Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 148 (2021)

EXPLICIT EXPONENTIAL LOWER BOUNDS FOR EXACT HYPERPLANE COVERS

Benjamin E. Diamond
Applied Cryptography
Coinbase
benediamond@gmail.com

Amir Yehudayoff*
Department of Mathematics
Technion – IIT
amir.yehudayoff@gmail.com

Abstract

We describe an explicit and simple subset of the discrete hypercube which cannot be exactly covered by fewer than exponentially many hyperplanes. The proof exploits a connection to communication complexity, and relies heavily on Razborov's lower bound for disjointness.

1 Introduction

The relationship between hyperplanes in Euclidean space and the discrete hypercube is fundamental and important. One basic problem entails understanding the number of hyperplanes which one must use to cover various subsets of the cube. A *hyperplane* is specified by a perpendicular vector $(a_1, \ldots, a_n) \in \mathbb{R}^n$ and an offset $w \in \mathbb{R}$, defined specifically as the set $H = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n a_i \cdot x_i = w\}$. A collection of hyperplanes $H_1, H_2, \ldots, H_N$ *exactly covers* a set $S \subseteq \{0,1\}^n$ if $\bigcup_{i=1}^N H_i \cap \{0,1\}^n = S$. The *exact cover number* $\mathrm{ec}(S)$ of a set S is the minimum cardinality N attained across all hyperplane configurations exactly covering S.

The exact cover number $\mathrm{ec}(\{0,1\}^n)$ of the full cube is 2. Seminal work of Alon and Füredi [AF93] shows that removing a single point makes exact covering much harder; the exact cover number of $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ is n. The study of exact covers of arbitrary sets appears in recent work of Aaronson, Groenland, Grzesik, Johnston, and Kielak [AGG+21]. That work focuses on worst-case cardinalities of the form $\mathrm{ec}(n) := \max_{S\subseteq\{0,1\}^n} \mathrm{ec}(S)$; it proves that $\mathrm{ec}(n)$ is between $\frac{2^n}{n^2}$ and $2^{n-\lfloor \log n \rfloor} < 2 \cdot \frac{2^n}{n}$. The work's lower bound on $\mathrm{ec}(n)$ is not explicit, as it relies on a generic counting argument.

In this work, we analyze $\mathrm{ec}(S)$ for concrete subsets $S \subseteq \{0,1\}^n$. Beyond this problem's intrinsic appeal, an additional strong source of motivation stems from forthcoming work by the first-listed author [Dia22], which links exact hyperplane covers to secure two-party computation. That work shows that exact covers yield protocols for secure computation by two malicious parties. The work's protocols, moreover, are efficient when the relevant exact cover numbers are small. It is of interest, therefore, to determine which set families are and are not efficiently coverable.

Our main result exhibits a concrete set whose exact cover number is exponentially large.

Definition. We write $D_n \subseteq \{0,1\}^{n/2} \times \{0,1\}^{n/2} \cong \{0,1\}^n$ for the set

$$D_n = \left\{(x,y) \in \{0,1\}^{n/2} \times \{0,1\}^{n/2} \;\middle|\; \left(\bigvee_{i=1}^{n/2} x_i \wedge y_i\right) = 0\right\}.$$

In other words, D_n consists exactly of those pairs (x,y) for which x and y—interpreted as *sets*—are disjoint.

Theorem. $\mathrm{ec}(D_n) \geq 2^{\Omega(n)}$.

Though we focus throughout this work on hyperplanes defined over the real numbers, our results in fact carry through to *any* field, and even to "hyperplanes" defined over $\mathbb{Z}$ (i.e., to rank-$n-1$ submodules of $\mathbb{Z}^n$).

*Research funded by NSF–BSF grant number 2021674.

ISSN 1433-8092

We briefly discuss a further interpretation of the theorem. Exact cover numbers can be understood as furnishing a sort of complexity measure on boolean functions. In computational complexity, functions are computed by *devices*, where each device has a *cost*; the complexity of a *function* is the minimum cost of a device computing it. In our setting, the "devices" are hyperplane covers, and the "cost" of a hyperplane cover is its cardinality. Specifically, to each boolean function $f : \{0,1\}^n \to \{0,1\}$, we may associate, for example, the complexity measure $\mathrm{ec}(f) := \max\left\{\mathrm{ec}(f^{-1}(0)), \mathrm{ec}(f^{-1}(1))\right\}$.

