hand,

Claim 21.2: If σ_{i+1} is selected by an honest player, then the expected cardinality of S_{i+1} is $|S_i|/2$.

The probability of hitting S_i is bounded by $|S_i|$ times the probability of hitting any specific element of S_i, so

Claim 21.3: With the above notation, the probability that the output of the protocol is in S (or, equivalently, in S_i) does not exceed $|S_i| \cdot 2^{-(l-i-1)/2}$.

proof: Clearly $|U_i| = 2^{l-i}$ and there remain $r \stackrel{\text{def}}{=} l - i$ rounds to termination, of which σ will be chosen by an honest player at least $\lfloor r/2 \rfloor$ times. Any $s \in S_i$ survives each such round with probability $\frac{1}{2}$, and is the output with probability at most $\cdot 2^{-\lfloor r/2 \rfloor}$, as claimed. □

In case $|S| < p^{-\frac{1}{2}}$ the proposition follows by using Claim 21.3; namely, the probability for output in S is bounded by

$$\begin{aligned} |S_0| \cdot 2^{-l/2} &= \sqrt{|S| \cdot \frac{|S|}{2^l}} \\ &= \sqrt{|S| \cdot p} \\ &< p^{\frac{1}{4}} \end{aligned}$$

So in what remains we consider the case $|S| \geq p^{-\frac{1}{2}}$. Let the protocol be executed for $t \stackrel{\text{def}}{=} \log_2 |S| - \frac{1}{2}\log_2(1/p) \geq 0$ rounds. In the rest of the proof we essentially show that, at this stage, $|S_t| \approx p^{-\frac{1}{2}}$. Using Claim 21.3 at this point, we obtain (again) the upper bound of $|S_t| \cdot 2^{-(l-t)/2} = p^{\frac{1}{4}}$ (using $l - t = l - \log_2 |S| + \frac{1}{2}\log_2(1/p) = (1 + \frac{1}{2}) \cdot \log_2(1/p)$).

We assume, without loss of generality, that the honest party picks the partitions at the even rounds. Also, there is no loss in assuming that his opponent plays a pure (i.e., deterministic) strategy. Since the honest party's strategy is fixed, the adversary's optimal move maximizes his expected payoff. On even-numbered rounds he selects one side of a partition presented by the honest player, while on round $2i + 1$ he selects a partition that is determined by a function Π_i. Formally, each of his moves is a function of the history of the execution, but this whole history is encoded by the current residual sample space. Thus, we may view each Π_i as a mapping $\Pi_i : 2^U \mapsto 2^U$, where U_{2i-2}, the residual sample space after $2i - 2$ rounds is partitioned into $(\Pi_i(U_{2i-2}), U_{2i-2} - \Pi_i(U_{2i-2}))$. Having fixed the adversary's strategy, the residual sample space after j rounds, U_j is a well-defined random variable. The following two sequences of random variables, depend now only on the coin tosses of the honest party:

1. π_i is the cardinality of $S \cap \Pi_i(U_{2i-2})$, for $i \geq 1$;
2. ζ_j is the cardinality of $S \cap U_j$, for $j \geq 0$ (where, $\zeta_0 = |S|$ is constant.)

The following facts are immediate by the definitions and Claims 21.1 and 21.3.

Claim 21.4: For every $i \geq 1$,

1. (*effect of round* $2i-1$: adversary presents partition)
   $\mathrm{Prob}(\zeta_{2i-1} = \pi_i) = \mathrm{Prob}(\zeta_{2i-1} = \zeta_{2i-2} - \pi_i) = \frac{1}{2}$.
2. (*effect of round* $2i$: adversary selects side)
   $|\zeta_{2i} - \frac{\zeta_{2i-1}}{2}| < \zeta_{2i-1}^{\frac{9}{10}}$ with probability at least $1 - \zeta_{2i-1}^{-\frac{4}{5}}$. Always $0 \leq \zeta_{2i} \leq \zeta_{2i-1}$.
3. (*termination*: as a function of the situation after $t \stackrel{\text{def}}{=} \log_2|S| - \frac{1}{2}\log_2(1/p)$ rounds) the protocol terminates with output in S with probability at most

$$\mathrm{Exp}(\zeta_t) \cdot 2^{-(l-t)/2} = \mathrm{Exp}(\zeta_t) \cdot p^{3/4}$$

   the expectation being over the coin tosses of the honest player in the first t rounds.

In proving Item (3), use $\mathrm{Exp}(\zeta_t \cdot 2^{-(l-t)/2}) = \mathrm{Exp}(\zeta_t) \cdot 2^{-(l-t)/2}$ and $l-t = l - \log_2|S| + \frac{1}{2}\log_2(1/p) = (1+\frac{1}{2}) \cdot \log_2(1/p)$. It remains to use Items (1) and (2) in order to prove:

Claim 21.5: Let $t \stackrel{\text{def}}{=} \log_2|S| - \frac{1}{2}\log_2(1/p)$ and suppose $t \geq 0$. Then

$$\mathrm{Exp}(\zeta_t) = O(p^{-1/2})$$

the expectation being over the coins tossed by the honest player in the first t rounds.

proof: Using Item (2) of Claim 21.4, we obtain

$$\begin{aligned} \mathrm{Exp}(\zeta_{2i+2}) &\leq \mathrm{Exp}\left(\frac{\zeta_{2i+1}}{2} + \zeta_{2i+1}^{\frac{9}{10}} + \zeta_{2i+1}^{-\frac{4}{5}} \cdot \zeta_{2i+1}\right) \\ &\leq \mathrm{Exp}\left(\frac{\zeta_{2i+1}}{2} + 2 \cdot \zeta_{2i+1}^{\frac{9}{10}}\right) \end{aligned}$$

On the other hand, using Item (1) of Claim 21.4, we obtain both

$$\mathrm{Exp}(\zeta_{2i+1}) = \frac{1}{2} \cdot \mathrm{Exp}(\zeta_{2i})$$

and

$$\mathrm{Exp}(\zeta_{2i+1}^{\frac{9}{10}}) = \frac{1}{2} \cdot \mathrm{Exp}(\pi_i^{\frac{9}{10}}) + \frac{1}{2} \cdot \mathrm{Exp}((\zeta_{2i} - \pi_i)^{\frac{9}{10}})$$

