Electronic Colloquium on Computational Complexity, Report No. 48 (2020)

Improved lifting theorems via robust sunflowers

Shachar Lovett*
Computer Science Department
UC San Diego
slovett@cs.ucsd.edu

Raghu Meka
Computer Science Department
UC Los Angeles
email@email.edu

Jiapeng Zhang†
Computer Science Department
Harvard University
jpeng.zhang@gmail.com

April 16, 2020

Abstract

Lifting theorems are a generic way to lift lower bounds in query complexity to lower bounds in communication complexity, with applications in diverse areas, such as combinatorial optimization, proof complexity, game theory. Lifting theorems rely on a gadget, where smaller gadgets give stronger lower bounds. However, existing proof techniques are known to require somewhat large gadgets.

We focus on one of the most widely used gadgets, the index gadget. For this gadget, existing lifting techniques are known to require at least a quadratic gadget size. We develop a new approach to prove lifting theorems for the indexing gadget, based on a novel connection to the recently developed robust sunflower lemmas. We show that this allows to reduce the gadget size to linear. We conjecture that it can be further improved to poly-logarithmic, similar to the known bounds for the corresponding robust sunflower lemmas.

1 Introduction

A lifting theorem is a *meta-theorem* which characterizes the communication complexity of a task in terms of an easier property such as *query complexity*. Within communication complexity, most functions of interest—e.g., equality, set-disjointness, inner-product, gap-hamming (c.f. [18, 20])—are *lifted functions* of the form $F \equiv f \circ g^n$ where $f\colon \{0,1\}^n \to \{0,1\}$ and $g\colon X \times Y \to \{0,1\}$ is a *small* two-party function, often called a *gadget*. Here, Alice and Bob are given inputs $x \in X^n$, $y \in Y^n$ respectively; their goal is to compute $F(x,y) = f \circ g^n(x,y)$.

A longstanding goal has been to characterize the communication complexity of lifted functions F as above in terms of a suitable *complexity-measure* of the "outer" function f. We want gadgets g such that for each boolean function $f : \{0,1\}^n \to \{0,1\}$ that has large decision tree complexity, the two-party function $F = f \circ g^n$ has large communication complexity.

As analyzing complexity of functions is often easier than analyzing communication problems, lifting theorems give generic ways to prove lower bounds in various models. Further, lifted functions (and closely related variants) are one of the main techniques we have to separate complexity classes and construct hard functions.

*Research supported by NSF Award CCF-1909634.
†Research Supported by NSF grant CCF-1763299 and Salil Vadhan's Simons Investigator Award.

ISSN 1433-8092

The framework of lifting theorems has been extensively studied in last two decades [4, 7, 11–13, 16, 25, 27, 28, 30] and several important advances in communication complexity in the past few years have used such an approach. These include the resolution of the *clique vs independent set* [10] problem of Yannakakis [31], exponential lower bounds for linear secret-sharing schemes [26], sub-exponential lower bounds on the size required for LPs to solve CSPs [19], refutation of the log approximate-rank conjecture [6], lower bound on the communication complexity of Nash equilibrium [2, 17]. In addition, study of lifting theorems has catalyzed a new wave of progress in query complexity [16].

The most commonly used gadget is the *indexing gadget* (e.g., [3, 15, 21, 25, 27]) and this is what we will work with. Another commonly used gadget is the *inner-product gadget* ([13, 19]; see also [5] for related but more general gadgets). The indexing gadget (formally defined below) is a *universal gadget* in the sense that if a lifting theorem is true for some gadget, then it also implies a similar result for indexing with different parameters. Further, the indexing gadget is more useful in applications (e.g., [11], [3], [21], [19]) as it preserves the underlying combinatorial structure[1].

The main point of our paper is to reduce the 'size' of a gadget $g : X \times Y \to \{0,1\}$ defined as $\min(|X|, |Y|)$. Gadget size is a fundamental parameter in lifting theorems and their applications. This is because often in applications, one loses factors that depend polynomially on the gadget size (as defined above). An ideal lifting theorem – one with constant gadget size – would give a unified way to prove tight lower bounds for most functions. Taking [19] as an example, improving their gadget size from $\text{poly}(n)$ to $O(1)$ (or even $\text{poly}(\log n)$) would improve their lower bounds for extended formulations from $2^{n^{\Omega(1)}}$ to $2^{\Omega(n)}$ (or $2^{\Omega(n/\text{poly}(\log n))}$ respectively). This would be quite remarkable.

In spite of the tremendous progress in lifting theorems, most generic lifting theorems require gadget sizes that are polynomial in n.[2] For example, we now have lifting theorems for the three classical models of communication: deterministic [16, 25], non-deterministic [13] and randomized [15]. In each of these results (and related ones), the gadget size has to be at least $\Omega(n^2)$.

Can we circumvent this quadratic barrier? Our main result is such a lifting theorem for almost linear-size indexing gadget for deterministic communication complexity. Moreover, our approach does not seem to have the same bottleneck as previous approaches, and we conjecture that it can be pushed to poly-logarithmic gadget sizes. We discuss this more in Section 1.3.

1.1 Our contribution

The indexing gadget $\text{Ind}_q : [q] \times \{0,1\}^q$ is defined as $\text{Ind}_q(x, y) = y_x$. When q is clear in the context, we write it succinctly as Ind. We study lifting for $\mathbf{P}$, namely, lifting deterministic decision tree complexity to deterministic communication complexity. Lifting for $\mathbf{P}$ was first studied by Raz and McKenzie [25], where they used this lifting theorem to prove monotone circuit lower bounds. More recently, Göös, Pitassi and Watson [16] refined their approach, and gave further applications in communication complexity. Formally, they proved the following theorem.

Theorem 1.1 ([16, 25])**.** *Let $n \geq 1, q \geq n^c$ where $c > 0$ is a large constant. For any boolean function $f : \{0,1\}^n \to \{0,1\}$ it holds that*

$$\mathbf{P}^{\text{cc}}(f \circ Ind_q^n) = \Theta\left(\mathbf{P}^{\text{dt}}(f) \cdot \log n\right).$$

Here $\mathbf{P}^{\text{cc}}(f \circ \text{Ind}_q^n)$ means the deterministic communication complexity of $f \circ \text{Ind}_q^n$, and $\mathbf{P}^{\text{dt}}(f)$ means the deterministic decision tree complexity of f.

As mentioned above, we want to minimize the gadget size q in Theorem 1.1, which currently needs to be polynomially large in n. It is an open problem whether a similar theorem holds for smaller q, ideally of constant size, but even the sub-polynomial case is widely open.

[1] For instance, [19] proves a lifting theorem for the inner-product gadget, but then translate it to the indexing gadget for the application.

[2] Some notable exceptions for specific models of communication with better gadget size are [14, 23, 28, 29].

There are two main approaches used in previous query-to-communication lifting theorems. The first approach is the Raz-McKenzie's simulation approach [16, 25, 30]. The second one is the related *extractor-based* approach introduced in [13] (also see [4, 12]). Either of these two approaches can be used to prove Theorem 1.1, however, both of them can be shown to fail if q is not large enough; concretely, if $q = o(n^2)$. We discuss this in more detail in Section 1.3. Therefore, a new approach seems necessary to reduce gadget size significantly.

Motivated by this goal, we initiate a different approach to prove lifting theorems, which builds on a new connection between lifting theorems and the robust sunflower lemma [1, 24]. The following is our main theorem, which shows promise of this approach, by allowing us to decrease the gadget size to near-linear.

Theorem 1.2. *Let $n \geq 1, q = \Omega(n \cdot \log^2 n)$. For any $f : \{0,1\}^n \to \{0,1\}$ it holds*

$$\mathbf{P}^{\mathrm{cc}}(f \circ Ind_q^n) = \Omega\left(\mathbf{P}^{\mathrm{dt}} \cdot \log(q/n)\right).$$

If we assume that $q \geq n^{1+\varepsilon}$ for any $\varepsilon > 0$ then

$$\mathbf{P}^{\mathrm{cc}}(f \circ Ind_q^n) = \Theta\left(\mathbf{P}^{\mathrm{dt}} \cdot \log q\right).$$

We note that, similar to Theorem 1.1, Theorem 1.2 holds in more general settings: f can have a non-boolean output, be a partial function or a relation, in which case the protocol for $f \circ \mathrm{Ind}$ should satisfy the analogous guarantees.

Possibly more interesting than the concrete result, the new approach has potential to go below previous bottlenecks. In fact, we conjecture that it can be pushed to gadget size $q = O(\log n)$, which is the limit of the robust sunflower lemmas. We discuss this and some promising initial results in this direction in Section 1.3.