Exact cover complexity relates in surprising ways to other, more standard, complexity measures, including, for example, that of *constant-depth circuit complexity*, a classically studied metric (see e.g. [Vol99, § 3.3]). We demonstrate this through the following example. A function f is *symmetric* if it is invariant under arbitrary permutations of its inputs (*parity* and *majority* are symmetric functions, for example). Symmetric functions in general admit only *superpolynomially sized* constant-depth circuits (see e.g. Vollmer [Vol99, Cor. 3.32]). On the other hand, if f is symmetric, then $\mathrm{ec}(f)$ is at most $n+1$. This latter fact can be seen in the following way. For each symmetric f, $f^{-1}(1)$ is a union of sets of the form $S_j := \{x \in \{0,1\}^n \mid \sum_{i=1}^n x_i = j\}$, for constants $j \in \{0, \ldots, n\}$. Each individual set S_j can be exactly covered by a single hyperplane, so that $f^{-1}(1)$ can be exactly covered by $n+1$ hyperplanes. A similar argument applies to $f^{-1}(0)$. We see that symmetric functions have *small* exact cover complexity, despite their *large* constant-depth circuit complexity. In the opposite direction, our theorem shows that the exact cover complexity of a polynomially-sized, depth-two, monotone circuit can be exponential in n. This discrepancy demonstrates a strong separation between exact cover complexity and constant-depth circuit complexity.

We prove our lower bound on the exact cover number of D_n by upper-bounding the sizes of certain sets of the form $H \cap D_n$, where $H \subseteq \mathbb{R}^n$ is a hyperplane. No useful such upper bound, of course, can possibly hold for *all* hyperplanes. Slightly abusing terminology, we say that a hyperplane $H \subseteq \mathbb{R}^n$ is *contained* in a set $S \subseteq \{0,1\}^n$ if $H \cap \{0,1\}^n \subseteq S$. If a hyperplane H is contained in S, then, trivially, $|H \cap \{0,1\}^n| \leq |S|$ holds. The following lemma establishes an exponential improvement in the *particular* case of D_n:

Lemma. *If a hyperplane $H \subseteq \mathbb{R}^n$ is contained in D_n, then $|H \cap \{0,1\}^n| \leq 2^{-\Omega(n)} \cdot |D_n|$.*

The lemma implies the theorem by means of a covering argument, which we presently sketch. Each hyperplane H which *participates* in an exact cover of D_n must be *contained* in D_n. The lemma entails that each particular such H may alone cover at most a proportion of $2^{-\Omega(n)}$ among D_n's points. It follows that at least $2^{\Omega(n)}$ hyperplanes must be used in any configuration exactly covering D_n.

The lemma may be interpreted as a strong—though restricted—anti-concentration result. Classical anti-concentration results concern expressions of the form $\max_{w \in \mathbb{R}} \Pr\left[\sum_{i=1}^n a_i \cdot X_i = w\right]$, where X is uniformly distributed in $\{0,1\}^n$. The *Littlewood–Offord problem* entails establishing anti-concentration when all of the coordinates of $(a_1, \ldots, a_n)$ are assumed to be nonzero [LO43]; the problem's original motivation arose from the study of roots of random polynomials. Littlewood and Offord proved the preliminary upper bound of $O\left(\frac{\log n}{\sqrt{n}}\right)$. In a celebrated and sharp result, Erdős [Erd45] solved the Littlewood–Offord problem, proving the upper bound $2^{-n} \cdot \binom{n}{\lfloor n/2 \rfloor} = \Theta\left(\frac{1}{\sqrt{n}}\right)$ using Sperner's theorem on the sizes of antichains. Kleitman [Kle65], Frankl and Füredi [FF88], Griggs [Gri93], and others subsequently generalized the problem.

The lemma says that for each normal $a \in \mathbb{R}^n$, we have $\max_w \Pr\left[\sum_{i=1}^n a_i \cdot X_i = w\right] \leq 2^{-\Omega(n)} \cdot \Pr[X \in D_n]$, where the maximum is taken *not* over all constants $w \in \mathbb{R}$, but rather over only those for which the hyperplane $\{x \in \mathbb{R}^n \mid \sum_{i=1}^n a_i \cdot x_i = w\}$ is contained in D_n. The lemma holds for all a, and guarantees *exponentially* strong anti-concentration; on the other hand, the bound holds only for certain w. The fact that the bound holds only for some among the values w makes it difficult to use known techniques to prove anti-concentration (like extremal combinatorics, or Fourier analysis).

Our main high-level contribution is a bridge between this sort of restricted anti-concentration and two-party communication complexity (see the textbook [RY20] and references within). Our proof follows the ideas of Razborov's [Raz92] famous lower bound for the distributional two-party communication complexity of disjointness. This bridge is built by the means of a certain decomposition. For each hyperplane $H = \left\{(x, y) \in \{0,1\}^{n/2} \times \{0,1\}^{n/2} \;\middle|\; \sum_{i=1}^{n/2} a_i \cdot x_i + \sum_{i=1}^{n/2} b_i \cdot y_i = w\right\}$, we have:

$$H \cap D_n = \bigcup_{k \in \mathbb{R}} (A_k \times B_k) \cap D_n, \tag{1}$$

where $A_k := \left\{x \in \{0,1\}^{n/2} \;\middle|\; \sum_{i=1}^{n/2} a_i \cdot x_i = k\right\}$ and $B_k := \left\{y \in \{0,1\}^{n/2} \;\middle|\; \sum_{i=1}^{n/2} b_i \cdot y_i = w - k\right\}$.

