Electronic Colloquium on Computational Complexity, Report No. 156 (2025)

Unifying the Landscape of Super-Logarithmic Dynamic Cell-Probe Lower Bounds

Young Kun Ko

Department of Computer Science and Engineering, Pennsylvania State University

Email: ykko@psu.edu

October 22, 2025

Abstract

We prove a general translation theorem for converting one-way communication lower bounds over a product distribution to dynamic cell-probe lower bounds.

Specifically, we consider a class of problems considered in [Pat10] where:

- $S_1, \ldots, S_m \in \{0,1\}^n$ are given and publicly known.
- $T \in \{0,1\}^n$ is a sequence of updates, each taking t_u time.
- For a given $Q \in [m]$, we must output $f(S_Q, T)$ in t_q time.

Our main result shows that for a "hard" function f, for which it is difficult to obtain a non-trivial advantage over random guessing with one-way communication under some product distribution over S_Q and T (for example, a uniform distribution), the above explicit dynamic cell-probe problem must have $\max\{t_u, t_q\} \geq \tilde{\Omega}(\log^{3/2}(n))$ if $m = \Omega(n^{0.99})$. This result extends and unifies the super-logarithmic dynamic data structure lower bounds from [LWY20] and [LY25] into a more general framework.

From a technical perspective, our approach merges the cell-sampling and chronogram techniques developed in [LWY20] and [LY25] with the new static data structure lower bound methods from [KW20] and [Ko25], thereby merging all known state-of-the-art cell-probe lower-bound techniques into one.

As a direct consequence of our method, we establish a super-logarithmic lower bound against the Multiphase Problem [Pat10] for the case where the data structure outputs the Inner Product (mod 2) of S_Q and T. We suspect further applications of this general method towards showing super-logarithmic dynamic cell-probe lower bounds. We list some example applications of our general method, including a novel technique for a one-way communication lower bound against small-advantage protocols for a product distribution using average min-entropy, which could be of independent interest.

ISSN 1433-8092

1 Introduction

We consider the following general class of dynamic data structure problems, initiated by the so-called Multiphase Program [Pat10].

- $S_1, \ldots, S_m \in \{0,1\}^n$ are given and publicly known.[1]
- $T \in \{0,1\}^n$ is given as a sequence of updates, using t_u time per update.
- For any given $i \in [m]$, output $f(S_i, T)$ using t_q time.

We analyze this dynamic data structure problem in the cell-probe model [Yao81], the most powerful model of computation for data structures. An input is pre-processed into a data structure of s cells, each with a word of size w-bits. In this model, we only charge for memory accesses (probes) of cells, while all computation on top of the probed cells is free. Because of its unimaginable strength, a lower bound in this model applies to any reasonable data structure.

The Multiphase Conjecture [Pat10, Tho13] states that if f is two-party Disjointness between S_i and T with $m = \text{poly}(n)$, then $\max\{t_u, t_q\} \geq n^\delta$ for some constant $\delta > 0$. The Multiphase Conjecture then implies polynomial lower bounds for Graph Reachability (for directed graphs), Dynamic Shortest Path (for undirected graphs), and other interesting dynamic problems.

In this work, we consider a generalized version of the Multiphase Program, considering f to be any function from a general class of functions. For instance, we consider f to be a "lifted" version of some function ψ. Suppose we divide n into blocks of size k each, denoted by $S_i[j], T[j]$ where $j \in [n/k]$. Let g be some inner gadget, $g : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}$. Some standard inner gadgets g include indexing functions (where the first k bits have only one 1) or functions that output the inner product of the two k-bit strings. Then consider f's of the form

$$f(S_i, T) := \psi(g(S_i[1], T[1]), \ldots, g(S_i[n/k], T[n/k]).$$

For example, this captures both Disjointness and Inner Product (mod 2), by setting $k = 1$ and g as the bitwise-AND function. If we take ψ as $\bigvee$, f is then Disjointness, while if we take ψ as $\bigoplus$, f is then Inner Product (mod 2). This is the class of functions studied extensively in so-called lifting theorems (see [RY18, BT22] and references therein).

1.1 Our Result

Our main result is to show a "lifting" type result for the dynamic data structure problem. If f is "hard" against an $n/\text{poly}\log(n)$ length message (i.e., a lower bound against one-way communication) under a product distribution, then the Multiphase Problem for such a function f has $\max\{t_u, t_q\} \geq \tilde{\Omega}(\log^{3/2}(n))$. Thus we lift f which is hard against one-way communication (a more tractable task) to hard against dynamic data structures, albeit only up to super-logarithmic hardness. As a corollary, we obtain the state-of-the-art lower bound against the Multiphase Conjecture [Pat10, CGL15, BL15, KW20, DL20, Ko25] when the function in question is Inner Product (mod 2) instead of Disjointness, improving upon the previously known $\Omega(\log(n))$ bound [CGL15, BL15].