Combining the three (in)equalities, we get

$$\begin{aligned} \mathrm{Exp}(\zeta_{2i+2}) &\leq \frac{1}{4} \cdot \mathrm{Exp}(\zeta_{2i}) + \mathrm{Exp}(\pi_i^{\frac{9}{10}}) + \mathrm{Exp}((\zeta_{2i} - \pi_i)^{\frac{9}{10}}) \\ &< \frac{1}{4} \cdot \mathrm{Exp}(\zeta_{2i}) + 2 \cdot \mathrm{Exp}(\zeta_{2i}^{\frac{9}{10}}) \end{aligned}$$

For $0 < \alpha < 1$, the function x^α over $x \geq 0$ is concave, so we may apply Jensen's inequality, and conclude

$$\mathrm{Exp}(\zeta_{2i+2}) < \frac{1}{4} \cdot \mathrm{Exp}(\zeta_{2i}) + 2 \cdot \mathrm{Exp}(\zeta_{2i})^{\frac{9}{10}}$$

Setting $z_i \stackrel{\mathrm{def}}{=} \mathrm{Exp}(\zeta_{2i})$, a minor adaptation of Claim 5 yields $\mathrm{Exp}(\zeta_t) = O(\frac{\zeta_0}{2^t})$. Recall now that $t = \log_2 |S| - \frac{1}{2}\log_2(1/p))$ and $\zeta_0 = |S|$, to conclude the claim. □

The proposition follows. ■

Remark 1 : The above bound is not tight, but it suffices for the purpose of sampling partitions in the generic protocol (See the proof of Theorem 22). Much better protocols can be obtained - see Theorem 23. These (more complex) sampling protocols use the above protocol and the bound from Proposition 21 as a bootstrapping step. In our best sampling protocol, if one party plays honestly, the probability for the protocol protocol to land in any set of density p does not exceed $O(\sqrt{p})$.

Remark 2 : Our two-party sampling protocol is very similar to *interactive hashing*, a protocol, that was discovered independently by Ostrovsky et. al. [20]. However, in interactive hashing one party always picks the partition and the other always chooses the side. Also, interactive hashing terminates after $l-1$ (rather than l) rounds. Interactive hashing was invented for completely different purposes and consequently its analysis as in [20] (and subsequent studies), is very different from what appears above. Interactive hashing was used in implementing various types of commitment protocols (cf. [20, 18, 21, 10]).

Main Result

Combining Propositions 20 and 21 with Theorem 15, we get

Theorem 22 (efficient protocol meeting the lower bound): *There exists a (generic) two-party protocol, for evaluating an arbitrary bivariate function f. This protocol is performed by a pair of uniform probabilistic polynomial-time programs with a single oracle call to the function f and satisfies the following properties:*

- *If both parties play honestly and their inputs are x and y respectively, then the output is $f(x, y)$.*
- *For every value v in the range of f, if one party plays honestly then the outcome of the protocol is v with probability at most*

$$O(\log^6(1/p_v) \cdot \max\{q_v, \sqrt{p_v}\})$$

Furthermore, in case $q_v = p_v$, this bound can be improved to $O(\sqrt{p_v})$.

proof: The protocol is an implementation of the generic protocol where the partitions are determined by poly(n)-degree polynomials that are selected using the sampling protocol described above. This proves the first item. For the second item we consider the event in which during the execution of the protocol (with at least one party being honest) a partition was selected which does not satisfy all v-balanced properties. Using Propositions 20 and 21, the probability of this event is $O(\Delta_v \cdot p_v)$. (Here we use the fact that Proposition 20 bounds the sum of the fourth-root of the density of "bad" partitions.) In the complementary case, when every partition that is used satisfies all v-balance properties, Theorem 15 applies, and the main part of the second item follows.

A bound of $O(\sqrt{p_v}\log^2(1/p_v))$ for the special case of $q_v = p_v$ can be obtained by using Corollary 9 instead of Theorem 15. The better bound of $O(\sqrt{p_v})$ requires a slightly more careful analysis that we turn to perform.

We slightly change the classification of rounds as appearing in the motivating discussion (subsection 4.1). We first consider the situation after $i \stackrel{\text{def}}{=} n - \log_2(1/p_v) - 4\log_2 \Delta_v$ rounds. Following the ideas in the proof of Lemma 6 (and using Proposition 20 and 21), we first observe that, with probability $\geq 1-p_v$, the number of v-entries in each row (column) of the residual matrix (i.e., $\#_v(x,Y_i) \leq 2\Delta_v^4$, $\forall x \in X_i$). (Here and below the probability space is comprised of runs of the generic protocol in which polynomial-partitions are selected using the sampling protocol of Proposition 21.) Next, we consider the situation after an additional $\ell \stackrel{\text{def}}{=} \frac{1}{2}\log_2(1/p_v)$ rounds. Using similar ideas (this time following Lemma 8), we conclude that, with probability $\geq 1 - p_v$, the total number of v-entries in the entire residual matrix, is at most $(4\Delta_v^4+1)\cdot 2\Delta_v^4 < 9\Delta_v^8$ (i.e., $\#_v(X_{i+\ell},Y_{i+\ell}) < 9\Delta_v^8$). Furthermore, with probability at least $1-p_v$, the residual matrix at this point is of size approximately $\frac{\Delta_v^4}{\sqrt{p_v}}$ by $\frac{\Delta_v^4}{\sqrt{p_v}}$. In the original analysis, we did not try to argue that the number of v-entries in each row/column decreases during these additional ℓ rounds. But this is most likely to be the case as shown below.