In the language of communication complexity, for each k, the set $A_k \times B_k$ is a *rectangle* (i.e., a product set in $\{0,1\}^{n/2} \times \{0,1\}^{n/2}$); the assumption $H \subseteq D_n$ entails precisely that each rectangle $A_k \times B_k$ is *monochromatic* (i.e., it satisfies $A_k \times B_k \subseteq D_n$). In a nutshell, we see that if $\mathrm{ec}(D_n)$ were small, then D_n would resemble the on-set of a function with low communication complexity. The full picture is in fact subtler, as we explain below.

The crux of Razborov's lower bound asserts that, in the setting of disjointness, all 1-monochromatic rectangles are small. To assess their sizes, Razborov carefully constructs a probability measure ρ on the cube, and proves that each 1-monochromatic rectangle has exponentially small mass—say, $2^{-\varepsilon \cdot n}$—under this measure. In our language, Razborov proves that if a rectangle $A \times B$ is contained in D_n, then $\rho(A \times B) \leq 2^{-\varepsilon \cdot n}$. We must surmount a few barriers in order to apply Razborov's ideas in our context.

The first among these barriers concerns the possible number r of rectangles involved (i.e., the number of distinct values k which appear in the union expression (1) above). In Razborov's setting, the bound of $2^{-\varepsilon \cdot n}$ on the mass of each *individual* rectangle—together with a union bound—implies that $r \geq 2^{\varepsilon \cdot n}$. This immediately implies the communication complexity lower bound. An exponential lower bound on r, however, is useless in our setting; we must not lower-bound *the total number of rectangles used*, but rather upper-bound *the total probability mass they represent.* The key observation which overcomes this barrier is that our rectangles have a very specific structure; indeed, the sets $\{A_k\}_{k \in \mathbb{R}}$ are themselves pairwise disjoint in $\{0,1\}^{n/2}$, and likewise for the sets $\{B_k\}_{k \in \mathbb{R}}$. This observation, together with a more careful analysis, allows us to overcome the first barrier. Interestingly, we exploit the structure of H as a hyperplane during our proof *only* in our use of this simple property of the set families $\{A_k\}_{k \in \mathbb{R}}$ and $\{B_k\}_{k \in \mathbb{R}}$.

A second barrier that we must overcome stems from the fact that the distribution on D_n we consider is uniform; Razborov's argument exploits the carefully constructed distribution ρ. This difference introduces several technical difficulties. Roughly speaking, we use *measure concentration* to reduce our problem to a setting closer to Razborov's; we then analyze a "perturbed" variant of his distribution (see Claim 2.1 below).

Our main lemma above conceals an implicit small linear constant within its expression $\Omega(n)$, which is moreover ineffective throughout our proof. We suspect that the following precise variant of our main result holds:

Conjecture. *If a hyperplane $H \subseteq \mathbb{R}^n$ is contained in D_n, then $|H \cap \{0,1\}^n| \leq 2^{n/2}$.*

That is, we suspect the precise variant of our above lemma whereby $|H \cap \{0,1\}^n| \leq 2^{(1-\log_2 3) \cdot n/2} \cdot |D_n|$. This conjecture is sharp, in that there exist hyperplanes H contained in D_n for which $|H \cap \{0,1\}^n| = 2^{n/2}$. For example, we may take as H the hyperplane $\left\{(x,y) \in \{0,1\}^{n/2} \times \{0,1\}^{n/2} \;\middle|\; \sum_{i=1}^{n/2} x_i = 0\right\}$. The intersection of this H with $\{0,1\}^{n/2} \times \{0,1\}^{n/2}$ is exactly the set $\{(0,\ldots,0)\} \times \{0,1\}^{n/2}$ consisting of pairs (x,y) for which x is empty. This set is obviously contained in D_n, and consists of exactly $2^{n/2}$ points.

2 Proving the Lemma

We fix even n and a hyperplane

$$H = \left\{(x,y) \in \{0,1\}^{n/2} \times \{0,1\}^{n/2} \;\middle|\; \sum_{i=1}^{n/2} a_i \cdot x_i + \sum_{i=1}^{n/2} b_i \cdot y_i = w\right\} \subseteq \mathbb{R}^n,$$

which we moreover assume is contained in $D := D_n$. We write μ for the uniform distribution on $D \subseteq \{0,1\}^n$. Throughout, we frequently interpret elements x and y of $\{0,1\}^{n/2}$ as *subsets* of $\{1,\ldots,\frac{n}{2}\}$, and use corresponding notation. For example, we write $|x|$ and $|y|$ for the *cardinalities*—that is, the Hamming weights—of x and y. We partition D along the sizes of its two constituent sets, in the following way. For integers ℓ_x and ℓ_y in $\{0,\ldots,\frac{n}{2}\}$, we set:

$$D_{\ell_x,\ell_y} := \left\{(x,y) \in \{0,1\}^{n/2} \times \{0,1\}^{n/2} \;\middle|\; |x| = \ell_x \wedge |y| = \ell_y\right\}.$$

We first argue that all but an exponentially vanishing proportion of the mass of μ is concentrated within those D_{ℓ_x,ℓ_y} for which both ℓ_x and ℓ_y are *simultaneously* near $\frac{n}{6}$.