Formally, we consider the following set of explicit dynamic data structure problems, which we denote as the generalized Multiphase Problem:

[1] In the original Multiphase Problem, these are given as pre-processing inputs and are pre-processed into a data structure.

- $S_1, \ldots, S_m \in \{\{0,1\}^k\}^{n/k}$ are given, where each S_Q is divided into n/k blocks $S_Q[1], \ldots, S_Q[n/k]$, each of size k.
- $T = (T[1], \ldots, T[n/k]) \in \{\{0,1\}^k\}^{n/k}$ is given as a sequence of updates, using t_u time per update (in the cell-probe model with word size w).
- For any given $Q \in \mathcal{Q} = [m]$, output

$$f(S_Q, T) = \psi\left(g(S_Q[1], T[1]), \ldots, g(S_Q[n/k], T[n/k])\right)$$

  in t_q time.

Problem 1: Generalized Multiphase Problem

Note that the functions f we consider are precisely the set of functions used in the Pattern Matrix Method [She11], and Lifting Theorems ([RY18, BT22] and references therein). Our main result is showing super-logarithmic lower bounds (as in [LWY20, LY25]) for a general class of "hard" f's which includes inner product (mod 2).

Before fully stating our main result, we would like to formally describe which f's are "hard" for our proof. As we also use the **chronogram method**, we will divide the n/k blocks of T into ℓ epochs, $\{T_i\}_{i=1}^{\ell}$, where T_i consists of n_i blocks and $|T_i| = k \cdot n_i \approx k \cdot \gamma^i$. γ is a parameter to be defined within the proof, and $\sum n_i = n/k$. If we fix all epochs other than i, this results in our function f becoming

$$f_i(S_Q, T_i) := f(S_Q, T)$$

With fixed S_Q^{-i} and T_{-i}, we consider the real-valued matrix $\Psi_i|_{S_Q^{-i}, T_{-i}} \in [-1,+1]^{2^{|S_Q^i|} \times 2^{|T_i|}}$ (assuming without loss of generality that $\Pr_{S_Q^i, T_i}[f(S_Q,T) = +1] \geq \Pr_{S_Q^i, T_i}[f(S_Q,T) = -1]$)

$$\Psi_i(S_Q^i, T_i) := \begin{cases} \frac{1}{\Pr_{T_i}[f(S_Q,T)=+1]} - 1 & \text{if } f(S_Q,T) = +1 \\ -1 & \text{otherwise} \end{cases}$$

Observe that Ψ_i has been normalized so that for every setting of S_Q^i, $\mathbb{E}_{T_i}\left[\Psi_i(S_Q^i, T_i)\right] = 0$.

We need two technical components regarding Ψ_i. First, we cannot have $\min_{S_Q^i, T_i}\left|\Psi_i(S_Q^i, T_i)\right|$ be too small, which corresponds to having a lower bound of β (we will set the bound on β later in the section) on

$$\min\left\{\frac{\Pr_{T_i}[f(S_Q,T) = -1]}{\Pr_{T_i}[f(S_Q,T) = +1]}, \frac{\Pr_{T_i}[f(S_Q,T) = +1]}{\Pr_{T_i}[f(S_Q,T) = -1]}\right\} \geq \beta$$

Otherwise, we cannot use the simulation theorem from [LY25]. This is equivalent to having the function somewhat balanced, i.e.,

$$\left|\mathbb{E}_{T_i}\left[f_i(S_Q^i, T_i)\right]\right| \leq \frac{1-\beta}{1+\beta}$$

As the next component, we would like Ψ_i to be resilient against a short message. If a short message dependent only on T_i is sent, we would like the corresponding distribution on Ψ_i to have small discrepancy. This results in the following definition.

Definition 1.1 (One-way Discrepancy)**.** *Fix S_Q^{-i}, T_{-i}. We define one-way discrepancy of Ψ_i under the distribution μ over T_i as*

$$\mathsf{odisc}_{\mu}(\Psi_i) := \mathbb{E}_{S_Q^i} \left| \sum_{T_i} \Psi_i(S_Q^i, T_i) \cdot \mu(T_i) \right|$$

with the parameter $C > 0$ as follows

$$\mathsf{odisc}_{C}(\Psi_i) := \max_{M(T_i):|M|<C} \mathbb{E}_M \left[\mathsf{odisc}_{T_i|M}(\Psi_i) \right]$$

An intuitive interpretation of one-way discrepancy is to use the following communication game. Suppose Alice is given S_Q^i and Bob is given T_i, distributed independently (i.e. a product distribution). Bob generates a message M of length at most C (note that the message is independent of S_Q^i). Alice then generates a guess $\{\pm 1\}$ that maximizes the advantage in [LY25] for the balanced version of f. [2] Since Bob's message is independent of Alice's input, we can simply take the expectation over S_Q^i to measure the expected advantage. The intuition of a "hard" function is f which is resilient against a short message by Bob.