1.2 Proof overview

Our proof follows the main framework of the Raz-McKenzie simulation [16, 25]. The main difference is that we use robust sunflowers to maintain a disperser property. Given any communication protocol that computes $f \circ \mathrm{Ind}_q^n$, we simulate this protocol and convert it to a decision tree that computes f. We use $X \in [q]^n$ and $Y \in (\{0,1\}^q)^n$ to denote the inputs of Alice and Bob, respectively, and $z \in \{0,1\}^n$ to denote the input to f.

During each simulation step, we maintain two sets $\mathbf{X} \subseteq [q]^n$ and $\mathbf{Y} \subseteq (\{0,1\}^q)^n$ of inputs that are consistent with the current simulation step. We call $\mathbf{X}$ k-spread[3] if for any $I \subset [n]$ and $S \in [q]^I$, it satisfies

$$|\{X \in \mathbf{X} : X_I = S\}| \leq k^{-|I|} \cdot |\mathbf{X}|.$$

Spreadness is a crucial notion in the following lemma, which is a direct corollary of the recent robust sunflower lemmas [1, 9, 24].

Lemma 1.3 (Robust sunflowers, see Theorem 2.11)**.** *Let $\gamma > 0$ and assume that $\mathbf{X}$ is $O(\log(n/\gamma))$-spread. Then for each $z \in \{0,1\}^n$,*

$$\Pr_{Y \sim (\{0,1\}^q)^n}[\exists X \in \mathbf{X}, Ind(X,Y) = z] > 1 - \gamma.$$

Notice that this lemma is non-trivial already for $k = \Omega(\log n)$. However in this paper, we choose $k = \Omega(n)$, because in this regime we have the following strong corollary.

Corollary 1.4 (Disperser property, see Lemma 3.1)**.** *Assume that $\mathbf{X}$ is $\Omega(n)$-spread and $|\mathbf{Y}| \geq 2^{-n} \cdot 2^{qn}$. Then*

$$\{Ind(X,Y) : X \in \mathbf{X}, Y \in \mathbf{Y}\} = \{0,1\}^n.$$

In order to ensure the correctness of our simulation process, we preserve the following properties of $\mathbf{X}$ and $\mathbf{Y}$ in each simulation step,

[3] This is closely related to *blockwise-denseness* from [13].

- $\mathbf{X}$ is k-spread, for a k that is slightly larger than n;
- $|\mathbf{Y}| \geq 2^{-n} \cdot 2^{qn}$.

At a high level, as long as these properties are preserved, the current set of inputs $\mathbf{X}, \mathbf{Y}$ project to all possible inputs z to the boolean function f. It is simple to maintain the property that $\mathbf{Y}$ is big; however, it could be the case that $\mathbf{X}$ loses the k-spread property. In this case, we show that we can query a set of input bits z_i and maintain the k-spread property. This increases the density of $\mathbf{X}$ in its ambient space, which allows us to bound the total number of bits queried as a function of the communication complexity of the lifted function. Implementing this strategy turns out to be more intricate and more details are provided in the actual proofs (which are not too technical).

1.3 Discussions and future directions

Comparison to Raz-McKenzie simulation. The idea of the Raz-McKenzie simulation [16, 25] is to convert a communication protocol for $f \circ \mathrm{Ind}$ into a decision tree for f. In each simulation step, let $X = (X_1, \ldots, X_n) \subset [q]^n$ be a set of "consistent" inputs for Alice. Raz-McKenzie checks the min-entropy $\mathcal{H}(X_i \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ for each i. There are two possible cases,

1. For each i, $\mathcal{H}(X_i \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ is very close to maximal, namely, $\log q$.
2. There is an i, such that $\mathcal{H}(X_i \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ is much smaller than $\log q$.

In the first case, they are able to simulate the communication without making a query in the decision tree protocol. On the other hand, if there exists a coordinate X_i such that $\mathcal{H}(X_i \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ is small, then we can query the i-th coordinate in the decision tree model. In the Raz-McKenzie simulation, there is a crucial *average-to-worst* reduction (known as the "thickness lemma"), which requires $q = \Omega(n^2)$ to maintain the correctness of the simulation.

In comparison to our approach, we use a *higher moment argument* in the following sense. Raz-McKenzie's simulation checks $\mathcal{H}(X_i \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ for each single $i \in [n]$. In our approach, we check $\mathcal{H}_\infty(X_I)$ each set $I \subseteq [n]$. If there is a set of coordinates I that has small min-entropy, then we query all of these coordinates. One of the main advantage of using this high moment argument is to avoid the average-to-worst reduction. Therefore we are able to avoid the main bottleneck of Raz-McKenzie. A side bonus is that our simulation is also simpler than Raz-McKenzie simulation.

In Erdős and Rado [8] proof of the sunflower lemma, it also checks only the links of size 1. In this sense, Raz-McKenzie's simulation is an analogue of Erdős and Rado proof of the sunflower lemma. By contrast, our simulation is an analogue of Alweiss-Lovett-Wu-Zhang [1] new proof with improved bounds. We believe our method can further improve the gadget size to $q = (\log n)^{O(1)}$, similar to the improvement for the sunflower lemma of [1]. The general approach is to replace the assumption that $\mathbf{Y}$ is large with some pseudo-randomness property. This could remove the requirement of taking $k = \Omega(n)$, and would allow us to get closer to the $k = O(\log n)$ limit of the robust sunflowers lemma. We leave exploring this direction to future work.

Extractors vs Dispersers. The second approach of lifting theorems is based on *extractors* [4, 12, 13]. The main structural lemma in this approach is the following. Let $q \geq n^c$ for a large constant c, let $\mathbf{X}$ be a distribution on $[q]^n$ and $\mathbf{Y}$ be distribution on $(\{0,1\}^q)^n$. If both $\mathbf{X}$ and $\mathbf{Y}$ have large spreadness (think $\mathbf{X}$ being $(k \sim q^{0.9})$-spread), then the distribution $\mathrm{Ind}(\mathbf{X}, \mathbf{Y})$ is *multiplicatively* close to the uniform distribution. In particular, $\mathrm{Supp}(\mathrm{Ind}(\mathbf{X}, \mathbf{Y})) = \{0,1\}^n$.

The uniformity of $\mathrm{Ind}(\mathbf{X}, \mathbf{Y})$ implies several nice properties allowing us to simulate the communication protocol with few queries. Interestingly, Li, Lovett and Zhang [22] also used this lemma to reprove the Erdős-Rado sunflower lemma. The journal version of that paper contains examples giving barriers to this approach, showing that $q = \Omega(n)$ is necessary.

In this paper, we first replace extractors with *dispersers*. That is, we seek conditions on $\mathbf{X}, \mathbf{Y}$ such that $\mathrm{Ind}(\mathbf{X}, \mathbf{Y})$ has full support, but without requiring the distribution to be close to uniform. This is in fact very

similar to the approach used in [1] towards the sunflower conjecture. However, this introduces several other challenges in using earlier simulation arguments. For readers familiar with the literature, at a high-level, "extractor-like" properties make conditioning on partial assignments easier; whereas, this is problematic with disperser like conditions. However, we can overcome this hurdles to lift $\mathbf{P}$ using robust sunflower lemmas leading to much smaller gadgets.

Toward lifting theorems with a smaller gadget. As we mentioned, it is a main open problem to improve the gadget size. Hence it is very important to ask what is the limitation of our new approach. Similar to the robust sunflower lemma, we believe that our approach is able to further improve the gadget size to $(\log n)^{O(1)}$. In fact, there is some evidence to show this may be possible.

Lifting theorems for NP are analogous to lifting theorems for P, except that decision tree complexity is changed to *certificate complexity*, and deterministic communication complexity is changed to *rectangle covering number* (cf., [20] for definitions). In an ongoing work of the authors, we can show that in the same regime of parameters as Theorem 1.2, namely when $q = \Omega(n \log^2 n)$, it holds that

$$\mathbf{NP}^{\mathrm{cc}}(f \circ \mathrm{Ind}_q) = \Omega\left(\mathbf{NP}^{\mathrm{dt}}(f) \cdot \log q\right).$$

In addition, if we assume that f is monotone, then this results holds even for $q = (\log n)^{O(1)}$. We believe that it might be possible to push this for a lifting theorem for NP for any function breaking the polynomial-gadget size bottleneck. However, it shows that the limitations of previous approaches (which also hold for their NP analogs) do not hold in our new approach, at least in some settings.