Claim 2.1. *For each constant $\delta > 0$, we have $\mu\left(\bigcup_{n \cdot (\frac{1}{6}-\delta) \leq \ell_x, \ell_y \leq n \cdot (\frac{1}{6}+\delta)} D_{\ell_x,\ell_y}\right) > 1 - 2^{-\Omega(n)}$.*

Proof. A pair (X, Y) distributed according to μ may be sampled using $\frac{n}{2}$ *i.i.d.* pairs $(X_1, Y_1), \ldots, (X_{n/2}, Y_{n/2})$, each uniform in $\{(0,0), (0,1), (1,0)\}$. The expected values of $\sum_{i=1}^{n/2} X_i$ and of $\sum_{i=1}^{n/2} Y_i$ are exactly $\frac{n}{6}$. The union bound, together with a standard application of Chernoff's bound, completes the proof. $\square$

We write μ_{ℓ_x,ℓ_y} for the distribution μ conditioned on D_{ℓ_x,ℓ_y}. In light of Claim 2.1, it suffices to prove that $\mu_{\ell_x,\ell_y}(H) \leq 2^{-\Omega(n)}$ holds whenever ℓ_x and ℓ_y simultaneously reside in $\left[n \cdot (\frac{1}{6} - \delta), n \cdot (\frac{1}{6} + \delta)\right]$, for some appropriate fixed $\delta > 0$. Throughout the remainder of the proof, we fix $\delta := \frac{1}{300}$, as well as arbitrary integers ℓ_x and ℓ_y in $\left[\frac{49}{300} \cdot n, \frac{51}{300} \cdot n\right]$. To bound $\mu_{\ell_x,\ell_y}(H)$, we use the decomposition (1). For fixed $k \in \mathbb{R}$, we write:

$$A_k := \left\{x \in \{0,1\}^{n/2} \;\middle|\; \textstyle\sum_{i=1}^{n/2} a_i \cdot x_i = k\right\} \quad \text{and} \quad B_k = \left\{y \in \{0,1\}^{n/2} \;\middle|\; \textstyle\sum_{i=1}^{n/2} b_i \cdot y_i = w - k\right\}.$$

A key step in our proof involves sampling from μ_{ℓ_x,ℓ_y} in a more informative way.

Remark. We denote random variables by capital letters, and by lowercase letters the values they attain. In what follows, we impose the further assumption whereby $\frac{n}{2}$ is odd, so that $m := \frac{n-2}{4}$ is an integer. This restriction is not necessary, but simplifies notation below.

We write $T := (Z_x, Z_y, \{I\})$ for a uniformly random partition of $\{1, \ldots, \frac{n}{2}\}$ into subsets sized exactly m, m, and 1, respectively. We write X for a uniformly random subset of $Z_x \cup \{I\}$ of cardinality exactly ℓ_x and Y for a uniformly random subset of $Z_y \cup \{I\}$ of cardinality exactly ℓ_y (X and Y are chosen independently). We moreover write X_0 and Y_0 for X and Y conditioned on $I \notin X$ and $I \notin Y$, respectively. Finally, we write X_1 and Y_1 for X and Y conditioned on $I \in X$ and $I \in Y$, respectively.

We intuitively characterize the role which this sampling procedure plays in our proof. The core of our argument shows that for *each* arbitrary rectangle $A \times B \subseteq \{0,1\}^{n/2} \times \{0,1\}^{n/2}$ (including, crucially, the rectangles $A_k \times B_k$ constructed above), the probability that $(X_0, Y_0) \in A \times B$ cannot significantly exceed the probability that $(X_1, Y_1) \in A \times B$. This implies in particular that $H \cap D_n$ is small for any $H \subseteq \mathbb{R}^n$ contained in D_n, since (X_1, Y_1) never falls within D_n, whereas (X_0, Y_0) always does. Intuitively, a rectangle of mass at least $2^{-\varepsilon \cdot n}$ cannot dramatically affect the distribution of *most* of (X, Y)'s coordinates; meanwhile, the variables (X_0, Y_0) and (X_1, Y_1) differ at just a single random coordinate. We provide full details in Subsection 2.1 below.

2.1 Good partitions

In this subsection, we introduce and analyze the notion of "good" partitions, and describe their basic properties.