The above model is not new. Such a communication model was originally studied in [KNR95], that is, the lower bound for one-way communication under a product distribution but in the small constant-error regime.[3] Here, we want to consider the low-advantage (high-error) regime.

The aforementioned two conditions can be summarized in the following definition of a hard f:

Definition 1.2. *We say f is "hard" if for each $i \in [\ell/3, \ell]$, with probability $\geq g_i \geq \Omega(1)$ over S_Q^{-i} and T_{-i}, f_i and its corresponding Ψ_i satisfy both*

- *(Balanced)* $\Pr_{S_Q^i} \left[\left| \mathbb{E}_{T_i} \left[f_i(S_Q^i, T_i) \right] \right| \geq \frac{1-\beta}{1+\beta} \right] \leq n^{-2}$ *with* $\beta \geq 2^{-o(\sqrt{\log n})}$
- *(Low Discrepancy)* $\mathsf{odisc}_C(\Psi_i) \leq n^{-2}$ *for* $C \leq n_i/\mathrm{poly}\log(n)$

Intuitively, this corresponds to saying that f is "hard" if for any given epoch $i \in [\ell/3, \ell]$, the corresponding f_i is balanced and resistant against a short one-way message with good probability over the remaining coordinates S_Q^{-i} and T_{-i}.

Then for such a hard f, we prove the following super-logarithmic lower bound on $\max\{t_u, t_q\}$.

Theorem 1.3 (Informal)**.** *For the explicit dynamic data structure Problem 1 equipped with a hard f, the update time t_u and query time t_q with $m = \Omega(n^{0.99})$ must have*

$$t_q \geq \Omega\left(\frac{\log^{3/2} n}{\log^2(wt_u)} \right).$$

Therefore, our work essentially reduces proving a super-logarithmic dynamic cell-probe lower bound to proving a lower bound against one-way communication over a product distribution (i.e., showing that the function f is "hard").

[2] If the function is a balanced function, Ψ_i is simply a ± 1 matrix (or the usual communication matrix for f_i).

[3] [KNR95] made a connection between the VC-dimension d of f_i with fixed S_Q^i in the small constant-error regime. However, their bound does not apply in the low-advantage regime.

1.1.1 Comparison with Previous Results

Progress on dynamic cell-probe lower bounds has been slow but steady over the past three decades. The seminal work of Fredman and Saks [FS89] provided the first $\Omega(\log n/\log\log n)$ dynamic cell-probe lower bound. It took another 15 years to remove the $\log\log n$ factor from the denominator. Pătraşcu and Demaine [PD04, PD06] gave an $\Omega(\log n)$ bound for an explicit dynamic problem. Yet achieving a super-logarithmic lower bound (i.e., $\omega(\log n)$ time lower bound on update/query time) seemed elusive [Tho13].

Larsen [Lar12] gave an $\Omega((\log n/\log\log n)^2)$ bound for the 2-dimensional range sum problem, breaking the logarithmic barrier for dynamic data structure lower bounds. Larsen's result, however, has a caveat. The number of output bits per query is $\Theta(\log n)$. Therefore, $\max\{t_u, t_q\}$ *per output bit* is still $\tilde{\Omega}(\log n)$. It took roughly an extra decade to finally provide a super-logarithmic lower bound for a Boolean dynamic data structure problem, established in [LWY20], which gave an $\tilde{\Omega}(\log^{3/2} n)$ lower bound for the dynamic 2-dimensional range parity sum problem. [LY25] then further extended the technique to dynamic graph s-t reachability (in a directed acyclic graph).

Multiphase Program On the other hand, the most promising avenue of attack in achieving a super-logarithmic lower bound (or even a polynomial bound) on $\max\{t_u, t_q\}$ for the past decades has been the Multiphase Program [Pat10, Tho13], the holy grail in dynamic cell-probe lower bounds.

Pătraşcu's original approach for the Multiphase Program was to provide a lower bound for the following communication game.

- Alice is given all the inputs but $T \in \{0,1\}^n$. Bob is given all the inputs but $\vec{S}$ consisting of $S_1, \ldots, S_m \in \{0,1\}^n$. Merlin is given all the inputs but $Q \in [m]$.
- Merlin sends a message of length $t_u \cdot n \cdot w$ to Bob. Alice and Bob then proceed in a standard two-party communication protocol to output $f(S_Q, T)$ after communicating $t_q w$ bits.