We decided to not include this proof in the current version. The reason is that in this paper we use the robust sunflower lemma of [1, 24] as a black box, and focus on connecting it to lifting theorems. To get the lifting result for NP, one needs to "open the box", prove a generalized version of the robust sunflower lemma, and apply the stronger lemma to the NP lifting setting. In specific, current robust sunflower lemma studied the case $\mathbf{Y}$ is uniform, and in a generalized version, we study the case $\mathbf{Y}$ is "pseudo-uniform". We postpone presenting the details to future work, where hopefully we can prove it for any outer function f.

Paper organization. We present background and results on set systems and robust sunflower lemmas in Section 2. We adapt these to the indexing gadget in Section 3. We prove our main theorem, Theorem 1.2, in Section 4.

2 Set systems

2.1 Notation

We denote the ground set by $[N] = \{1, \ldots, N\}$. We use capital letters S for sets $S \subset [N]$ and bold letters $\mathbf{S}$ for set systems (families of sets). We say that $\mathbf{S}$ is an n-set system if all sets $S \in \mathbf{S}$ have size $|S| \leq n$. A weighted set system is a pair $(\mathbf{S}, \mathcal{D})$, where $\mathbf{S}$ is a set system and $\mathcal{D}$ a distribution supported on $\mathbf{S}$. Given $S \in \mathbf{S}$, we denote by $\mathcal{D}(S)$ the probability that $\mathcal{D}$ assigns to S. Given $\mathbf{S}' \subset \mathbf{S}$, we denote $\mathcal{D}(\mathbf{S}') = \sum_{S \in \mathbf{S}'} \mathcal{D}(S)$. Given a set system $\mathbf{S}$, we denote by $\mathcal{U}_{\mathbf{S}}$ the uniform distribution over $\mathbf{S}$. We shorthand $\Pr_{S \sim \mathcal{U}_{\mathbf{S}}}[\cdot]$ with $\Pr_{S \in \mathbf{S}}[\cdot]$.

2.2 Spread set systems

Informally, spread set systems are set systems where no element, or a small number of elements, is contained in too many sets in the set system.

Definition 2.1 (Spread set systems)**.** *Let $\alpha, k \geq 1$. A weighted set system $(\mathbf{S}, \mathcal{D})$ is (α, k)-spread if for any set $T \subset [N]$:*

$$\Pr_{S \sim \mathcal{D}}[T \subset S] \leq \alpha \cdot k^{-|T|}.$$

A set system $\mathbf{S}$ is (α, k)-spread if there exists a distribution $\mathcal{D}$ supported on $\mathbf{S}$ such that $(\mathbf{S}, \mathcal{D})$ is (α, k)-spread.

We shorthand $(1,k)$-spread as k-spread throughout. Note that if a set system is (α,k)-spread then it is also (k/α)-spread.

Definition 2.2 (Link)**.** *Let* $\mathbf{S}$ *be a set system,* $T \subset [N]$*. The link of* $\mathbf{S}$ *at* T *is*

$$(\mathbf{S})_T = \{S \setminus T : S \in \mathbf{S}, T \subset S\}.$$

Definition 2.3 (Uniform spread set systems)**.** *Let* $\alpha, k \geq 1$*. A set system* $\mathbf{S}$ *is* uniform (α,k)*-spread if* $(\mathbf{S}, \mathcal{U}_{\mathbf{S}})$ *is* (α,k)*-spread. Equivalently, if for any set* $T \subset [N]$ *it holds that*

$$|(\mathbf{S})_T| \leq \alpha \cdot k^{-|T|}|\mathbf{S}|.$$

The following claim shows that large subsets of uniform spread set systems are also uniform spread.

Claim 2.4. *Let* $\mathbf{S}$ *be a set system which is uniform* (α,k)*-spread. Let* $\mathbf{S}' \subset \mathbf{S}$ *of size* $|\mathbf{S}'| = c|\mathbf{S}|$*. Then* $\mathbf{S}'$ *is uniform* $(\alpha/c,k)$*-spread.*

Proof. For any $T \subset [N]$ we have $|(\mathbf{S}')_T| \leq |(\mathbf{S})_T| \leq \alpha k^{-|T|}|\mathbf{S}| \leq (\alpha/c)k^{-|T|}|\mathbf{S}'|$. □

2.3 Covers

Definition 2.5 (Cover)**.** *Let* $\mathbf{S}, \mathbf{R}$ *be set systems. We say that* $\mathbf{R}$ *is a cover for* $\mathbf{S}$ *if for any* $S \in \mathbf{S}$ *there is* $R \in \mathbf{R}$ *such that* $R \subseteq S$*.*

The next claim shows that covers of spread set systems are themselves spread[4].

Claim 2.6. *Let* $\mathbf{S}, \mathbf{R}$ *be set systems, where* $\mathbf{R}$ *is a cover of* $\mathbf{S}$*. Assume that* $\mathbf{S}$ *is* (α,k)*-spread. Then* $\mathbf{R}$ *is also* (α,k)*-spread.*

Proof. Let $\mathcal{D}$ be a distribution supported on $\mathbf{S}$ so that $(\mathbf{S},\mathcal{D})$ is (α,k)-spread. Let $\pi : \mathbf{S} \to \mathbf{R}$ be a map so that $\pi(S) \subset S$ for all $S \in \mathbf{S}$. Define a distribution $\mathcal{D}'$ on $\mathbf{R}$ as follows:

$$\mathcal{D}'(R) = \mathcal{D}(\pi^{-1}(R)).$$

We claim that $(\mathbf{R},\mathcal{D}')$ is (α,k)-spread. To see that, fix $T \subset [N]$. We have:

$$\Pr_{R\sim\mathcal{D}'}[T \subset R] = \Pr_{S\sim D}[T \subset \pi(S)] \leq \Pr_{S\sim D}[T \subset S] \leq \alpha \cdot k^{-|T|}.$$

□

We next refine the definition of cover to *tight covers*, where each element in the cover covers a large fraction of the set system.

Definition 2.7 (Tight cover)**.** *Let* $\beta \leq 1, k \geq 1$*. Let* $\mathbf{S}, \mathbf{R}$ *be set systems where* $\mathbf{R}$ *is a cover of* $\mathbf{S}$*. We say that* $\mathbf{R}$ *is a* tight (β,k)*-cover of* $\mathbf{S}$ *if for any* $R \in \mathbf{R}$ *it holds that*

$$|(\mathbf{S})_R| \geq \beta \cdot k^{-|R|}|\mathbf{S}|.$$

If $\mathbf{S}$ is uniform spread and $\mathbf{R}$ is a tight cover of $\mathbf{S}$, then we get tight control over links of sets in $\mathbf{R}$.

Corollary 2.8. *Let* $\mathbf{S}$ *be a uniform* (α,k)*-spread set system, and let* $\mathbf{R}$ *be a tight* (β,k)*-cover for* $\mathbf{S}$*. Then for every* $R \in \mathbf{R}$ *it holds*

$$\beta \cdot k^{-|R|}|\mathbf{S}| \leq |(\mathbf{S})_R| \leq \alpha \cdot k^{-|R|}|\mathbf{S}|.$$

The next lemma shows that any set system contains either a large sub set system which is uniform spread; or a large sub set system with a tight cover, whose corresponding links are uniform spread.

[4]In the application to lifting, we will apply Claim 2.6 to set systems $\mathbf{S}$ which are uniform spread. Note that the cover $\mathbf{R}$ is spread, but not uniform spread.

Lemma 2.9. *Let $\mathbf{S}$ be a set system and let $k \geq 1$. Then there exists a set system $\mathbf{S}^* \subset \mathbf{S}$ of size $|\mathbf{S}^*| \geq |\mathbf{S}|/2$ for which one of the following holds:*

(i) $\mathbf{S}^*$ is uniform k-spread; or

(ii) There exists a $(1/2, k)$-tight cover $\mathbf{R} = \{R_1, \ldots, R_t\}$ of $\mathbf{S}^*$. Moreover, for each $R_i \in \mathbf{R}$ there exists $\mathbf{S}^i \subset \mathbf{S}^*$ of size $|\mathbf{S}^i| \geq |\mathbf{S}|/2$ such that $\left(\mathbf{S}^i\right)_{R_i}$ is uniform k-spread.