Claim 22.1: There exists a constant c so that with probability at least $1-p_v$, after $i+\ell = n - \frac{1}{2}\log_2(1/p_v) - 4\log_2\Delta_v$ rounds, there are at most c v-entries in each residual row (resp., column) (i.e., $\#_v(x,Y_{i+\ell}) \leq c$, $\forall x \in X_{i+\ell}$).

proof: We consider again these additional ℓ rounds, assuming that previously (i.e., after i rounds) each residual row/column contains at most $2\Delta_v^4$ v-entries. We want to bound, for each individual row $x \in X_i$, the probability that $\#_v(x,Y_{i+\ell}) > c$. Say that a column partition is *good* if either there are fewer than c v-entries in the x-row, or each side of the partition contains at least one third of these entries. (In a *good* round, a good column partitions is performed). A *uniformly* selected polynomial partition fails to be good with probability that is exponentially small in the number of v-entries, since at this point, the degree of the polynomials that determine the partition exceeds the number of v-entries in row x. However, the polynomial partitions are selected using the sampling protocol of Proposition 21. As Proposition 21 states, the same remains valid also when using the sampling algorithm to select the partitions (at the cost of a different constant in the exponent). Therefore, there exists a constant c so that, as long as row x has more than c v-entries, the next round is good with probability at least 16/17 (a great underestimate for all but the very last rounds). On the other hand, if we go through at least $t \stackrel{\text{def}}{=} \log_{3/2}(2\Delta_v^4)$ good rounds, then row x has at most c v-entries. Thus $\#_v(x,Y_{i+\ell}) > c$ only if fewer than $t \ll \ell$ out of the last ℓ rounds are good, and the probability of this event is bounded above by

$$\binom{\ell}{t} \cdot (1/17)^{\ell-t} < (1/16)^{(1+\epsilon)\cdot\ell} = p_v^{(1+\epsilon)\cdot 2}$$

where $\epsilon > 0$ is some small constant, the inequality follows by $t = o(\ell)$ and the equality uses the definition of ℓ. Summing over all possible $x \in X_i$, the claim follows. $\square$

Combining Claim 22.1 with the discussion which precedes it, we conclude that after $i + \ell = n - \frac{1}{2}\log_2(1/p_v) - 4\log_2 \Delta_v$ rounds, with very high probability, the residual matrix contains at most $9\Delta_v^8$ entries of value v with at most c such entries in any row or column. Since we are seeking an $O(\sqrt{p_v})$ bound, we can, and will ignore those rare runs (of probability $O(p_v)$), for which this is not the case. Proceeding analogously to subsection 4.1, we could consider the situation after another $r = 4\log_2 \Delta_v$ rounds and bound by p_v the probability that after a total of $i + \ell + r = n - \frac{1}{2}\log_2(1/p_v)$ rounds the residual submatrix contains more than Δ_v entries of value v. This would yield a bound of $O(\Delta_v \cdot \sqrt{p_v})$ on the influence towards v. To obtain the better bound claimed above, we observe that it suffices to bound the *expected number* of v-entries in the residual matrix (rather than bounding the probability that too many v-entries remain). Specifically, we consider a standard coloring of the v-entries after $i+\ell$ rounds. This coloring uses at most $2 \cdot c + 1$ colors. Fixing one of these colors, we consider the next $r \stackrel{\rm def}{=} 4\log_2 \Delta_v$ rounds, and bound the expected number of the remaining v-entries. A *diagonal* is a set of entries in a matrix that has no more than a single element in common with any row/column.

Claim 22.2: Consider a diagonal D of at most $9\Delta_v^8$ entries in the residual matrix $(X_{i+\ell} \times Y_{i+\ell})$. Then the expected number of entries from D in the residual matrix $X_{i+\ell+r} \times Y_{i+\ell+r}$ is $O(1)$.

proof: It suffices to analyze a process in which $2r = 8\log_2 \Delta_v$ polynomial partitions, selected by the sampling protocol of Proposition 21, are applied to a space containing $9 \cdot \Delta_v^8$ elements so that after selecting each partition we proceed with the side containing more elements. Our claim is that the expected number of elements after applying these $2r$ partitions is $O(1)$. To prove this claim, let us consider first what happens after applying a single partition. Namely, let S be a subset (of some universe) and ζ be a random variable representing the number of S-elements in the S-heavier side (i.e., the side containing more S-elements) of a partition, selected by the sampling protocol. Clearly,

$$\mathrm{Exp}(\zeta) < \left[\frac{|S|}{2} + |S|^{3/4}\right] + \mathrm{Prob}\left(\zeta > \frac{|S|}{2} + |S|^{3/4}\right) \cdot |S|$$

For a uniformly selected polynomial-partition the probability that the S-heavy side contains more than $|S|/2 + |S|^{3/4}$ elements of S is exponentially small in $\sqrt{|S|}$ and by Proposition 21 the same holds (with a smaller constant in the exponent) when the polynomial-partition is selected by the sampling protocol. Thus, $\mathrm{Exp}(\zeta) < \frac{|S|}{2} + |S|^{3/4} + O(1)$. Hence, we have a sequence of random variables, $\zeta_0, \zeta_1, ..., \zeta_{2r}$, so that $\zeta_0 < 9\Delta_v^8$ and $\mathrm{Exp}(\zeta_i | \zeta_{i-1} = s) < \frac{s}{2} + s^{3/4} + O(1)$, for $i = 1, ..., 2r$. Manipulating the expectation operators (as in the proof of Claim 21.5), we conclude that $\mathrm{Exp}(\zeta_{2r}) = O(1)$ and the current claim follows. $\square$

Combining Claims 22.1 and 22.2, we conclude that with probability $1 - p_v$ we reach round $i + \ell + r = n - \frac{1}{2}\log_2(1/p_v)$ with an expected number of $O(1)$ entries of value v. Using the analysis of Corollary 9 (corresponding to stage 3 in the motivating discussion) we establish the claimed $O(\sqrt{p_v})$ bound and the theorem follows. $\blacksquare$

As stated in Remark 1, we have sampling protocols that improve on Proposition 21. This can be done either directly (with the techniques used in proving Theorem 22) or by applying Theorem 22 to any function f with $q_v = 2^{-l}$ ($\forall v \in \{0,1\}^l$). In either case, the resulting sampling protocols use the simple sampling protocol and the bound presented in Proposition 21 as a bootstrapping step.