Problem 2: Multiphase Communication Game

There are results directly attacking the above communication game (namely, [CEEP12, KW20, DL20, Ko25]). But we emphasize that attacking the communication game is a much harder problem than the underlying dynamic data structure problem [KW20, Ko25]. For instance, the Multiphase Communication Game has connections to a very old circuit lower bound problem [JS10].

Furthermore, there is a known separation between the communication model and the dynamic data structure problem in question [Ko25]. There exists f (namely indexing) that is easy in the communication model, but hard under the dynamic cell-probe model.

But even if we turn to a weaker dynamic cell-probe lower bound, only a logarithmic lower bound is known. When f is Disjointness or Inner Product (mod 2), [BL15, CGL15] showed $\max\{t_u, t_q\} \geq \Omega(\log n)$.[4]

In summary, even though the Multiphase Problem is conjectured to be harder than problems considered in [LWY20, LY25], no super-logarithmic lower bounds were known for the Multiphase Problem prior to our work, due to a technical challenge that we explain in Section 1.2.

[4] From a technical perspective, this is under a weaker model, where the query algorithm has no knowledge of $\vec{S}$ and therefore must probe for it as well. The setting we consider is slightly stronger as the querier already knows $\vec{S}$.

1.2 Technical Contribution

1.2.1 One way simulation theorem of [LWY20]

One technical ingredient of our work is the one-way simulation theorem from [LWY20, LY25] which pioneered "a recipe" for super-logarithmic dynamic cell-probe lower bounds. In this section, we highlight the contributions in [LWY20, LY25] to provide context for our technical contribution.

Consider a fixed dynamic data structure problem $f : \mathcal{Q} \times T \to \{\pm 1\}$, where $\mathcal{Q}$ is the set of queries, and T is the sequence of updates. Therefore if T is given as the sequence of updates, when $Q \in \mathcal{Q}$ is given as the query, the data structure must output $f(Q, T)$.

The first building block for the one-way simulation theorem is the seminal chronogram technique of Fredman and Saks [FS89]. A simple way to interpret chronogram techniques is as follows. As the dynamic data structure must be ready to handle queries at any point during the sequence of updates, in contrast to the static data structure, we can divide the update T into epochs of length $n_1, \ldots, n_\ell$ such that $\sum_{i=1}^{\ell} n_i = n/k$, where each $n_i = \gamma n_{i-1}$ for some parameter γ. Therefore, they are geometrically decreasing. The main observations in the chronogram technique are as follows: (i) The average number of cells that are probed which were last updated in epoch i is t_{tot}/ℓ; (ii) The number of cells changed by all subsequent epochs is small, as the number of updates decreases geometrically over the epochs.

Now we focus on a fixed epoch i to give a lower bound on t_{tot}/ℓ. Such a static data structure (for a fixed epoch i) has pre-initialized memory (updates performed in previous epochs) and a cache (updates performed in subsequent epochs). Observe that a data structure in epoch i, with n_i updates, which would generate $t_u n_i$ many updated cells, has $\sum_{j>i} t_u n_j \leq o(n_i)$ cells in cache, and $\sum_{j<i} t_u n_j$ cells in pre-initialized memory.

To prove a lower bound for such a static data structure, [LWY20, LY25] introduce a one-way communication game which can simulate a static data structure with pre-initialized memory and a cache, to then argue that no "too-good-to-be-true" one-way protocol can exist.

- Bob is given all the updates $\{T_i\}_{i=1}^{\ell}$, but not the desired query $Q \in \mathcal{Q}$.
- Alice is given all the updates, except the updates in epoch i (i.e., T_i), and the desired query.
- Bob sends a one-way message M of length C-bits to Alice. Then Alice announces $f(Q, T)$.

Protocol 1: One-way Communication Simulation G_f

Alice, without Bob's message, can only generate the pre-initialized memory on her own. In the extreme case where $C \geq w(t_u n_i + o(n_i))$, Bob can then send over the entire cache (which is at most $o(n_i)$ cells) and the memory state generated in epoch i. Therefore, Alice will announce the correct $f(Q, T)$ every time. In fact, just sending T_i would allow Alice to recover $f(Q, T)$ for all $Q \in \mathcal{Q}$. The goal is then to analyze the correctness of the protocol when $C \ll n_i$.