Proof. We define a nested sequence of set systems $\mathbf{S}_0 \supset \mathbf{S}_1 \supset \cdots \supset \mathbf{S}_t$ as follows. Initialize $\mathbf{S}_0 = \mathbf{S}$. Given $\mathbf{S}_m$ for $m \geq 0$, we proceed as follows. First, if $|\mathbf{S}_m| < |\mathbf{S}|/2$ then set $t = m$ and terminate. Next, if $\mathbf{S}_m$ is uniform k-spread, then set $\mathbf{S}^* = \mathbf{S}_m$ and terminate. The remaining case is that $\mathbf{S}_m$ is not uniform k-spread. In this case, by definition there exists some non-empty set $T \subset [N]$ such that $|(\mathbf{S}_m)_T| > k^{-|T|}|\mathbf{S}_m|$. We choose R_{m+1} to be such T of maximal size, and set $\mathbf{S}_{m+1} = \{S \in \mathbf{S}_m : R_{m+1} \not\subset S\}$ to be all the elements in $\mathbf{S}_m$ not covered by R_{m+1}.

Clearly, if during the process some $\mathbf{S}_m$ is uniform k-spread then we are in case (i). So consider the case where no $\mathbf{S}_m$ is uniform k-spread, and so the process ends with $\mathbf{S}_t$ of size $|\mathbf{S}_t| < |\mathbf{S}|/2$. In this case we define $\mathbf{S}^i = \mathbf{S}_{i-1} \setminus \mathbf{S}_i$ for $i = 1, \ldots, t$; $\mathbf{S}^* = \mathbf{S}^1 \cup \ldots \cup \mathbf{S}^t = \mathbf{S} \setminus \mathbf{S}_t$; and $\mathbf{R} = \{R_1, \ldots, R_t\}$. By construction, $\mathbf{R}$ is a cover of $\mathbf{S}^*$, as all sets in $\mathbf{S}^i$ are supersets of R_i. Moreover, $\mathbf{R}$ is a $(1/2, k)$-tight cover for $\mathbf{S}^*$ since

$$|(\mathbf{S}^*)_{R_i}| \geq |\left(\mathbf{S}^i\right)_{R_i}| = |(\mathbf{S}_{i-1})_{R_i}| > k^{-|R_i|}|\mathbf{S}_{i-1}| \geq k^{-|R_i|}|\mathbf{S}|/2 \geq k^{-|R_i|}|\mathbf{S}^*|/2.$$

To conclude, we claim that $\left(\mathbf{S}^i\right)_{R_i}$ is uniform k-spread. Assume not. In this case there is a set $T \subset [N]$ disjoint from R_i, such that

$$|\left(\mathbf{S}^i\right)_{R_i \cup T}| > k^{-|T|}|\left(\mathbf{S}^i\right)_{R_i}|.$$

However, as all the sets in $\mathbf{S}_i$ do not contain R_i by construction, we have $(\mathbf{S}_i)_{R_i} = \emptyset$ and hence $\left(\mathbf{S}^i\right)_{R_i} = (\mathbf{S}_{i-1})_{R_i}$ and $\left(\mathbf{S}^i\right)_{R_i \cup T} = (\mathbf{S}_{i-1})_{R_i \cup T}$. Thus we have

$$|(\mathbf{S}_{i-1})_{R_i \cup T}| > k^{-|T|}|(\mathbf{S}_{i-1})_{R_i}| > k^{-(|R_i|+|T|)}|\mathbf{S}_{i-1}|.$$

This violates the maximality of R_i, as we could have chosen instead $R_i \cup T$. □

2.4 Satisfiability

Definition 2.10. *Let $\mathbf{S}$ be a set system over $[N]$. We say that $\mathbf{S}$ is γ-satisfiable if the following holds. Let $W \subset [N]$ be a uniformly sampled subset. Then*

$$\Pr_W[\exists S \in \mathbf{S}, S \subset W] > 1 - \gamma.$$

The following theorem of Alweiss et al. [1], refined by Frankston et al. [9] and Rao [24], shows that spread set systems are satisfying. Below $K > 0$ is an absolute constant.

Theorem 2.11 (Spread set systems are satisfiable [1, 9, 24])**.** *Let $\mathbf{S}$ be an n-set system. Let $\gamma > 0$ and assume that $\mathbf{S}$ is k-spread for $k \geq K \log(n/\gamma)$. Then $\mathbf{S}$ is γ-satisfiable.*

Corollary 2.12. *Let $\mathbf{S}, \mathbf{Y}$ be set systems over $[N]$. Assume that $|\mathbf{Y}| \geq \gamma \cdot 2^N$ and that $\mathbf{S}$ is an n-set system which is k-spread for $k \geq K \log(n/\gamma)$. Then there exists $S \in \mathbf{S}$ and $Y \in \mathbf{Y}$ so that $S \subset Y$.*

Proof. Apply Theorem 2.11 to $\mathbf{S}$ to derive that $\mathbf{S}$ is γ-satisfiable. Define $\mathbf{W} = \{W \subset [N] : \exists S \in \mathbf{S}, S \subset W\}$, so that $|\mathbf{W}| > (1 - \gamma)2^n$. As $|\mathbf{Y}| \geq \gamma 2^n$, it must be that $\mathbf{Y}, \mathbf{W}$ intersect. □

We will need a generalization of Theorem 2.11 for DNFs.

Definition 2.13. *A DNF is a function $f : \{0,1\}^N \to \{0,1\}$ given as $f(x) = C_1(x) \vee \ldots \vee C_m(x)$, where each clause C_i is a conjunction (AND) of variables x_i or their negation $\neg x_i$. The width of the DNF is the maximal number of variables appearing in a clause.*

Let f be a DNF with clauses $C_1, \ldots, C_m$. We associate with it the set system $\mathbf{S}(f) = \{S_1, \ldots, S_m\}$, where S_i are the variables appearing in the clause C_i (either negated or not, it doesn't matter). Note that if f has width n then $\mathbf{S}(f)$ is an n-set system.

Lemma 2.14. *Let $f : \{0,1\}^N \to \{0,1\}$ be a DNF of width at most n, and let $\mathbf{S}(f)$ be its associated set system. Let $\gamma > 0$ and assume that $\mathbf{S}(f)$ is k-spread for $k \geq K \log(n/\gamma)$. Then*

$$\Pr_{x \in \{0,1\}^N}[f(x) = 1] > 1 - \gamma.$$

Before proving Lemma 2.14, we note that Theorem 2.11 is equivalent to a special case of monotone DNFs, where there are no negated variables.

Proof. The proof is by reduction to Theorem 2.11. Let f be a DNF, and let $\mathbf{S} = \mathbf{S}(f)$ be its associated set system. Let f_{mon} be the monotone DNF corresponding to $\mathbf{S}$, namely

$$f_{\mathrm{mon}}(x) = \bigvee_{S \in \mathbf{S}} \bigwedge_{i \in S} x_i.$$

In other words, if we associate $x \in \{0,1\}^N$ with its corresponding set $X \subset [N]$, then $f_{\mathrm{mon}}(x) = 1$ iff exists $S \in \mathbf{S}$ such that $S \subset X$. Theorem 2.11 then tells us that

$$\Pr[f_{\mathrm{mon}}(x) = 1] > 1 - \gamma.$$

In order to prove that lemma, we will show that the monotone case minimizes the satisfiability probability; namely that

$$\Pr[f(x) = 1] \geq \Pr[f_{\mathrm{mon}}(x) = 1].$$

In order to prove the lemma, we will iterate over $i = 1, \ldots, n$, and in the i-th iteration replace all negated variable $\neg x_i$ with x_i. We will show that this process can only decrease the satisfiability probability. Fix some $i \in [n]$, let g denote the DNF before switching $\neg x_i$ with x_i, and h denote the DNF after the switch. Let $x_{\neq i}$ denote all variables other than x_i, and consider any assignment $a \in \{0,1\}^{N-1}$ to $x_{\neq i}$. Define $g_a(x_i) = g(x_i, x_{\neq i} = a)$ and $h_a(x_i) = h(x_i, x_{\neq i} = a)$ to be the restricted DNFs. We claim that for each choice of a it holds that

$$\Pr[g_a(x_i) = 1] \geq \Pr[h_a(x_i) = 1].$$

Hence also $\Pr[g(x) = 1] \geq \Pr[h(x) = 1]$.