Definition 2.2. For each $k \in \mathbb{R}$, we define sets of "good" partitions in the following way:

$$G_x^k := \left\{t = (z_x, z_y, \{i\}) \;\middle|\; \Pr\left[X_1 \in A_k \mid T = t\right] \geq \tfrac{1}{70} \cdot \Pr\left[X_0 \in A_k \mid T = t\right] - 2^{-\varepsilon \cdot n}\right\}$$

and

$$G_y^k := \left\{t = (z_x, z_y, \{i\}) \;\middle|\; \Pr\left[Y_1 \in B_k \mid T = t\right] \geq \tfrac{1}{70} \cdot \Pr\left[Y_0 \in B_k \mid T = t\right] - 2^{-\varepsilon \cdot n}\right\},$$

where $\varepsilon := \frac{1}{500}$.

Roughly speaking, a partition $t = (z_x, z_y, \{i\})$ is "good" with respect to $k \in \mathbb{R}$ if, conditioned upon it, the distributions (X_0, Y_0) and (X_1, Y_1) do not intersect the sets A_k and B_k excessively differently. The main result of this subsection is the following proposition, which states that *most* partitions are "good". We write $\chi_x^k(t)$ and $\chi_y^k(t)$ for the indicator functions of the events $t \in G_x^k$ and $t \in G_y^k$, respectively.

Proposition 2.3. *For each $k \in \mathbb{R}$,*

$$\mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \cdot \Pr\left[Y_0 \in B_k \mid T\right] \cdot \chi_x^k(T) \cdot \chi_y^k(T)\right] \geq \tfrac{1}{10} \cdot \mathbb{E}_T\left[\Pr\left[X_0 \in A_k \mid T\right] \cdot \Pr\left[Y_0 \in B_k \mid T\right]\right].$$

Before proving the proposition, we establish a few preliminary claims. The first records a fact pertaining to the structure of this probability distribution upon conditioning.

Claim 2.4. *For each $k \in \mathbb{R}$, as the partition $t = (z_x, z_y, \{i\})$ varies, the numbers* $\Pr[X \in A_k \mid T = t]$ *and* $\Pr[Y_0 \in B_k \mid T = t]$ *depend only on z_y and the numbers* $\Pr[Y \in B_k \mid T = t]$ *and* $\Pr[X_0 \in A_k \mid T = t]$ *depend only on z_x.*

The following claim compares the probability that a random ℓ_x-element subset of $z_x \subseteq \{1, \dots \frac{n}{2}\}$ resides in $A_k \subseteq \{0,1\}^{n/2}$ with the probability that a random ℓ_x-element subset of $z_x \cup \{i\}$ does. It shows that the former cannot exceed the latter by much more than threefold. It exploits the fact that the probability that a random ℓ_x-element subset of $z_x \cup \{i\}$ does not contain i is close to $\frac{1}{3}$.

Claim 2.5. *For each fixed scalar $k \in \mathbb{R}$ and partition $t = (z_x, z_y, \{i\})$, we have*

$$\Pr[X_0 \in A_k \mid T = t] \le \tfrac{25}{8} \cdot \Pr[X \in A_k \mid T = t]$$

and

$$\Pr[Y_0 \in B_k \mid T = t] \le \tfrac{25}{8} \cdot \Pr[Y \in B_k \mid T = t].$$

Proof. We prove the first conclusion; the second is similar. By the definitions of the distributions X and X_0,

$$\begin{aligned}\Pr[X \in A_k \mid T = t] &= \Pr[i \notin X \mid T = t] \cdot \Pr[X_0 \in A_k \mid T = t] + \Pr[i \in X \mid T = t] \cdot \Pr[X_1 \in A_k \mid T = t] \\ &\ge \Pr[i \notin X \mid T = t] \cdot \Pr[X_0 \in A_k \mid T = t].\end{aligned}$$

The proportion of ℓ_x-element subsets of $z_x \cup \{i\}$ which do not contain i is $\Pr[i \notin X \mid T = t] = \frac{\binom{m}{\ell_x}}{\binom{m+1}{\ell_x}} = \frac{m+1-\ell_x}{m+1} \ge 1 - \frac{n}{n+2} \cdot \frac{17}{25} \ge \frac{8}{25}$. □

The following claim shows that, for each *fixed* $z_y \subseteq \{1, \dots, \frac{n}{2}\}$, as the index $i \in \{1, \dots, \frac{n}{2}\} - z_y$ varies, most among the resulting partitions $(z_x, z_y, \{i\})$ are "good"; a symmetrical statement holds for each *fixed* $z_x \subseteq \{1, \dots, \frac{n}{2}\}$ as the index $i \in \{1, \dots, \frac{n}{2}\} - z_x$ varies.