As a measure of correctness of the one-way communication protocol, [LY25] introduces the following quantity called the advantage of the protocol over the product distribution on the query set and update set $\mathcal{D} = \mathcal{D}_{\mathcal{Q}} \times \mathcal{D}_T$.

$$\overrightarrow{\mathsf{adv}}(G_f, \mathcal{D}_{\mathcal{Q}}, \mathcal{D}_T, C) := \max_M \mathbb{E}_{Q,M} \left[\left| \mathbb{E}_{T \sim \mathcal{D}_T | M} \left[\overline{f}(Q, T) | M \right] \right| \right]$$

where $\overline{f} : \mathcal{Q} \times T \to [-1, +1]$ is the normalized version of f so that for every $Q \in \mathcal{Q}$, $\mathbb{E}_{T \sim \mathcal{D}_T}[\overline{f}(Q, T)] = 0$. Note that the normalization is only necessary when f is not a balanced function for every Q.

If f is completely balanced, i.e., the setting for [LWY20], no normalization is necessary, and the above quantity exactly captures the small advantage over a fair random coin toss.

The remaining high-level argument of both [LWY20, LY25] is as follows: (i) show that a "too-good-to-be-true" data structure implies a "too-good-to-be-true" one-way communication protocol with a short one-way message (small C) and good advantage (i.e., the simulation theorem); (ii) show that such a "too-good-to-be-true" one-way communication protocol cannot exist.

As the main contribution in [LWY20, LY25] is the one-way simulation theorem (i.e., step (i)), let us focus on obtaining (i). To obtain (i), we would like to generate a short one-way message M from the static data structure. The key technique to reduce the length of the one-way message M is the cell-sampling technique [Sie04, PTW10]. A naive attempt is to sample each cell updated in epoch i with probability $1/(t_u w)^{\Theta(1)}$ independently at random. Bob then sends the sampled cells to Alice as the single-shot message, along with the cache. If all cells required to be read have been sampled, then Alice can answer correctly. If not, she just randomly guesses the answer. Since there is a non-negligible probability of sampling all required cells, this would result in some small advantage. However, the technical challenge is that Alice cannot distinguish between the above two cases just from the sampled cells alone. Alice cannot distinguish whether a cell she received from Bob's message has been touched during epoch i or not.

The main technical centerpiece to resolve the above issue is the so-called "Peak-to-Average" lemma [LWY20], which states the existence of a small subset of cells such that knowing their contents gives a nontrivial advantage. Using the lemma (which can be interpreted as a clever choice of cells for cell-sampling) and the chronogram technique [FS89], any "too-good-to-be-true" dynamic data structure yields a static data structure (with pre-initialized memory and a cache) that achieves a small advantage over random guessing.

The key insights in [Lar12, LWY20, LY25] are to carefully subsample cells updated in epoch i such that knowing the contents of such cells gives a non-trivial advantage over random guessing. The cache and sub-sampled cells constitute a short one-way message M that achieves a non-trivial advantage over random guessing.

All these technical components formally culminate in the following one-way simulation theorem for a static data structure with pre-initialized memory and a cache.

Theorem 1.4 ([LWY20, LY25])**.** *Let $f : \mathcal{Q} \times T \to [-1,+1]$ be a data structure problem with weights $[-1,+1]$. $\mathcal{D}_\mathcal{Q} \times \mathcal{D}_T$ be a (product) distribution over the queries and inputs such that $\mathbb{E}_{T\sim\mathcal{D}_T}[f(Q,T)] = 0$ for any $Q \in \mathcal{Q}$, and $|f(Q,T)| \geq \beta$ for some $\beta > 0$. If there exists a data structure with pre-initialized memory and a cache for f with at most S updated cells, S_{cac} cells in cache, expected query time t_{tot}, and expected query time into updated cells t_q, then for any $p \in (0,1)$,*

$$\begin{aligned}&\overrightarrow{\mathsf{adv}}(G_\mathcal{P}, \mathcal{D}_\mathcal{Q}, \mathcal{D}_\mathcal{I}, (8pS + S_{cac})w)\\ &\geq \exp\left(-O\left(\log(1/p)(t_q + \sqrt{t_{tot}(t_q \log(1/p) + \log(1/\beta))} + \log(1/\beta)\right)\right) - \exp\left(-\Omega(pS)\right)\end{aligned}$$

1.2.2 Our technical centerpiece

Our main point of departure from [LWY20, LY25] is (ii): showing that no such one-way communication protocol can exist.

The one-way communication game G_f used in [LWY20, LY25] is based on an *explicit* static data structure lower bound on the Butterfly Graph [Pă11]. While this was comparatively straightforward for [LWY20], the primary technical challenge in [LY25] lies in proving a new static lower bound on the Butterfly Graph, as this requires establishing a strong *explicit* static data structure lower

bound even for a small advantage over random guessing. This approach follows from a standard encoding-decoding (and therefore a counting) argument, inspired by the static data structure lower bound from [Pă11].