To see that, consider all clauses in g that survive when we set $x_{\neq i} = a$. If none survive that $g_a = h_a = 0$, and if a clause is satisfied then $g_a = h_a = 1$. Otherwise, if g_a contains only clauses with x_i then $h_a = g_a$, and if g_a contains only clauses with $\neg x_i$ then $h_a = \neg g_a$. In all of these cases $\Pr[g_a(x_i) = 1] = \Pr[h_a(x_i) = 1]$. The remaining case is where g_a contains a clause with x_i and a clause with $\neg x_i$, whereas h_a in this case contains only clauses with x_i. In this case $\Pr[g_a(x_i) = 1] = 1 \geq 1/2 = \Pr[h_a(x_i) = 1]$. □

The following corollary is analogous to Corollary 2.12, where we identify $Y \subset [N]$ with its indicator vector $y \in \{0,1\}^N$. The proof is identical so we omit it.

Corollary 2.15. *Let $f : \{0,1\}^N \to \{0,1\}$ be a DNF of width at most n, and let $\mathbf{S}(f)$ be its associated set system. Let $\mathbf{Y} \subset \{0,1\}^N$ be a set system of size $|Y| \geq \gamma \cdot 2^N$. Assume that $\mathbf{S}(f)$ is k-spread for $k \geq K \log(n/\gamma)$. Then there exists $Y \in \mathbf{Y}$ such that $f(Y) = 1$.*

3 Set systems for lifting the index gadget

In this section, we specialize the discussion of set systems to those that arise in the study of lifting of the index gadget.

3.1 Basic setup and definitions

Basic set systems. Fix $q, n \geq 1$, set $N = qn$ and identify the ground set $[N]$ with $[n] \times [q]$. We will work with $\mathbf{X} \subset [q]^n$ and $\mathbf{Y} \subset (\{0,1\}^q)^n$. We identify $X \in [q]^n$ with the set $S_X = \{(i, X_i) : i \in [n]\}$ and $Y \in (\{0,1\}^q)^n$ with the set $S_Y = \{(i,j) : (Y_i)_j = 1\}$ (so basically Y is the indicator vector of S_Y). Thus we view both $\mathbf{X}, \mathbf{Y}$ as set systems over $[n] \times [q]$. Note that $\mathbf{X}$ is an n-set system. We shorthand $Y_{i,j} = (Y_i)_j$.

Partial assignments. Let $I \subset [n]$ and let $I^c = [n] \setminus I$. We denote partial assignments to X by $X' \in [q]^I$ (sometimes we use R instead of X'). Given $X' \in [q]^I, X'' \in [q]^{I^c}$ we denote by $X' \circ X'' \in [q]^n$ their concatenation. Given a set $X' \in [q]^I$ and a set system $\mathbf{X}'' \subset [q]^{I^c}$ define their concatenation to be the set system

$$X' \circ \mathbf{X}'' = \{X' \circ X'' : X'' \in \mathbf{X}''\} \subset [q]^n.$$

We define partial assignments and concatenation for Y analogously.

Links. Let $\mathbf{X} \subset [q]^n$ and $X' \in [q]^I$ be a partial assignment. The link of $\mathbf{X}$ at X' is the set of all elements $X \in \mathbf{X}$ which are consistent with X', restricted to coordinates outside I:

$$(\mathbf{X})_{X'} = \{X'' \in [q]^{I^c} : X' \circ X'' \in \mathbf{X}\}.$$

Note that this is consistent with the link of the set system $\{S_X : X \in \mathbf{X}\}$ at $S_{X'}$. We define links for Y analogously.

Index function. Given $X \in [q]^n$ and $Y \in (\{0,1\}^q)^n$, their index function is

$$\mathrm{Ind}(X, Y) = (Y_{i,X_i} : i \in [n]) \in \{0,1\}^n.$$

More generally, if $X' \in [q]^I$ for $I \subset [n]$ define

$$\mathrm{Ind}(X', Y) = \left(Y_{i,X'_i} : i \in I\right) \in \{0,1\}^I.$$

Restrictions. Given $\mathbf{Y} \subset (\{0,1\}^q)^n$, $R \in [q]^I$ and $v \in \{0,1\}^I$, define the restriction of $\mathbf{Y}$ to R, v to be

$$[\mathbf{Y}]_{R,v} = \{Y \in \mathbf{Y} : \mathrm{Ind}(R, Y) = v\}.$$

Note that restrictions of $\mathbf{Y}$ is defined over the same domain as $\mathbf{Y}$, whereas links are defined over a small domain.

3.2 Useful lemmas

Given $\gamma > 0$ we set $k(\gamma) = 4K \log(2n/\gamma)$, where K is the absolute constant from Theorem 2.11.

The first lemma shows that if $\mathbf{X}$ is spread and $\mathbf{Y}$ is large, then $\mathrm{Ind}(X, Y)$ for $X \in \mathbf{X}, Y \in \mathbf{Y}$ attains all possible values.

Lemma 3.1. *Let $\mathbf{X} \subset [q]^n, \mathbf{Y} \subset (\{0,1\}^q)^n$,$\gamma > 0$ and set $k = k(\gamma)$. Assume that $\mathbf{X}$ is k-spread and $|\mathbf{Y}| \geq \gamma \cdot 2^{qn}$. Then*

$$\{\mathit{Ind}(X, Y) : X \in \mathbf{X}, Y \in \mathbf{Y}\} = \{0,1\}^n.$$

Proof. Fix $z \in \{0,1\}^n$. We will show that there exists $X \in \mathbf{X}, Y \in \mathbf{Y}$ such that $\mathrm{Ind}(X, Y) = z$. Given $Y \in \mathbf{Y}$, define a new set $Y_z \in (\{0,1\}^q)^n$ by $(Y_z)_{i,j} = Y_{i,j} \oplus z_i \oplus 1$. Observe that $\mathrm{Ind}(X, Y) = z$ iff $S_X \subset S_{Y_z}$. Define a set system $\mathbf{Y}_z = \{Y_z : Y \in \mathbf{Y}\}$ and note that $|\mathbf{Y}_z| = |\mathbf{Y}|$. Now, Corollary 2.12 implies that there is $X \in \mathbf{X}, Y_z \in \mathbf{Y}_z$ such that $S_X \subset S_{Y_z}$. This implies that $\mathrm{Ind}(X, Y) = z$. □

The next lemma shows that given $\mathbf{Y}$, we can remove at most half its elements, such that in the resulting set system each restriction is either empty or large.

Lemma 3.2. *Let $\mathbf{Y} \subset (\{0,1\}^q)^n$. Then there exists a set system $\mathbf{Y}' \subset \mathbf{Y}$ of size $|\mathbf{Y}'| \geq |\mathbf{Y}|/2$ with the following property. For every non-empty $I \subset [n]$, partial assignment $R \in [q]^I$ and value $v \in \{0,1\}^I$, the restriction $[\mathbf{Y}']_{R,v}$ is either empty or satisfies*

$$|[\mathbf{Y}']_{R,v}| \geq |\mathbf{Y}|/(8qn)^{|I|}.$$

Proof. We apply a greedy pruning procedure. Set $m = 8qn$. Initialize $\mathbf{Y}' = \mathbf{Y}$. As long as there exists a non-empty $I \subset [n]$, partial assignment $R \in [q]^I$ and value $v \in \{0,1\}^I$ such that $0 < |[\mathbf{Y}']_{R,v}| < |\mathbf{Y}|/m^{|I|}$, remove the elements in $[\mathbf{Y}']_{R,v}$ from $\mathbf{Y}'$. We claim that this process stops after removing at most half the elements from $\mathbf{Y}$, giving us the claimed $\mathbf{Y}'$.

To see why, note that at each phase, if we consider a partial assignment $R \in [q]^I$ and value $v \in \{0,1\}^I$, then we remove at most $|\mathbf{Y}|/m^{|I|}$ elements. Clearly, each pair (R, v) can be used at most once. The number of pairs (R, v) with $|I| = k$ is $\binom{n}{k}(2q)^k$. Thus we can bound the fraction of elements removed by

$$\frac{|\mathbf{Y} \setminus \mathbf{Y}'|}{|\mathbf{Y}|} \leq \sum_{k=1}^{n} \binom{n}{k} (2q)^k m^{-k} \leq \sum_{k=1}^{n} (2qn/m)^k \leq \sum_{k=1}^{n} 4^{-k} \leq 1/2.$$

□

The final lemma shows that if $\mathbf{X}, \mathbf{Y}$ are such that $\mathbf{X}$ is somewhat spread, then we can restore spreadness by restricting to a partial assignment for which the restriction of $\mathbf{Y}$ is large.