Claim 2.6. *For each fixed scalar $k \in \mathbb{R}$, and arbitrary fixed subsets z_x and z_y of $\{1, \dots, \frac{n}{2}\}$, we have:*

$$\Pr\left[T \notin G_x^k \mid Z_y = z_y\right] < \tfrac{1}{7}$$

and

$$\Pr\left[T \notin G_y^k \mid Z_x = z_x\right] < \tfrac{1}{7}.$$

Proof. We prove only the first inequality, as the second is similar. We first handle the case in which $\Pr[X \in A_k \mid Z_y = z_y] < 2^{-\varepsilon \cdot n}$. In light of Claim 2.4, we note that $\Pr[X \in A_k \mid Z_y = z_y] = \Pr[X \in A_k \mid T = t]$ holds for each *particular* partition t drawn from the distribution $(Z_x, z_y, \{I\})$. Using Claim 2.5, we see that if any particular such t moreover satisfied $t \notin G_x^k$, then we would have:

$$\Pr[X_1 \in A_k \mid T = t] < \tfrac{1}{70} \cdot \Pr[X_0 \in A_k \mid T = t] - 2^{-\varepsilon \cdot n} \le \tfrac{1}{20} \cdot \Pr[X \in A_k \mid T = t] - 2^{-\varepsilon \cdot n} < 0,$$

a contradiction, so that $\Pr\left[T \notin G_x^k \mid Z_y = z_y\right] = 0$, and the claim is proved.

We thus assume that $\Pr[X \in A_k \mid Z_y = z_y] \ge 2^{-\varepsilon \cdot n}$. We write $\overline{z_y}$ for the complement of z_y in $\{1, \dots, \frac{n}{2}\}$, and abbreviate $\widehat{A}_k := A_k \cap \binom{\overline{z_y}}{\ell_x}$ for the set of ℓ_x-element subsets of $\overline{z_y}$ which reside in A_k. We note that:

$$\Pr[X \in A_k \mid Z_y = z_y] = \frac{\left|\widehat{A}_k\right|}{\binom{m+1}{\ell_x}},$$

so that $\left|\widehat{A}_k\right| = \binom{m+1}{\ell_x} \cdot \Pr[X \in A_k \mid Z_y = z_y] \ge \binom{m+1}{\ell_x} \cdot 2^{-\varepsilon \cdot n}$. We moreover record the lower bound

$$\log_2 \binom{m+1}{\ell_x} \ge \log_2 \binom{m+1}{\lfloor \frac{51}{300} \cdot n \rfloor} \ge \log_2 \binom{m+1}{\lfloor \frac{17}{25} \cdot (m+1) \rfloor} \ge (0.9 - o(1)) \cdot (m+1);$$

the last inequality is a standard consequence of Stirling's approximation, together with the binary entropy inequality $H\left(\frac{17}{25}\right) > 0.9$. These facts together imply that $\log_2\left(\left|\widehat{A}_k\right|\right) \ge (0.9 - o(1) - 4 \cdot \varepsilon) \cdot (m+1)$.

We write $\widehat{X}$ for a uniformly random element of $\widehat{A}_k$. We fix a partition $t = (z_x, z_y, \{i\})$, and abbreviate $\widehat{A}_{k,1}$ for the set of elements of $\widehat{A}_k$ which contain i; we now have that:

$$\Pr[X \in A_k \mid T = t] \cdot \Pr\left[i \in \widehat{X}\right] = \frac{\left|\widehat{A}_k\right|}{\binom{m+1}{\ell_x}} \cdot \frac{\left|\widehat{A}_{k,1}\right|}{\left|\widehat{A}_k\right|} = \frac{\left|\widehat{A}_{k,1}\right|}{\binom{m}{\ell_x-1}} \cdot \frac{\binom{m}{\ell_x-1}}{\binom{m+1}{\ell_x}} = \Pr[X_1 \in A_k \mid T = t] \cdot \frac{\ell_x}{m+1},$$

so that $\Pr[X \in A_k \mid T = t] \cdot \Pr\left[i \in \widehat{X}\right] \leq \Pr[X_1 \in A_k \mid T = t] \cdot \frac{17}{25}$. By this inequality and Claim 2.5, we see that if *moreover* $t \notin G_x^k$ holds, then we have:

$$\Pr[X \in A_k \mid T = t] \cdot \Pr[i \in \widehat{X}] < \tfrac{1}{100} \cdot \Pr[X_0 \in A_k \mid T = t] \leq \tfrac{1}{30} \cdot \Pr[X \in A_k \mid T = t],$$

so that $\Pr\left[i \in \widehat{X}\right] < \frac{1}{30}$. Equivalently, if the partition $t = (z_x, z_y, \{i\})$ is not "good", then the component of the joint distribution $\widehat{X}$ corresponding to the element $i \in \overline{z_y}$ has success probability less than $\frac{1}{30}$.