Theorem 23 (a better two-party sampling protocol): *There exists a protocol for sampling $\{0,1\}^l$ that is performed by a pair of uniform probabilistic polynomial-time programs, so that: For every $S \subseteq \{0,1\}^l$ of density p, if one party plays honestly, the outcome of the protocol is in S with probability at most $O(\sqrt{p})$.*

proof (using the second alternative): Let $f: \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}^l$ satisfy $q_v = 2^{-l}$ for every $v \in \{0,1\}^l$. For example, $f(x,y) = x + y \bmod 2^l$, where x and y are viewed as residues mod 2^n (and $n \geq l$, say $n = l$). An honest party is supposed to select its input uniformly in $\{0,1\}^n$ and to invoke the protocol of Theorem 22. The current theorem follows from the (furthermore part of) Theorem 22, by considering the indicator function $\chi_S(v) = 1$ if $v \in S$ (and $\chi_S(v) = 0$ otherwise). Namely, we consider the function $g(x,y) \stackrel{\text{def}}{=} \chi_S(f(x,y))$ and take advantage of the fact that the protocol in Theorem 22 is generic (i.e., determines a pair of inputs (x,y) for the function independently of the function). ■

5 Towards the Multi-Party Case

We believe that the ideas developed in the two-party case will prove useful also for the multi-party case. However, even the problem of computing a 3-argument function by a 3-party protocol in the presence of one dishonest party is much more involved than the problem of computing a bivariate function by a 2-party protocol as in the previous section. A natural extension of our two-party protocol is to let each round consist of three steps (rather than two) and refer to three partitions of the three residual input spaces. In each step, a predetermined party announce in which side of the partition its input lies, and by doing so makes its residual input-space smaller. We believe that, this (generic) protocol when used with random partitions, nearly minimizes the advantage of any dishonest party, regardless of the function that is being computed. We also believe that this protocol nearly minimizes the advantage of any coalition of two dishonest players. However, this seems to require a much more complex analysis, and additional parameters of the function need to be taken into account. In particular, the advantage of a single adversary towards a value v depends not only on the density of v-entries in the entire function (denoted p_v above) and on the density of v-entries in the function restricted by the best input (denoted q_v). For example, a single party can influence any protocol for computing the function $f(x,y,z) = x+y+z \bmod N$ to produce output 0 (or any other residue mod N) with probability $N^{-2/3}$ (and the generic protocol can be shown to bound the advantage of a dishonest party to about this value). On the other hand, a single party can influence any protocol for computing the function $g(x,y,z) = x + y \bmod N$ to produce output 0 with probability $N^{-1/2}$ (and again the generic protocol meets this bound). However, both function have the same $p_v = q_v = 1/N$.

Another difficulty which arises in the context of multi-party protocols is that, when the number of parties is large, we cannot afford to let the parties reveal information in a predetermined order (as in the two-party case and the three-part case above). This difficulty is best demonstrated in the special case where each input is one bit (i.e., $\mathit{Domain}(f) = \{0,1\} \times \{0,1\} \cdots \times \{0,1\}$.)

Here, the influence of parties which are last to reveal their input is more substantial than the influence of parties which reveal their input first. This calls for choosing a random permutation to determine the order of playing. Thus, the role of a sampling protocol in the multi-party case is more fundamental than in the two-party situation. (Recall that in the two-party protocols, sampling was introduced only for increased efficiency.)

In this paper we confine ourselves to the presentation of an efficient fault-tolerant multi-party sampling protocol. Namely,

Theorem 24 (multi-party sampling protocol): *There exists an m-party sampling protocol that is performed by m (identical) uniform probabilistic polynomial-time programs, so that:*
For every set $S \subseteq \{0,1\}^l$, if $m-t$ parties play honestly, then the outcome of the protocol is in S with probability at most $O(\log(1/p) \cdot p^{1-O(\frac{t}{m})})$, where $p \stackrel{\rm def}{=} |S|/2^l$.

Our proof of Theorem 24 adapts the ideas used in Theorem 23 to the multi-party context. Namely, our protocol uses partitions which are in turn selected by a lower-quality sampling protocol. Specifically, the protocol proceeds in l rounds. In each round, the m parties first select at random (using a simpler sampling protocol) a poly$(n \cdot m)$-degree polynomial specifying a partition of the residual sample space, and next use the collective-coin tossing protocol of Alon and Naor [1] to choose one side of this partition. The sampling protocol used to choose poly(nm)-degree polynomials is similar except that the partitions are specified by linear transformations (as in the protocol of Proposition 21). These linear transformations are selected using a trivial sampling protocol which consists of selecting each bit individually by the collective-coin tossing protocol of Alon and Naor [1].

We prefer an alternative presentation of our proof, in which the construction of multi-party sampling protocols is reduced to the construction of sampling algorithms that use an SV-source as their source of randomness. Recall that an SV-source with parameter $\gamma \geq \frac{1}{2}$ (cf. [22]) is a sequence of Boolean random variables, $X_1, X_2, \ldots$, so that for each i and every $\alpha \in \{0,1\}^i$ and every $\sigma \in \{0,1\}$:

$$\mathrm{Prob}(X_{i+1}=\sigma | X_1 \cdots X_i = \alpha) \leq \gamma$$

Theorem 24 follows from

Proposition 25 (sampling with a SV-source): *For every constant γ, $\frac{1}{2} \leq \gamma < \frac{1}{\sqrt{2}}$, there exist a probabilistic polynomial-time algorithm, A_1, which on input 1^n uses any SV-source with parameter γ for its internal coin tosses and satisfies, for every sufficiently large n and every set $S \subseteq \{0,1\}^n$,*

$$\mathrm{Prob}(A_1(1^n) \in S) = O(\log(1/p) \cdot p^{\log_2(1/\gamma)})$$

where $p \stackrel{\rm def}{=} \frac{|S|}{2^n}$, and the probability is taken over an arbitrary SV-source with parameter γ.

In particular, for $\gamma = \frac{1}{2}(1+\epsilon)$, we have $\log_2(1/\gamma) = 1 - \log_2(1+\epsilon) \geq 1 - \frac{1}{\ln 2} \cdot \epsilon$. Thus, observing that the Alon-Naor protocol implements an SV-source with parameter $\gamma = \frac{1}{2}(1+O(\frac{t}{n}))$, we derive Theorem 24 as a corollary of Proposition 25. Furthermore, Proposition 25 yields an alternative way of recognizing BPP languages in polynomial-time using an arbitrary SV-source with parameter $\gamma < \frac{1}{\sqrt{2}} \approx 0.7$. Consider, without loss of generality, an algorithm A that using

n (perfect) random coins errs with probability at most $1/n$. In order to utilize A when only an SV-source is available, we first use algorithm A_1 (with the SV-source) to generate a "somewhat random" n-bit string, r, and then invoke algorithm A with the string r as a substitute for the n coins required by A. We stress that algorithm A is only invoked once. To analyze the performance of the new algorithm, let S be the set of coin-sequences on which A errs. By our hypothesis $|S| \leq 2^n/n$ and thus using Proposition 25 a string $r \in S$ is generated with probability at most $O(\log(n)/n^{\log_2(1/\gamma)}) < 1/3$. Thus, using an SV-source (with parameter $\gamma < \frac{1}{\sqrt{2}}$), our algorithm errs with probability at most 1/3.