Lemma 3.3. *Let $\mathbf{X} \subset [q]^n, \mathbf{Y} \subset (\{0,1\}^q)^n$, $\gamma > 0$ and set $k = k(\gamma)$. Assume that $\mathbf{X}$ is uniform $(2,k)$-spread and $|\mathbf{Y}| \geq \gamma 2^{qn}$. Then there exists a partial assignment $R \in [q]^I$ and a subset $\tilde{\mathbf{X}} \subset \mathbf{X}$ of size $|\tilde{\mathbf{X}}| \geq |\mathbf{X}|/2$ such that the following holds:*

1. *Let $\mathbf{X}' = \left(\tilde{\mathbf{X}}\right)_R$. Then $|\mathbf{X}'| \geq (1/4)k^{-|I|}|\mathbf{X}|$ and $\mathbf{X}'$ is uniform k-spread.*
2. *For any $v \in \{0,1\}^I$ we have $|[\mathbf{Y}]_{R,v}| \geq (8qn)^{-|I|}|\mathbf{Y}|$.*

Proof. Apply Lemma 2.9 to $\mathbf{X}$ and let $\mathbf{X}^* \subset \mathbf{X}$ be the guaranteed subset of size $|\mathbf{X}^*| \geq |\mathbf{X}|/2$. If case (i) holds, namely $\mathbf{X}^*$ is uniform k-spread, then we take $\tilde{\mathbf{X}} = \mathbf{X}^*$ and R to be the empty partial assignment. In this case $I = \emptyset$, $\mathbf{X}' = \tilde{\mathbf{X}}$, $\mathbf{Y}' = \mathbf{Y}$ and both conditions hold.

So, assume that case (ii) holds. Namely, there is a $(1/2,k)$-tight cover $\mathbf{R} = \{R_1, \ldots, R_t\}$ of $\mathbf{X}^*$, such that for every R_i there exists $\mathbf{X}^i \subset \mathbf{X}^*$ of size $|\mathbf{X}^i| \geq |\mathbf{X}|/2$ for which $\left(\mathbf{X}^i\right)_{R_i}$ is uniform k-spread. Observe that each R_i corresponds to a partial assignment since each R_i is a subset of S_X for some $X \in \mathbf{X}$. Thus we identify R_i with an element of $[q]^{I_i}$ for some non-empty I_i.

Next, apply Lemma 3.2 to $\mathbf{Y}$ to obtain the claimed sub set system $\mathbf{Y}' \subset \mathbf{Y}$. Our goal will be to find $R = R_i \in \mathbf{R}$ with the following property: for each $v \in \{0,1\}^I$ where $I = I_i$, the set $[\mathbf{Y}']_{R,v}$ is non empty. Assume for a minute we can guarantee the existence of such R. In this case we set $\tilde{\mathbf{X}} = \mathbf{X}^i$. Note that $\mathbf{X}' = \left(\mathbf{X}^i\right)_{R_i}$ and hence $\mathbf{X}'$ is guaranteed to be uniform k-spread. By the assumption on R, for each $v \in \{0,1\}^I$ we know that $[\mathbf{Y}']_{R,v}$ is not empty, and hence by the property of $\mathbf{Y}'$ guaranteed by Lemma 3.2 we have that $|\mathbf{Y}'| \geq (8qn)^{-|I|}|\mathbf{Y}|$. Lastly, we need to verify that $|\mathbf{X}'| \geq (1/4)k^{-|I|}|\mathbf{X}|$. This follows since $\mathbf{R}$ is a $(1/2,k)$-tight cover of $\mathbf{X}^*$, and hence $|\mathbf{X}'| \geq (1/2)k^{-|I|}|\mathbf{X}^*|$, combined with the fact that $|\mathbf{X}^*| \geq |\mathbf{X}|/2$.

Finally, we need to show that there exists $R_i \in \mathbf{R}$ as claimed. Assume not; then for every $R_i \in \mathbf{R}$ with $R_i \in [q]^{I_i}$, there exists $v_i \in \{0,1\}^{I_i}$ such that $\mathbf{Y}'_{R_i,v_i}$ is empty. We next construct a DNF checking this. Recall that $N = qn$ and that we identify $Y \in (\{0,1\}^q)^n$ with $y \in \{0,1\}^N$. Given each (R_i, v_i), let $C_i(y)$ be the clause specifying that $\mathrm{Ind}(R_i, Y) = v_i$. Note that C_i is a conjunction of $|I_i|$ variables $y_{i,j}$ or their negations. Define

$$f(y) = C_1(y) \vee \cdots \vee C_t(y)$$

and note that f is a DNF of width at most n. By assumption, $f(Y) = 0$ for all $Y \in \mathbf{Y}'$. On the other hand, we will show that $\mathbf{S}(f)$ is spread, and obtain a contradiction by using Corollary 2.15.

Note that $\mathbf{S}(f) = \mathbf{R}$. We know that $\mathbf{R}$ is a cover of $\mathbf{X}^*$, that $\mathbf{X}^*$ is a subset of $\mathbf{X}$ with at least half the elements, and then $\mathbf{X}$ is uniform $(2,k)$-spread. Thus by Claim 2.4 we have that $\mathbf{X}^*$ is uniform $(4,k)$-spread, and by Claim 2.6 that $\mathbf{R}$ is $(4,k)$-spread. This implies that $\mathbf{R}$ is $(k/4)$-spread. In addition $|\mathbf{Y}'| \geq |\mathbf{Y}|/2 \geq (\gamma/2)2^N$. Hence we can apply Corollary 2.15 to it, as we set $k = k(\gamma) = 4K\log(2n/\gamma)$. This shows that it is impossible that $f(Y) = 0$ for all $Y \in \mathbf{Y}'$, completing the proof. □

4 Lifting for P

We use the properties of set systems from previous sections to prove a deterministic lifting theorem when the gadget size is nearly linear. Let $f : \{0,1\}^n \to \{0,1\}$ be an arbitrary boolean function. We define a two-party function $(f \circ \mathrm{Ind})(X,Y) = f(\mathrm{Ind}(X,Y))$ where $X \in [q]^n, Y \in (\{0,1\}^q)^n$.

Theorem 4.1 (Main lifting theorem)**.** *Let $f : \{0,1\}^n \to \{0,1\}$, $q = \Omega(n\log^2 n)$ and set $F = f \circ Ind_q^n$. Then*

$$\mathbf{P}^{\mathrm{cc}}(F) = \Omega(\mathbf{P}^{\mathrm{dt}}(f)) \cdot \log\left(\frac{q}{n\log^2 q}\right).$$

Before proving Theorem 4.1, we make a couple of comments. It is straightforward that $\mathbf{P}^{\mathrm{cc}}(F) = O(\mathbf{P}^{\mathrm{dt}}(f)\log q)$, and hence our bound is tight up to logarithmic factors. Furthermore, if $q \geq n^{1+c}$ for any $c > 0$ then we obtain $\mathbf{P}^{\mathrm{cc}}(F) = \Omega(\mathbf{P}^{\mathrm{dt}}(f)\log q)$, in which case our bound is tight up to constants. This in particular recovers the result of Raz and McKenzie [25].

We now move the prove Theorem 4.1. Our goal is to show that if F has a low communication protocol, then f has low depth decision tree. Suppose we have a protocol Π for F with at most Δ bits of communication. For ease of notation, we assume (without loss of generality) that in each round only one bit is communicated. We will use the protocol Π for F to simulate a decision tree for f with small depth. Consider an arbitrary input string $z \in \{0,1\}^n$ on which we are trying to compute f. Fix a parameter $k = \Theta(n\log^2 n) \leq q/2$ to be chosen later.

Setup. We maintain set systems $\mathbf{X} \subset [q]^n, \mathbf{Y} \subset (\{0,1\}^q)^n$, corresponding to inputs consistent with the current execution of the protocol. In addition, we maintain a restriction $\rho \in \{0,1,*\}^n$ which specifies which input bits were read, and what values they have. We denote by $\mathrm{Fix}(\rho) = \rho^{-1}(\{0,1\})$ and $\mathrm{Free}(\rho) = \rho^{-1}(*)$ the partition of $[n]$ to coordinates fixed by ρ or that are still free.

Consistency. We will maintain consistency of the restriction ρ with the the sets $\mathbf{X}, \mathbf{Y}$ and the input z.