We write $\widehat{X}_j$ for the indicator function of the event $i_j \in \widehat{X}$, where $\{i_1, \ldots, i_{m+1}\}$ are the elements of $\overline{z_y}$, so that $\widehat{X} = (\widehat{X}_1, \ldots, \widehat{X}_{m+1})$. We observe that if the claim were false—and, in particular, $\Pr[i \in \widehat{X}] < \frac{1}{30}$ held for a proportion consisting of at least $\frac{1}{7}$ among the $m+1$ elements $i \in \overline{z_y}$—then the binary entropy of $\widehat{X}$ would satisfy:

$$\begin{aligned}
(0.9 - o(1) - 4 \cdot \varepsilon) \cdot (m+1) &\leq H(\widehat{X}) && \text{(by } H(\widehat{X}) = \log_2\left(\left|\widehat{A}_k\right|\right) \text{ and the above)} \\
&\leq \sum_{j=1}^{m+1} H(\widehat{X}_j) && \text{(by the sub-additivity of entropy)} \\
&< \left(\tfrac{6}{7} + \tfrac{1}{7} \cdot H(\tfrac{1}{30})\right) \cdot (m+1) && \text{(by the assumption that the claim is false)} \\
&\leq 0.89 \cdot (m+1).
\end{aligned}$$

This contradiction completes the proof of the claim. ☐

We are now ready to prove the main proposition.

Proof of Proposition 2.3. Because $1 - \chi_x^k(t) \cdot \chi_y^k(t) \leq 1 - \chi_x^k(t) + 1 - \chi_y^k(t)$ holds for each t, and by linearity of expectation and symmetry, it suffices to prove that

$$\mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \cdot \Pr[Y_0 \in B_k \mid T] \cdot \left(1 - \chi_x^k(T)\right)\right] \leq \tfrac{9}{20} \cdot \mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \cdot \Pr[Y_0 \in B_k \mid T]\right].$$

To prove this, it in turn suffices to show that, for each *fixed* m-element subset $z_y \subseteq \{1, \ldots, \frac{n}{2}\}$, it holds that:

$$\begin{aligned}
&\mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \cdot \Pr[Y_0 \in B_k \mid T] \cdot \left(1 - \chi_x^k(T)\right) \mid Z_y = z_y\right] \\
&\quad \leq \tfrac{9}{20} \cdot \mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \cdot \Pr[Y_0 \in B_k \mid T] \mid Z_y = z_y\right].
\end{aligned}$$

We prove this latter claim in the following way:

$$\begin{aligned}
&\mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \cdot \Pr[Y_0 \in B_k \mid T] \cdot \left(1 - \chi_x^k(T)\right) \mid Z_y = z_y\right] \\
&\quad = \Pr[Y_0 \in B_k \mid Z_y = z_y] \cdot \mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \cdot \left(1 - \chi_x^k(T)\right) \mid Z_y = z_y\right] && \text{(by Claim 2.4)} \\
&\quad \leq \tfrac{25}{8} \cdot \Pr[Y_0 \in B_k \mid Z_y = z_y] \cdot \mathbb{E}_T\left[\Pr[X \in A_k \mid T] \cdot \left(1 - \chi_x^k(T)\right) \mid Z_y = z_y\right] && \text{(by Claim 2.5)} \\
&\quad = \tfrac{25}{8} \cdot \Pr[Y_0 \in B_k \mid Z_y = z_y] \cdot \Pr[X \in A_k \mid Z_y = z_y] \cdot \mathbb{E}_T\left[1 - \chi_x^k(T) \mid Z_y = z_y\right] && \text{(by Claim 2.4)} \\
&\quad \leq \tfrac{25}{8} \cdot \tfrac{1}{7} \cdot \Pr[Y_0 \in B_k \mid Z_y = z_y] \cdot \Pr[X \in A_k \mid Z_y = z_y] && \text{(by Claim 2.6)} \\
&\quad = \tfrac{25}{56} \cdot \Pr[Y_0 \in B_k \mid Z_y = z_y] \cdot \mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \mid Z_y = z_y\right].
\end{aligned}$$

The final equality above amounts to the following calculation:

$$\Pr[X \in A_k \mid Z_y = z_y] = \frac{\left|A_k \cap \binom{\overline{z_y}}{\ell_x}\right|}{\binom{m+1}{\ell_x}} = \frac{1}{m+1} \cdot \frac{\left|A_k \cap \binom{\overline{z_y}}{\ell_x}\right| \cdot (m+1-\ell_x)}{\binom{m}{\ell_x}} = \frac{1}{m+1} \cdot \sum_{i \in \overline{z_y}} \frac{\left|A_k \cap \binom{\overline{z_y} - \{i\}}{\ell_x}\right|}{\binom{m}{\ell_x}},$$

where the rightmost expression is precisely $\mathbb{E}_T\left[\Pr[X_0 \in A_k \mid T] \mid Z_y = z_y\right]$, by definition, and the final equality stems from a double-counting argument; indeed, $\sum_{i \in \overline{z_y}} \left|A_k \cap \binom{\overline{z_y} - \{i\}}{\ell_x}\right|$ counts each distinct element of $A_k \cap \binom{\overline{z_y}}{\ell_x}$ exactly $m+1-\ell_x$ times. An additional application of Claim 2.4 completes the proof. ☐

2.2 Completing the argument

We record the following claim:

Claim 2.7. $\sum_{k\in\mathbb{R}} \mathbb{E}_T \left[\Pr\left[X_0 \in A_k \mid T\right]\right] \leq 1$ *and* $\sum_{k\in\mathbb{R}} \mathbb{E}_T \left[\Pr\left[Y_0 \in B_k \mid T\right]\right] \leq 1$.