proof: Following is a description of the **algorithm** A_1. The constant δ used in the description will be determined later (as a function of γ). On input 1^n, the algorithm proceeds in rounds, each round consisting of two steps. In the first step, algorithm A_1, uses a second sampling algorithm, denoted A_2, to select a succinct description of a "pseudorandom" partition of the residual sample space. In the second step, Algorithm A_1 uses the next bit of the SV-source to determine a side of this partition and so further restricts the residual sample space. We use two types of partitions. In the first $n - 4\log_2 \delta n$ rounds, algorithm A_1 uses partitions defined, as in Subsection 4.3, by a polynomial of degree $(\delta n)^4$ over $GF(2^n)$. In the remaining rounds, where the residual sample space is most likely to be smaller than $2(\delta n)^4$, algorithm A_1 uses partitions uniformly chosen from the set of all **perfectly balanced** partitions (i.e., bipartitions in which the cardinalities of the two sides are either equal, or differ by one). The two-step process is repeated until the residual sample space contains a unique element. We will see that algorithm $A_1(1^n)$ almost certainly halts after no more than $n+2$ rounds. (Longer executions can be truncated after $n+2$ rounds with an arbitrary output.)

We now turn to the description of **algorithm** A_2, which is invoked by $A_1(1^n)$ on input 1^m, where $m = (\delta n)^4 \cdot n$ for the first $n - 4\log_2 \delta n$ rounds of $A_1(1^n)$ and where m is the size of the residual sample space of A_1 later on. On input 1^m, algorithm A_2 proceeds in m rounds. In the i^{th} round, the algorithm uses a third sampling algorithm, denoted A_3, to selects a random m-dimensional binary vector v_i that is linearly independent of previously used vectors. Clearly, the candidate vectors constitute an $(m-(i-1))$-dimensional vector space over GF_2. The chosen vector partitions the residual sample space into two subsets of equal cardinality (as in Proposition 21). Algorithm A_2 uses the next bit of the SV-source to select a side of this partition.

Algorithm A_3, invoked by $A_2(1^m)$, on input 1^k (for $k = m, m-1, ..., 1$), is the trivial sampling algorithm which generates a sample point in $\{0,1\}^k$ by merely using the next k bits of the SV-source.

We now turn to the analysis of the sampling algorithm A_1. We first consider what happens if one replaces algorithm A_2 by an algorithm that uniformly selects the appropriate partitions (i.e., $(\delta n)^4$-degree polynomial for the first $n - 4\log_2 \delta n$ rounds and perfectly balanced partitions for later rounds). The analysis is done following the paradigm of the previous section. Namely, we first analyze the performance of the algorithm assuming it employs partitions which satisfy some combinatorial properties (cf., Claim 25.1), and next consider the probability that uniformly selected partitions satisfy these properties (cf., Claim 25.2).

Claim 25.1 (**A_1 with balanced partitions**): Let U_i be the residual sample space after round i, and $S_i \stackrel{\text{def}}{=} S \cap U_i$ ($U_0 \stackrel{\text{def}}{=} \{0,1\}^n$). Suppose that, for every i, algorithm A_1 partitions U_{i-1} in a way that is Δ-balanced with respect to S_{i-1} as well as to U_{i-1}. Furthermore, suppose that for every $i > n - 4\log_2 \Delta$, the i^{th} partition chosen for algorithm A_1 is perfectly balanced (i.e., $-1 \le 2|U_i| - |U_{i-1}| \le 1$.) Then

$$\text{Prob}(A_1(1^n) \in S) \le 2\Delta \cdot p^{\log_2(1/\gamma)}$$

In addition, $|U_{n-4\log_2 \delta n}| < 2(\delta n)^4$, provided that $\Delta \le \delta n$.

proof: The proof is analogous to the proof of Corollary 9. Using an argument analogous to one used in the proof of Lemma 6, we conclude that after $t \stackrel{\text{def}}{=} n - \log_2(1/p)$ rounds the residual sample space contains at most Δ elements of S (i.e., $|S_t| \le \Delta$). Actually, the argument only uses the hypothesis that the i^{th} partition is Δ-balanced with respect to S_{i-1}, for every $i \le t$, and is indifferent to the way in which the sides of the partitions are selected in these t rounds. Using the hypothesis that the i^{th} partition is Δ-balanced with respect to U_{i-1}, for every $i \le t$, we conclude that after these t rounds, the residual sample space contains at least $\frac{1}{2p}$ elements (i.e., $|U_t| \ge 1/2p$). Furthermore, using the hypothesis that also the following $r \stackrel{\text{def}}{=} \log_2(1/p) - 4\log_2 \Delta$ rounds use partitions which are Δ-balanced with respect to the residual sample space, we conclude that after $t + r = n - 4\log_2 \Delta$ rounds the residual sample space has cardinality at least $\frac{1}{2}\Delta^4$ (use Claim 5). Now, since all the remaining partitions are assumed to be perfectly balanced, there must be at least $l \stackrel{\text{def}}{=} (4\log_2 \Delta) - 1$ rounds until termination. We now return to the situation after t rounds, and consider the remaining rounds, which by the above are at least $r \stackrel{\text{def}}{=} t + l = \log_2(1/p) - 1$ in number. Since the side of the partition is selected by an SV-source with parameter γ, the probability that any specific element in U_t survives the remaining (i.e., at least r) rounds is at most γ^r. Thus, the probability that some element of S_t survives these rounds does not exceed

$$\begin{aligned} |S_t| \cdot \gamma^r &\le \Delta \cdot \gamma^{\log_2(1/p)-1} \\ &\le \Delta \cdot p^{\log_2(1/\gamma)} \cdot 2^{\log_2(1/\gamma)} \end{aligned}$$

But $\gamma \ge 1/2$, whence $\log_2(1/\gamma) \le 1$ and the main part of the claim follows.