A key distinction of our method is how it overcomes the limitations of the encoding-decoding (or counting) arguments used in prior work [LWY20, LY25] to establish one-way communication lower bounds. This approach is effective when the query distribution, $\mathcal{M}$, is uniform over a small support, as a simple counting argument suffices (see, e.g., Section 5.2 of [LY25]). However, the encoding-decoding strategy fails when the query set is sub-sampled from a distribution with a much larger support. In this scenario, the one-way message can depend arbitrarily on the specific queries chosen, adding a layer of complexity that makes a simple counting argument intractable. This challenge was the primary bottleneck to applying the simulation lemma from [LWY20] to the Multiphase Problem and necessitated our different approach.

More specifically, we need a lower bound for the following modified one-way communication game where both Alice and Bob have access to $\vec{S}$.

- Bob is given $\vec{S} = S_1, \ldots, S_m$ and all the updates $\{T_j\}_{j=1}^{\ell}$, but not the desired query $Q \in \mathcal{Q} = [m]$.
- Alice is given $\vec{S} = S_1, \ldots, S_m$, all updates except those in epoch i (i.e., T_i), and the desired query Q.
- Bob sends a one-way message M of length C bits to Alice. Then Alice announces $f(S_Q, T)$.

Protocol 2: Modified One-way Communication Game

The main technical challenge for the encoding-decoding argument is the dependence on $\vec{S}$. Alice and Bob both know $\vec{S}$, and M can be interpreted differently depending on the specific $\vec{S}$ that is chosen. Thus, we must account for all possible $\vec{S}$, which essentially invalidates the counting argument. The issue is compounded when proving the lower bound in the low-advantage regime, i.e., where the advantage is $n^{-\Omega(1)}$.

Our technical contribution is the use of information-theoretic tools from [KW20, Ko25] to essentially "remove" the dependency on $\vec{S}$ when S_i and T's are generated under a product distribution. Our work reduces the problem to a lower bound for the following one-way communication game, which has been studied in various contexts and is far more tractable.

- Bob is given all the updates $\{T_j\}_{j=1}^{\ell}$.
- Alice is given ς, chosen from some distribution $\mathcal{D}_{\mathcal{Q}}$, and all updates except those in epoch i (i.e., T_i).
- Bob sends a one-way message M of length $C = n_i/\operatorname{poly}\log(n)$ bits to Alice. Then Alice announces $f(\varsigma, T)$.

Protocol 3: Reduced One-way Communication Game

We remark that our definition of "hard" (Definition 1.2) exactly captures the lower bound for Protocol 3. Hard functions do not have a short protocol with good advantage for Protocol 3. In Section 4 and Section 5, we list some examples of these hard functions including the usual Inner Product mod 2. We remark that we need a lower bound of the form $C \geq n/\operatorname{poly}\log(1/\varepsilon)$ for an

ε-advantage (under a product distribution). We cannot expect such a lower bound for well-studied Indexing or Disjointness, as a simple upper bound with $n/\text{poly}(1/\varepsilon)$ exists. We introduce a new method of lower bounding Protocol 3 against small advantage under a product distribution using average min-entropy.

2 Preliminary

2.1 Information Theory

In this section, we provide the necessary background on information theory and information complexity that are used in this paper. For further reference, we refer the reader to [CT06].

Definition 2.1 (Entropy)**.** *The entropy of a random variable X is defined as*

$$H(X) := \sum_x \Pr[X = x] \log \frac{1}{\Pr[X = x]}.$$

Similarly, the conditional entropy is defined as

$$H(X|Y) := \mathbb{E}_Y \left[\sum_x \Pr[X = x|Y = y] \log \frac{1}{\Pr[X = x|Y = y]} \right].$$

Fact 2.2 (Conditioning Decreases Entropy)**.** *For any random variable X and Y*

$$H(X) \geq H(X|Y)$$

With entropy defined, we can also quantify the correlation between two random variables, or how much information one random variable conveys about the other.

Definition 2.3 (Mutual Information)**.** *Mutual information between X and Y (conditioned on Z) is defined as*

$$I(X; Y|Z) := H(X|Z) - H(X|YZ).$$

Similarly, we can also define how much one distribution conveys information about the other distribution.

Definition 2.4 (KL-Divergence)**.** *KL-Divergence between two distributions μ and ν is defined as*

$$\mathsf{D}_{KL}(\mu||\nu) := \sum_x \mu(x) \log \frac{\mu(x)}{\nu(x)}.$$

To bound mutual information, it suffices to bound KL-divergence, due to the following fact.