Proof. We prove the first inequality. Because the sets A_k—as k ranges throughout $\mathbb{R}$—are pairwise disjoint,

$$\sum_{k\in\mathbb{R}} \mathbb{E}_T \left[\Pr\left[X_0 \in A_k \mid T\right]\right] = \sum_{k\in\mathbb{R}} \Pr\left[X_0 \in A_k\right] \leq \Pr\left[X_0 \in \bigcup_{k\in\mathbb{R}} A_k\right] \leq 1.$$

This completes the proof. $\square$

We are now in a position to prove the lemma. Invoking finally the hypothesis whereby H is contained in D_n, we have that:

$$0 = \Pr[(X_1, Y_1) \in D_n] \geq \Pr[(X_1, Y_1) \in H].$$

Applying now the decomposition (1), we have:

$$0 = \Pr[(X_1, Y_1) \in H] \geq \sum_{k\in\mathbb{R}} \mathbb{E}_T \left[\Pr\left[X_1 \in A_k \mid T\right] \cdot \Pr\left[Y_1 \in B_k \mid T\right] \cdot \chi_x^k(T) \cdot \chi_y^k(T)\right].$$

Invoking Definition 2.2, we see further that:

$$0 \geq \sum_{k\in\mathbb{R}} \mathbb{E}_T \left[\left(\tfrac{1}{70} \cdot \Pr\left[X_0 \in A_k \mid T\right] - 2^{-\varepsilon\cdot n}\right) \cdot \left(\tfrac{1}{70} \cdot \Pr\left[Y_0 \in B_k \mid T\right] - 2^{-\varepsilon\cdot n}\right) \cdot \chi_x^k(T) \cdot \chi_y^k(T)\right].$$

Applying Claim 2.7, we have that:

$$0 \geq \tfrac{1}{70^2} \cdot \sum_{k\in\mathbb{R}} \mathbb{E}_T \left[\Pr\left[X_0 \in A_k \mid T\right] \cdot \Pr\left[Y_0 \in B_k \mid T\right] \cdot \chi_x^k(T) \cdot \chi_y^k(T)\right] - 2^{-\Omega(n)}.$$

Using Proposition 2.3, we conclude that:

$$\begin{aligned}
0 &\geq \tfrac{1}{10\cdot 70^2} \cdot \sum_{k\in\mathbb{R}} \mathbb{E}_T \left[\Pr\left[X_0 \in A_k \mid T\right] \cdot \Pr\left[Y_0 \in B_k \mid T\right]\right] - 2^{-\Omega(n)} \\
&= \tfrac{1}{10\cdot 70^2} \cdot \mathbb{E}_T \left[\sum_{k\in\mathbb{R}} \Pr\left[X_0 \in A_k \mid T\right] \cdot \Pr\left[Y_0 \in B_k \mid T\right]\right] - 2^{-\Omega(n)} \\
&= \tfrac{1}{10\cdot 70^2} \cdot \mu_{\ell_x,\ell_y}(H) - 2^{-\Omega(n)}.
\end{aligned}$$

In light of Claim 2.1, this calculation completes the proof of the lemma.

References

[AF93] Noga Alon and Zoltán Füredi. Covering the cube by affine hyperplanes. *European Journal of Combinatorics*, 14(2):79–83, 1993.

[AGG+21] James Aaronson, Carla Groenland, Andrzej Grzesik, Tom Johnston, and Bartłomiej Kielak. Exact hyperplane covers for subsets of the hypercube. *Discrete Mathematics*, 344(9), 2021.

[Dia22] Benjamin E. Diamond. Securely computing piecewise constant codes. `https://ia.cr/2021/146`, August 2022. manuscript in preparation.

[Erd45] Paul Erdős. On a lemma of Littlewood and Offord. *Bulletin of the American Mathematical Society*, 51:898–902, 1945.

[FF88] Peter Frankl and Zoltán Füredi. Solution of the Littlewood–Offord problem in high dimensions. *Annals of Mathematics*, 128(2):259–270, 1988.

[Gri93] Jerrold R. Griggs. On the distribution of the sums of residues. *Bulletin of the American Mathematical Society*, 28(2):329–333, 1993.

[Kle65] Daniel J. Kleitman. On a lemma of Littlewood and Offord on the distribution of certain sums. *Mathematische Zeitschrift*, 90(4):251–259, 1965.

[LO43] John E. Littlewood and Cyril Offord. On the number of real roots of a random algebraic equation (III). *Sbornik: Mathematics*, 12(3):277–286, 1943.

[Raz92] Alexander A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[RY20] Anup Rao and Amir Yehudayoff. *Communication Complexity and Applications*. Cambridge University Press, 2020.

[Vol99] Heribert Vollmer. *Introduction To Circuit Complexity: A Uniform Approach*. Texts in Theoretical Computer Science. Springer-Verlag, 1999.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il