The additional part (i.e., $|U_{n-4\log_2 \delta n}| < 2(\delta n)^4$) follows easily by using Claim 5. □

Claim 25.2 (A_1 – probability of balanced partitions): For every $\epsilon > 0$ there exists a $\delta > 0$ so that the following holds. Let π_i denote the probability that a *uniformly chosen* partition for round i is not $\delta \cdot \log_2(1/p)$-balanced with respect to either S_{i-1} or U_{i-1}. Then,

$$\sum_{i \ge 1} \pi_i^\epsilon < p$$

As in Proposition 20, it is very useful for the sequel (though, admittedly, not very natural) to raise the probabilities to the ϵ^{th} power.

proof: For $i \le n - 4\log_2 \delta n$, the proof is identical to the simpler cases (e.g., Properties (P0) and (P1)) considered in the proof of Proposition 20. For $i > n - 4\log_2 \delta n$, we observe that the probability of any event, assuming a uniformly selected *perfectly-balanced* partition is at most $\sqrt{|U_{i-1}|}$ times larger than its probability assuming a uniformly selected partition. Since the argument of Proposition 20, can tolerate such factors, the claim follows also for $i > n - 4\log_2 \delta n$. □

Combining Claims 25.1 and 25.2, we conclude that it suffices to show that for some constant $\epsilon > 0$ and for any set of "bad" partitions, $B \subseteq \{0,1\}^m$, the probability that $A_2(1^m)$ produces an output in B is at most $(|B|/2^m)^\epsilon$. Once this is done, the proposition follows by considering $B^{(i)}$, the set of partitions which are not $\delta \cdot \log_2(1/p)$-balanced with respect to either S_{i-1} or U_{i-1}, and noting that $\delta \cdot \log_2(1/p) \leq \delta n$ (which guarantees that in the last $4\log_2(\delta \log_2(1/p))$ rounds perfectly-balanced partitions are used). Namely,

$$\begin{aligned}
\text{Prob}(A(1^n)\in S) &< \text{Prob}(A(1^n)\in S|\forall i\ A(1^m)\not\in B^{(i)}) \\
&\quad +\text{Prob}(\exists i \text{ s.t. } A(1^m)\in B^{(i)}) \\
&< 2\delta\log(1/p)\cdot p^{\log_2(1/\gamma)} + \sum_{i\geq 1}\left(\frac{|B^{(i)}|}{2^m}\right)^\epsilon \\
&< 3\delta\log(1/p)\cdot p^{\log_2(1/\gamma)}
\end{aligned}$$

where the second inequality is based on Claim 25.1 and our hypothesis concerning A_2 and the last inequality follows from Claim 25.2. Also note that Claim 25.1 guarantees that the residual sample space after $n-4\log_2(\delta n)$ rounds has size at most $\text{poly}(n)$ whence it is possible to represent and generate random partitions of it. Thus, we turn to the analysis of algorithm A_2. Recall that our goal is to show that for some ϵ, (that depends on γ), and for every $B \subseteq \{0,1\}^m$ of cardinality $q \cdot 2^m$,

$$\text{Prob}(A_2(1^m)\in B) = O(q^\epsilon) \tag{15}$$

Let $\epsilon \stackrel{\text{def}}{=} \log_2(1/\gamma) - \frac{1}{2} > 0$ and $\beta \stackrel{\text{def}}{=} \frac{1}{1+\epsilon} < 1$ (recall that $\gamma < \frac{1}{\sqrt{2}}$ is assumed). Also, $\epsilon \leq \frac{1}{2}$ and $\beta \geq \frac{2}{3}$, since $\gamma \geq \frac{1}{2}$. Henceforth, we fix an arbitrary set $B \subseteq \{0,1\}^m$ and let $q \stackrel{\text{def}}{=} \frac{|B|}{2^m}$ (as above). We separately analyze the performance of A_2 throughout the first t rounds (hereafter referred to as phase 1), and in the remaining $m-t$ rounds (phase 2), where

$$t \stackrel{\text{def}}{=} \max\{0, m - \frac{2\beta}{2\beta-1}\log_2(1/q)\} \tag{16}$$

Let B_i denote the residual set of bad polynomials after i rounds of algorithm A_2 (e.g., $B_0 = B$).

Claim 25.3 (A_2 – phase 1):

$$\text{Prob}(|B_t| > 2q\cdot 2^{m-t}) \leq O(q^{2\epsilon})$$

Remark: By definition of t, we have $m-t = \frac{2\beta}{2\beta-1}\cdot\log_2(1/q)$ and $q\cdot 2^{m-t} = 2^{(m-t)/2\beta}$.

proof: For every i, let $b_i \stackrel{\text{def}}{=} \frac{|B|}{2^i}$. Our plan is to prove that with very high probability, $|B_i| \approx b_i$ for every $i \leq t$, which would establish our claim. We consider the first time when $|B_i| \not\approx b_i$. Thus, the probability that $|B_t| > 2q\cdot 2^{m-t}$ is bounded above by

$$\text{Prob}\left[\exists i < t : \left(\left||B_{i+1}| - \frac{|B_i|}{2}\right| > |B_i|^{\frac{1+\beta}{2}}\right) \wedge \left(\forall j < i : \left||B_{j+1}| - \frac{|B_j|}{2}\right| \leq |B_j|^{\frac{1+\beta}{2}}\right)\right]$$

Now, using Chebyshev's Inequality (as in the proof of Proposition 21), we can show that *for a uniformly chosen random linear partition*,

$$\text{Prob}\left(\left||B_{i+1}| - \frac{|B_i|}{2}\right| > |B_i|^{\frac{1+\beta}{2}}\right) < \frac{1}{|B_i|^\beta}$$