Fact 2.5 (KL-Divergence and Mutual Information)**.** *The following equality between mutual information and KL-Divergence holds*

$$I(A; B|C) = \mathbb{E}_{B,C} \left[\mathsf{D}_{KL}(A|_{B=b,C=c}||A|_{C=c}) \right].$$

Fact 2.6 (Pinsker's Inequality)**.** *For any two distributions P and Q,*

$$\|P - Q\|_{TV} = \frac{1}{2}\|P - Q\|_1 \leq \sqrt{\frac{1}{2 \log e} D(P||Q)}$$

We also make use of the following facts on Mutual Information throughout the paper.

Fact 2.7 (Chain Rule)**.** *For any random variable A, B, C and D*

$$I(AD; B|C) = I(D; B|C) + I(A; B|CD).$$

Fact 2.8. *For any random variable A, B, C and D, if $I(B; D|C) = 0$*

$$I(A; B|C) \leq I(A; B|CD).$$

Proof. By the chain rule and non-negativity of mutual information,

$$I(A; B|C) \leq I(AD; B|C) = I(B; D|C) + I(A; B|CD) = I(A; B|CD).$$

□

Fact 2.9. *For any random variable A, B, C and D, if $I(B; D|AC) = 0$*

$$I(A; B|C) \geq I(A; B|CD).$$

Proof. By the chain rule and non-negativity of mutual information,

$$I(A; B|CD) \leq I(AD; B|C) = I(A; B|C) + I(B; D|AC) = I(A; B|C).$$

□

3 Proof of Main Theorem

Recall that the class of explicit dynamic problems that we would like to give a lower bound for is the following.

- $S_1, \ldots, S_m \in \{\{0,1\}^k\}^{n/k}$ are given, each S_Q divided into n/k blocks $S_Q[1], \ldots, S_Q[n/k]$, each of size k.
- $T = (T[1], \ldots, T[n/k]) \in \{\{0,1\}^k\}^{n/k}$ is given as a sequence of updates, using t_u time per update.
- For any given $Q \in \mathcal{Q} = [m]$, output

  $$f(S_Q, T) = \psi\left(g(S_Q[1], T[1]), \ldots, g(S_Q[n/k], T[n/k])\right)$$

  in t_q time.

which we denote as the generalized Multiphase Problem. The original Multiphase Problem can be easily represented as $k = 1$ with g being the standard bit-wise AND, ψ being $\bigvee$, or $\bigoplus$ if f is computing inner product over $\mathbb{F}_2$.

We formally prove the following result for the generalized Multiphase Problem.

Theorem 3.1. *Suppose f is hard (as in Definition 1.2) and $m = \Omega(n^{0.99})$. Then it must be the case that*

$$t_q \geq \Omega\left(\frac{\log^{3/2} n}{\log^2(wt_u)}\right).$$

The main intuition behind phrasing our result in a generalized manner is that our result "lifts" some hard function ψ (in a restricted model) to a super-logarithmic data structure lower bound.

Epochs To give a super-logarithmic lower bound, the first main technical component is the chronogram technique [FS89], where we divide the sequence of n/k updates into epochs, each containing $n_i := \gamma^i$ blocks (of k bits each), and $\sum_{i=1}^{\ell} n_i = n/k$. We denote the epochs of updates as $\{T_i\}_{i=1}^{\ell}$, and we process the updates in reverse order.

Then we also consider the associated coordinates in S_Q. In particular, we denote S_Q^i as the part of S_Q associated with the i-th epoch T_i. This notation can be further extended to work with multiple indices in $[m]$ and $[\ell]$. S_Q^{-i} denotes S_Q except S_Q^i, analogously T_{-i}. We use $\vec{S}$ to denote the bundle over all Q's. Similarly, if we use a subset of $[m]$ in the subscript, that denotes S_Q's in the subscript.

Notations and Hard Distribution We will use τ_{tot} as the random variable for the number of cells (total, across all epochs) probed in the query algorithm. τ_q is used as the random variable for the number of cells updated in epoch i that are probed.

We assume each $S_Q[i]$ is distributed i.i.d. for all $i \in [n/k]$ and $Q \in [m]$, as well as all the coordinates in T. This implies that $S_Q[i]$'s and T_i's are all independent.

We will use the overline notation to denote the normalized version of f. That is, with fixed S_Q^{-i}, T_{-i} assuming without loss of generality that $\Pr_{S_Q^i,T_i}[f(S_Q,T)=+1] \geq \Pr_{S_Q^i,T_i}[f(S_Q,T)=-1]$,

$$\overline{f}_i(S_Q^i, T_i) := \begin{cases} \frac{1}{\Pr_{T_i}[f(S_Q,T)=+1]} - 1 & \text{if } f(S_Q,T)=+1 \\ -1 & \text{otherwise} \end{cases}$$

and vice-versa if $\Pr_{S_Q^i,T_i}[f(S_Q,T)=+1] \leq \Pr_{S_Q^i,T_i}[f(S_Q,T)=-1]$.