Definition 4.2 ($\mathbf{X}$-consistency)**.** *We say that $\mathbf{X} \subset [q]^n$ is consistent with $\rho \in \{0,1,*\}^n$ if $\mathbf{X} = X' \circ \mathbf{X}''$, where $X' \in [q]^{Fix(\rho)}$ and $\mathbf{X}'' \subset [q]^{Free(\rho)}$.*

Definition 4.3 ($\mathbf{Y}$-consistency)**.** *We say that $\mathbf{Y} \subset (\{0,1\}^q)^n$ is consistent with $\rho \in \{0,1,*\}^n$ if $\mathbf{Y} = Y' \circ \mathbf{Y}''$, where $Y' \in (\{0,1\}^q)^{Fix(\rho)}$ and $\mathbf{Y}'' \subset (\{0,1\}^q)^{Free(\rho)}$.*

Definition 4.4 (z-consistency)**.** *We say that $z \in \{0,1\}^n$ is consistent with $\rho \in \{0,1,*\}^n$ if $\rho_i \in \{*, z_i\}$ for all $i \in [n]$.*

Variables used in the simulation. We initialize $\mathbf{X}_0 = [q]^n$, $\mathbf{Y}_0 = (\{0,1\}^q)^n$, $\rho_0 = *^n$. At the end of round r, we will have that $\mathbf{X}_r \subset [q]^n, \mathbf{Y}_r \subset (\{0,1\}^q)^n$ are subsets of the inputs consistent with the protocol so far, and $\rho_r \in \{0,1,*\}^n$ is a restriction.

Invariants. We will maintain the following invariant: $\mathbf{X}_r, \mathbf{Y}_r, z$ are consistent with ρ_r. We use the following notation: $\mathbf{X}_r = X'_r \circ \mathbf{X}''_r$ where $X'_r \in [q]^{\mathrm{Fix}(\rho_r)}, \mathbf{X}'' \subset [q]^{\mathrm{Free}(\rho_r)}$ and $\mathbf{Y}_r = Y'_r \circ \mathbf{Y}''_r$ where $Y'_r \in (\{0,1\}^q)^{\mathrm{Fix}(\rho_r)}, \mathbf{Y}'' \subset (\{0,1\}^q)^{\mathrm{Free}(\rho_r)}$. In addition, we maintain the invariant that $\mathrm{Ind}(X'_r, Y'_r) = z_{\mathrm{Fix}(\rho_r)}$, and that $\mathbf{X}''_r$ is uniform k-spread.

Protocol structure. Assume we are at the beginning of round r, and have sets $\mathbf{X}_{r-1}, \mathbf{Y}_{r-1}$ of possible inputs for Alice and Bob. If it is Alice's turn, then the protocol Π partitions $\mathbf{X}_{r-1}$ it into two sets $\mathbf{X}_{r-1,0}, \mathbf{X}_{r-1,1}$, depending on the bit sent by Alice. Similarly, if it is Bob's turn, then the protocol partitions $\mathbf{Y}_{r-1}$ is split into $\mathbf{Y}_{r-1,0}, \mathbf{Y}_{r-1,1}$, depending on the bit sent by Bob.

Simulating the protocol. For each round $r = 1, \ldots, \Delta$ of the protocol Π, given $\mathbf{X}_{r-1}, \mathbf{Y}_{r-1}$ which are the output of the previous round, do the following:

1. If it is Bob's turn in the protocol, then Bob sends a bit $b \in \{0,1\}$ that maximizes the size of $\mathbf{Y}_{r-1,b}$. Set $\mathbf{X}_r = \mathbf{X}_{r-1}, \mathbf{Y}_r = \mathbf{Y}_{r-1,b}$ and $\rho_r = \rho_{r-1}$.
2. If it is Alice's turn in the protocol, then Alice sends a bit $b \in \{0,1\}$ that maximizes the size of $\bar{\mathbf{X}}_r = \mathbf{X}_{r-1,b}$. Next, do the following:
   (a) Observe that $\bar{\mathbf{X}}_r = X'_{r-1} \circ \bar{\mathbf{X}}''_r$ for some $\bar{\mathbf{X}}''_r \subset \mathbf{X}''_{r-1}$.
   (b) Apply Lemma 3.3 to $\bar{\mathbf{X}}''_r, \mathbf{Y}''_{r-1}$ (we will argue later that they satisfy the conditions of the lemma). Find $R_r \in [q]^{I_r}, \tilde{\mathbf{X}}_r$ that satisfy the conclusions of the lemma.
   (c) Query z_i for $i \in I_r$.
   (d) Define ρ_r as follows: $(\rho_r)_i = (\rho_{r-1})_i$ if $i \notin I_r$, $(\rho_r)_i = z_i$ for $i \in I_r$.
   (e) Set $\mathbf{X}_r = \mathbf{X}'_{r-1} \circ \left(\tilde{\mathbf{X}}_r\right)_{R_r}$.
   (f) Set $\bar{\mathbf{Y}}_r = \{Y \in \mathbf{Y}_{r-1} : \mathrm{Ind}(Y, R_r) = z_{I_r}\}$. Choose $Y'_r \in (\{0,1\}^q)^{\mathrm{Fix}(\rho_r)}$ to maximize the size of $\mathbf{Y}''_r = \left(\bar{\mathbf{Y}}_r\right)_{Y'_r}$. Set $\mathbf{Y}_r = Y'_r \circ \mathbf{Y}''_r$.

When the protocol terminates, we output the value computed by the protocol as the value of $f(z)$. We first show the following invariants about the simulation.

Lemma 4.5. *There is a constant c_1 such that if $q > k > c_1 \Delta \log q$, then the above algorithm maintains the following invariants at all times $r \geq 0$:*

(1) $\mathbf{X}_r, \mathbf{Y}_r, z$ *are consistent with* ρ_r.

(2) $Ind(X'_r, Y'_r) = z_{Fix(\rho_r)}$.

(3) $\mathbf{X}''_r$ *is uniform* k*-spread.*

(4) $|\mathbf{X}''_r| \geq 8^{-r}(q/k)^{|Fix(\rho_r)|} \cdot q^{|Free(\rho_r)|}$. *In particular,* $|Fix(\rho_r)| \leq 3r/\log(q/k)$.

(5) $|\mathbf{Y}''_r| \geq 2^{-r}(8qn)^{-|Fix(\rho_r)|} \cdot 2^{q|Free(\rho_r)|}$.

Proof. The claim clearly holds for round 0. Suppose the conditions hold at the end of round $r-1$. We next argue that the conditions hold at the end of round r.

If it is Bob's turn in the protocol, then the claim follows easily as $|\mathbf{Y}_r| \geq |\mathbf{Y}_{r-1}|/2$ and the other quantities do not change, so suppose it is Alice's turn in the protocol. By parts (4) and (5) of the inductive assumption,

$$|\mathbf{Y}''_{r-1}| \geq 2^{-(r-1)}(8qn)^{-3(r-1)/\log(q/k)} 2^{q|\mathrm{Free}(\rho_{r-1})|} \equiv \gamma 2^{q|\mathrm{Free}(\rho_{r-1})|},$$

where $\gamma = 2^{-(r-1)}(8qn)^{-3(r-1)/\log(q/k)} \geq q^{-O(r)}$.

Observe that $\bar{\mathbf{X}}''_r$ computed in step 2 is uniform $(2,k)$-spread by Claim 2.4, since $|\bar{\mathbf{X}}''_r| \geq |\mathbf{X}''_{r-1}|/2$ and since by induction $\mathbf{X}''_{r-1}$ is uniform k-spread. In order apply Lemma 3.3 as in Step 2 (b) of the simulation, We further need $k \geq 4K \log(2n/\gamma)$. Now,

$$4K \log(2n/\gamma) = O(r \log q) = O(\Delta \log q).$$

Therefore, if $k > c_1 \Delta \log q$ for some absolute constant c_1, then $\bar{\mathbf{X}}''_r, \mathbf{Y}''_{r-1}$ satisfy the conditions of Lemma 3.3 as desired.

Finally, note that $\mathrm{Fix}(\rho_r)$ is the disjoint union of $\mathrm{Fix}(\rho_{r-1})$ and I_r, and in particular $|\mathrm{Fix}(\rho_r)| = |\mathrm{Fix}(\rho_{r-1})| + |I_r|$. That the invariants hold after round r now follows easily. Part (1) follows from the specific way $\mathbf{X}_r, \mathbf{Y}_r$ are chosen. Parts (2), (3), (4) then follow inductively from the guarantees of Lemma 3.3 as used in steps 2(b) - (e) of the algorithm.