Call a linear partitions for round $i+1$ **bad**, if $\left||B_{i+1}| - \frac{|B_i|}{2}\right| > |B_i|^{\frac{1+\beta}{2}}$. We now know that the number of bad partitions is bounded by $\frac{1}{|B_i|^\beta} \cdot 2^{m-i}$. We need to bound the probability that $A_3(1^{m-i})$ selects a bad partition in round $i+1$. The union bound, the definition of A_3 and Claim 5 (for $|B_i|$), imply

$$\begin{aligned} \mathrm{Prob}(A_3(1^{m-i}) \text{ is bad}) &\leq \frac{2^{m-i}}{|B_i|^\beta} \cdot \gamma^{m-i} \\ &< 2 \cdot \gamma^{m-i} \cdot \frac{2^{m-i}}{b_i^\beta} \end{aligned}$$

where the last inequality uses our assumption that all previous partitions are good (whence for each $j < i$, $|B_{j+1}| > \frac{|B_j|}{2} - |B_j|^{\frac{1+\beta}{2}}$ and consequently $|B_i| > b_i/2$). Since $b_i = 2^{t-i} \cdot b_t$ and $b_t^\beta = (q2^{m-t})^\beta = 2^{(m-t)/2}$ (see remark above), we get

$$\begin{aligned} \mathrm{Prob}(A_3(1^{m-i}) \text{ is bad}) &< 2 \cdot \gamma^{m-i} \cdot \frac{2^{m-i}}{2^{(t-i)\beta} \cdot 2^{(m-t)/2}} \\ &= 2 \cdot \gamma^{m-i} \cdot 2^{(m-i) - \frac{m-t}{2} - \beta(t-i)} \\ &= 2 \cdot \left(\gamma\sqrt{2}\right)^{m-i} \cdot 2^{-(\beta - \frac{1}{2})\cdot(t-i)} \end{aligned}$$

Letting $\rho \stackrel{\text{def}}{=} \gamma \cdot \sqrt{2} < 1$ (as $\gamma < \frac{1}{\sqrt{2}}$) and using $m - i \geq m - t > 2\log_2(1/q)$ (as $m - t = \frac{2\beta}{2\beta - 1}\log_2(1/q)$ and $\beta < 1$), we get

$$\begin{aligned} \mathrm{Prob}(A_3(1^{m-i}) \text{ is bad}) &< 2 \cdot \rho^{2\log_2(1/q)} \cdot 2^{-(\beta - \frac{1}{2})\cdot(t-i)} \\ &= 2 \cdot q^{2\log_2(1/\rho)} \cdot 2^{-(\beta - \frac{1}{2})\cdot(t-i)} \\ &= a \cdot b^{t-i} \end{aligned}$$

where $a \stackrel{\text{def}}{=} 2q^{2\log_2(1/\rho)}$ and $b \stackrel{\text{def}}{=} 2^{-(\beta-\frac{1}{2})} < 1$ (as $\beta > \frac{1}{2}$). Hence, the probability that A_3 chooses a bad partition for some round i, throughout phase 1, is bounded by $\sum_{i=1}^{t} a \cdot b^{t-i} < \frac{a}{1-b}$. Using $\epsilon = \log_2(1/\gamma) - \frac{1}{2} = \log_2(1/\rho) \leq \frac{1}{2}$ and $\beta = \frac{1}{1+\epsilon}$, we get

$$\begin{aligned} \frac{a}{1-b} &= \frac{2q^{2\epsilon}}{1 - 2^{-\frac{1-\epsilon}{2(1+\epsilon)}}} \\ &\leq \frac{2q^{2\epsilon}}{1 - 2^{-1/6}} \\ &< 20 \cdot q^{2\epsilon} \end{aligned}$$

and the claim follows.□

Claim 25.4 (A_2 – **phase 2**): Let B_t be the residual target set after t rounds and consider an execution of the $m - t$ remaining rounds. Suppose that $|B_t| \leq 2b_t$, where $b_t \stackrel{\text{def}}{=} \frac{|B|}{2^t}$ (as in Claim 25.3). Then the probability that $A_2(1^m)$ terminates with output in B_t is at most $2q^\epsilon$.

proof: We consider the executions of rounds $t+1$ through m. Regardless of which linear partitions are used in the remaining $m-t$ rounds, the probability that a particular element of B_t is output by $A_2(1^m)$ is bounded by γ^{m-t}. Hence,

$$\begin{aligned}
\text{Prob}(A_2(1^m) \text{ hits } B_t) &\leq |B_t| \cdot \gamma^{m-t} \\
&\leq 2b_t \cdot \gamma^{m-t} \\
&= 2 \cdot 2^{\frac{m-t}{2\beta}} \cdot \gamma^{m-t} \\
&= 2 \cdot \left(\gamma \cdot 2^{\frac{1}{2\beta}}\right)^{m-t}
\end{aligned}$$

Setting (as before) $\rho = \gamma\sqrt{2}$, and using $\epsilon = \log_2(1/\rho)$ and $\beta = \frac{1}{1+\epsilon}$, we get $2^{\frac{1}{2\beta}} = \sqrt{2/\rho}$. Hence, using again $\rho < 1$ and $m - t > 2\log_2(1/q)$, we get

$$\begin{aligned}
\text{Prob}(A_2(1^m) \text{ hits } B_t) &\leq 2 \cdot \left(\gamma \cdot \sqrt{\frac{2}{\rho}}\right)^{m-t} \\
&< 2 \cdot \sqrt{\rho}^{\,2\log_2(1/q)} \\
&= 2q^{\log_2(1/\rho)} \\
&= 2q^{\epsilon}
\end{aligned}$$

and the claim follows. □

Combining Claims 25.3 and 25.4, we have established Eq. (15) and the proposition follows. ■

Remark 3 Actually, the result of Proposition 25 can be improved using a slightly more careful analysis of the algorithm A_1 provided in the above proof. The improved analysis is analogous to the proof of the tighter bound for the case $q_v = p_v$ of Theorem 22. Namely, we replace Claims 25.1 and 25.2 by the following three claims. *In the first two claims we assume that algorithm A_2 satisfies* Eq. (15).

claim 1: *With probability at least $1-p$, after $i \stackrel{\text{def}}{=} n - \log_2(1/p) - 4\log_2(\delta\log_2(1/p))$ rounds the residual sample space contains at most $2(\delta\log_2(1/p))^4$ elements of S; namely,*

$$\text{Prob}(|S_i| > 2(\delta\log_2(1/p))^4) < p$$

claim 2: *Consider an arbitrary subset S' of U_t so that $|S'| \leq 2(\delta\log_2(1/p))^4$. Then the expected number of elements of S' which survive an additional number of $4\log_2(\delta\log_2(1/p))$ rounds is bounded above by $O(1)$.*

claim 3: *Let $t \stackrel{\text{def}}{=} n - \log_2(1/p)$. Then,*

$$\text{Prob}(A(1^n) \in S) \leq \text{Exp}(|S_t|) \cdot \gamma^{\log_2(1/p)-1}$$

(Here, we do use a part of the proof of Claim 25.1 to assert that with probability $1-p$ the protocol does not terminate before $n-1$ rounds.)