For part (5), we have by Lemma 3.3 that $|\bar{\mathbf{Y}}_r| \geq (8qn)^{-|I_r|}|\mathbf{Y}''_{r-1}| = (8qn)^{-|I_r|}|\mathbf{Y}_{r-1}|$ and hence by induction that $|\bar{\mathbf{Y}}_r| \geq 2^{-r}(8qn)^{-|\mathrm{Fix}(\rho_r)|} \cdot 2^{q|\mathrm{Free}(\rho_{r-1})|}$. By averaging there must exist $Y'''_r \in (\{0,1\}^q)^{I_r}$ such that $|(\bar{\mathbf{Y}}_r)_{Y'''_r}| \geq 2^{-r}(8qn)^{-|\mathrm{Fix}(\rho_r)|} \cdot 2^{q|\mathrm{Free}(\rho_r)|}$. To conclude, observe that the definition in step 2 (f) in the protocol is equivalent to taking $Y'_r = Y'_{r-1} \circ Y'''_r$, from which the bound holds. □

Lemma 4.6. *There is a constant c_2 such that if $q > k > c_2 \Delta \log q$, then the above simulations computes $f(z)$ correctly and queries at most $O(\Delta/\log(q/k))$ bits of z.*

Proof. We take $c_2 \geq c_1$, so we can apply Lemma 4.5. Let $I = \mathrm{Fix}(\rho_\Delta)$ and $J = \mathrm{Free}(\rho_\Delta)$. Consider the termination of the simulation. By part (4) of Lemma 4.5, the number of queries made on z is at most $|I| \leq 3\Delta/\log(q/k)$.

In order to show correctness, we argue that $z \in \{\mathrm{Ind}(X,Y) : X \in \mathbf{X}_\Delta, Y \in \mathbf{Y}_\Delta\}$. This implies the simulation computes $f(z)$ correctly. First, by part (2) of Lemma 4.5, we have that $\mathrm{Ind}(X',Y')_I = z_I$. We would next apply Lemma 3.1 to obtain that

$$\{\mathrm{Ind}(X'',Y'') : X'' \in \mathbf{X}''_\Delta, Y'' \in \mathbf{Y}''_\Delta\} = \{0,1\}^J,$$

and hence in particular $z = \mathrm{Ind}(X,Y)$ for some $X \in \mathbf{X}_\Delta, Y \in \mathbf{Y}_\Delta$.

We next verify the conditions needed to apply Lemma 3.1. Part (3) of Lemma 4.5 gives that $\mathbf{X}''_\Delta$ is k-spread. The above arguments and part (5) of Lemma 4.5 gives

$$|\mathbf{Y}''_\Delta| \geq 2^{-\Delta}(8qn)^{-3\Delta/\log(q/k)} 2^{q|J|} \equiv \gamma 2^{q|J|},$$

where $\gamma = 2^{-\Delta}(8qn)^{-3\Delta/\log(q/k)}$. In order to apply Lemma 4.5, we need $k \geq 4K\log(2n/\gamma) = O(\Delta \log(q))$. Thus as long as $k \geq c_1 \Delta \log q$ for a large enough constant c_1, the conditions holds. □

Our main lifting theorem follows easily from the above:

Proof of Theorem 4.1. Let $d = \mathbf{P}^{\mathrm{dt}}(f)$. Then, clearly $\mathbf{P}^{\mathrm{cc}}(F) = \Delta \leq d \log q$. Let Π be such a protocol and run a simulation as outlined in the section for $k = 2c_1 n \log^2 q$ for c_1 as in the previous lemma. We can do so as $2c_1 \Delta \log q \leq 2c_1 d \log^2 q \leq 2c_1 n \log^2 q < q$ if $q > c_2 n \log^2 n$ for a sufficiently big c_2.

The simulation gives us a decision tree for f with depth at most $O(\Delta/\log(q/k))$. Therefore, we must have

$$\Delta = \Omega(\mathbf{P}^{\mathrm{dt}}(f)\log(q/k)) = \Omega(\mathbf{P}^{\mathrm{dt}}(f)) \cdot \log\left(\frac{q}{n\log^2 q}\right).$$

□

References

[1] R. Alweiss, S. Lovett, K. Wu, and J. Zhang. Improved bounds for the sunflower lemma. *arXiv preprint arXiv:1908.08483*, 2019.

[2] Y. Babichenko and A. Rubinstein. Communication complexity of approximate nash equilibria. In H. Hatami, P. McKenzie, and V. King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 878–889. ACM, 2017.

[3] S. O. Chan, J. R. Lee, P. Raghavendra, and D. Steurer. Approximate constraint satisfaction requires large LP relaxations. *J. ACM*, 63(4):34:1–34:22, 2016.

[4] A. Chattopadhyay, Y. Filmus, S. Koroth, O. Meir, and T. Pitassi. Query-to-communication lifting for bpp using inner product. *arXiv preprint arXiv:1904.13056*, 2019.

[5] A. Chattopadhyay, Y. Filmus, S. Koroth, O. Meir, and T. Pitassi. Query-to-communication lifting using low-discrepancy gadgets. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:103, 2019.

[6] A. Chattopadhyay, N. S. Mande, and S. Sherif. The log-approximate-rank conjecture is false. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 42–53, 2019.

[7] S. F. de Rezende, O. Meir, J. Nordström, T. Pitassi, R. Robere, and M. Vinyals. Lifting with simple gadgets and applications to circuit and proof complexity. *arXiv preprint arXiv:2001.02144*, 2020.

[8] P. Erdős and R. Rado. Intersection theorems for systems of sets. *Journal of the London Mathematical Society*, 35(1):85–90, 1960.

[9] K. Frankston, J. Kahn, B. Narayanan, and J. Park. Thresholds versus fractional expectation-thresholds. *arXiv preprint arXiv:1910.13433*, 2019.

[10] M. Göös. Lower bounds for clique vs. independent set. In V. Guruswami, editor, *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 1066–1076. IEEE Computer Society, 2015.

[11] M. Göös, R. Jain, and T. Watson. Extension complexity of independent set polytopes. *SIAM Journal on Computing*, 47(1):241–269, 2018.

[12] M. Göös, P. Kamath, T. Pitassi, and T. Watson. Query-to-communication lifting for p np. *computational complexity*, 28(1):113–144, 2019.

[13] M. Göös, S. Lovett, R. Meka, T. Watson, and D. Zuckerman. Rectangles are nonnegative juntas. *SIAM Journal on Computing*, 45(5):1835–1869, 2016.

[14] M. Göös and T. Pitassi. Communication lower bounds via critical block sensitivity. In D. B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 847–856. ACM, 2014.

[15] M. Göös, T. Pitassi, and T. Watson. Query-to-communication lifting for BPP. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:53, 2017.

[16] M. Göös, T. Pitassi, and T. Watson. Deterministic communication vs. partition number. *SIAM Journal on Computing*, 47(6):2435–2450, 2018.

[17] M. Göös and A. Rubinstein. Near-optimal communication lower bounds for approximate nash equilibria. In M. Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 397–403. IEEE Computer Society, 2018.

[18] S. Jukna. *Boolean function complexity: advances and frontiers*, volume 27. Springer Science & Business Media, 2012.

[19] P. K. Kothari, R. Meka, and P. Raghavendra. Approximating rectangles by juntas and weakly-exponential lower bounds for LP relaxations of csps. In H. Hatami, P. McKenzie, and V. King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 590–603. ACM, 2017.

[20] E. Kushilevitz. Communication complexity. In *Advances in Computers*, volume 44, pages 331–360. Elsevier, 1997.

[21] J. R. Lee, P. Raghavendra, and D. Steurer. Lower bounds on the size of semidefinite programming relaxations. In R. A. Servedio and R. Rubinfeld, editors, *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 567–576. ACM, 2015.

[22] X. Li, S. Lovett, and J. Zhang. Sunflowers and quasi-sunflowers from randomness extractors. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[23] T. Pitassi and R. Robere. Strongly exponential lower bounds for monotone computation. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:188, 2016.

[24] A. Rao. Coding for sunflowers. *arXiv preprint arXiv:1909.04774*, 2019.

[25] R. Raz and P. McKenzie. Separation of the monotone nc hierarchy. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 234–243. IEEE, 1997.

[26] R. Robere, T. Pitassi, B. Rossman, and S. A. Cook. Exponential lower bounds for monotone span programs. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 406–415. IEEE, 2016.

[27] A. A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In C. Dwork, editor, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 85–94. ACM, 2008.

[28] A. A. Sherstov. The pattern matrix method. *SIAM J. Comput.*, 40(6):1969–2000, 2011.

[29] A. A. Sherstov. Communication lower bounds using directional derivatives. *J. ACM*, 61(6):34:1–34:71, 2014.

[30] X. Wu, P. Yao, and H. S. Yuen. Raz-mckenzie simulation with the inner product gadget.

[31] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, 1991.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il