Consequently, we get

- *For every constant γ, $\frac{1}{2} \leq \gamma < \frac{1}{\sqrt{2}}$, the algorithm A_1 appearing in the proof of Proposition 25 satisfies, for every set $S \subseteq \{0,1\}^n$,*

$$\mathrm{Prob}(A_1(1^n) \in S) = O(p^{\log_2(1/\gamma)})$$

  where $p \stackrel{\mathrm{def}}{=} \frac{|S|}{2^n}$, and the probability is taken over an arbitrary SV-source with parameter γ.

- Theorem 24 can be improved analogously. Namely, *for every set $S \subseteq \{0,1\}^l$, if $m-t$ parties plays honestly then the outcome of the protocol is in S with probability bounded above by $O(p^{1-O(\frac{t}{m})})$, where $p \stackrel{\mathrm{def}}{=} |S|/2^l$.*

Acknowledgment

Oded Goldreich would like to thank the Computer Science Department of the Hebrew University for providing him shelter in times of war; surprisingly some shelters are more pleasant than home! The authors wish to thank the Computer Science Department of Tel-Aviv University for use of its computing facilities.

References

[1] Alon N., and M. Naor, "Coin-flipping games immune against linear-sized coalitions", *SIAM J. on Computing*, to appear. Extended abstract in *Proc. of 31st FOCS*.

[2] Beaver D., and S. Goldwasser, "Distributed Computation with Faulty Majority ", *FOCS*, 1989.

[3] Ben-or M., S. Goldwasser, and A. Wigderson, "Completeness Theorems for Non-Cryptographic Fault-Tolerant Distributed Computation", *Proc. of 20th STOC*, 1988, pp. 1–10.

[4] Ben-or M., and N. Linial, "Collective coin flipping", *Randomness and Computation*, (S. Micali, ed.), JAI Press, pp. 91–115, 1989.

[5] Ben-Or, M., Linial, N., Saks, M., "Collective coin flipping and other models of imperfect randomness", *Colloq. Math Soc. Janos Bolyai* no. 52, Combinatorics Eger 1987, pp. 75–112.

[6] Chor, B., and O. Goldreich, "Unbiased Bits from Sources of Weak Randomness and Probabilistic Communication Complexity", *SIAM J. Computing*, Vol. 17, pp. 230–261, April 1988.

[7] Chaum, D., C. Crepeau, and I. Damgard, "Multiparty Unconditionally Secure Protocols", *Proc. of 20th STOC*, 1988, pp. 11–19.

[8] Chor, B., and E. Kushilevitz, "A Zero-One Law for Boolean Privacy", *SIAM Jour. on Disc. Math.*, Vol. 4, Feb. 1991, pp. 36–47.

[9] J. Cooper and N. Linial, Fast perfect-information leader election protocol with linear immunity, *Combinatorica, accepted for publication.* Extended abstract: *25th Symposium on the Theory of Computing, San Diego, 1993, 662 - 671.*

[10] I. Damgård, "Interactive Hashing can Simplify Zero-Knowledge Protocol Design Without Computational Assumptions", *Advances in Cryptology – Proceedings of Crypto93,*

[11] Galil, Haber, and Yung., "Cryptographic Computation: Secure Faulty-Tolerant Protocols and the Public Key Model", *Advances in Cryptology – Proceedings of Crypto89,*

[12] Goldreich, O., S. Goldwasser, and N. Linial, "Fault-tolerant Computation in the Full Information Model", TR-682, Computer Science Dept., Technion, Haifa, Israel, July 1991.

[13] Goldreich, O., S. Micali, and A. Wigderson, "Proofs that Yield Nothing but their Validity and a Methodology for Cryptographic Protocol Design", *JACM*, Vol. 38, July 1991, pp. 691–729. Extended abstract in *27th FOCS*, 1986, pp. 174–187.

[14] Goldreich, O., S. Micali, and A. Wigderson, "How to Play Any Mental Game", *Proc. of 19th STOC*, 1987, pp. 218–229.

[15] Goldwasser, S. and L.A. Levin, "Fair Computation of General Functions in Presence of Immoral Majority", *Advances in Cryptology – Proceedings of Crypto90*,

[16] J. Kahn, G. Kalai and N. Linial, "The influence of variables on boolean functions", *29th FOCS*, 1988, pp. 68–80.

[17] Kilian, J., "Founding Cryptography on Oblivious Transfer", 20^{th} STOC (1988), 20–29.

[18] M. Naor, R. Ostrovsky, R. Venkatesan and M. Yung, "Perfect Zero-Knowledge Arguments for NP Can Be Based on General Complexity Assumptions", *Advances in Cryptology – Proceedings of Crypto92*, Springer-Verlag Lecture Notes in Computer Science. To appear in *J. of Cryptology.*

[19] R. Ostrovsky, S. Rajagopalan, and U. Vazirani, "Simple and Efficient Leader Election in the Full Information Model", *STOC* 1994, 234–242.

[20] R. Ostrovsky, R. Venkatesan and M. Yung, "Fair Games Against an All-Powerful Adversary", Presented at DIMACS Complexity and Cryptography Workshop, October 1990, Princeton. *AMS DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, Vol. 13, Jin-Yi Cai editor, pp. 155–169, 1993.

[21] R. Ostrovsky, R. Venkatesan and M. Yung, "Interactive Hashing Simplifies Zero-Knowledge Protocol Design", *Eurocrypt93.*

[22] Santha, M., and U.V. Vazirani, "Generating Quasi-Random Sequences from Slightly-Random Sources", *25th Symp. on Foundation of Computer Science*, pp. 434–440, 1984.

[23] U.V. Vazirani, and V.V. Vazirani, "Random Polynomial Time is equal to Slightly-Random Polynomial Time", *26th Symp. on Foundation of Computer Science*, pp. 417–428, 1985.

[24] Yao, A. C., "Protocols for Secure Computations", *23th FOCS*, pp. 160–164, 1982.

[25] Yao, A. C., "How to Generate and Exchange Secrets", *27th FOCS*, 1986, pp. 162–167.