High Level Proof Strategy Before delving into the technical part, we would like to give an overview of our overall proof strategy.

1. First, we want to select a "good" epoch $i \in [\ell/3, 5\ell/6]$. We show that there exists some epoch i of the updates, and an associated event E_i and query set $\mathcal{Q}_i' \subset \mathcal{Q}$ which yield a "good" data structure problem with pre-initialized memory and a cache.

2. Such a "good" data structure then yields a "too-good-to-be-true" protocol for the following communication model from [LWY20, LY25]. Bob sends a one-way message M to Alice. Then

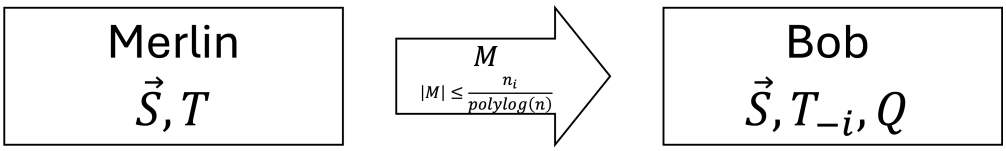

Figure 1: Communication Model

   Alice guesses the value of $f(S_Q,T) = f_i(S_Q^i, T_i)$. The performance of the protocol is then measured by the advantage of the protocol.

3. We show that a "too-good-to-be-true" dynamic data structure implies a "too-good-to-be-true" one-way simulation in Section 3.1. Then in Section 3.2, we show that hard f's cannot have a "too-good-to-be-true" one-way simulation using information-theoretic tools, thereby resulting in a contradiction.

3.1 One-way simulation Theorem

In this section, we will prove the following simulation theorem (an analog of Theorem 1.4 from [LWY20, LY25]) which converts a "too-good-to-be-true" dynamic data structure for f to a one-way communication protocol with a high advantage. Recall that our f in question satisfies the following property.

Definition 1.2. *We say f is "hard" if for each $i \in [\ell/3, \ell]$, with probability $\geq g_i \geq \Omega(1)$ over S_Q^{-i} and T_{-i}, f_i and its corresponding Ψ_i satisfy both*

- *(Balanced)* $\Pr_{S_Q^i}\left[\left|\mathbb{E}_{T_i}\left[f_i(S_Q^i, T_i)\right]\right| \geq \frac{1-\beta}{1+\beta}\right] \leq n^{-2}$ *with* $\log(1/\beta) \leq o(\log^{1/2}(n))$.
- *(Low Discrepancy)* $\mathsf{odisc}_C(\Psi_i) \leq n^{-2}$ *with* $C \leq n_i/\mathrm{poly}\log(n)$.

For such f, we prove the following simulation theorem.

Theorem 3.2. *Suppose for any $\vec{S}$, there exists a dynamic data structure for Problem 1 with f with update time $t_u = \mathrm{poly}\log(n)$, average query time $\mathbb{E}_{Q,T}[\tau_{tot}] = t_{tot} \leq o\left(\frac{\log^{3/2}(n)}{(\log\log n)^2}\right)$. Then there exists an epoch $i \in [\ell/3, 5\ell/6]$, an event E_i^1 and a query set $\mathcal{Q}_i \subset \mathcal{Q}$ that depend only on $\vec{S}^{-i}, T_{-i}$ with $\Pr_{\vec{S}^{-i},T_{-i}}[E_i^1] \geq \Omega(g_i)$ such that*

$$|\mathcal{Q}_i| \geq g_i m/2$$
$$\forall Q \in \mathcal{Q}_i, \quad \Pr_{S_Q^i}\left[|\mathbb{E}_{T_i}[f(S_Q, T)]| \geq \frac{1-\beta}{1+\beta}\right] \leq n^{-2}, \ \mathsf{odisc}_C(\Psi_i) \leq n^{-2}.$$

Furthermore, there exists a sub-event $E_i^2 \subset E_i^1$ with $\Pr[E_i^2|E_i^1] \geq 9/10$, and a query set $\mathcal{Q}_i' \subset \mathcal{Q}_i$ with $|\mathcal{Q}_i'| \geq \left(1 - \frac{10}{n^2}\right)|\mathcal{Q}_i|$ depending only on $\vec{S}, T_{-i}$, conditioned on which there exists a one-way communication protocol that obtains

$$\overrightarrow{\mathsf{adv}}(G_{\overline{f}_i}, \mathcal{D}_{\mathcal{Q}_i'}, \mathcal{D}_{T_i}, n_i/\mathrm{poly}\log(n)) \geq n^{-o(1)}.$$

We will use Theorem 1.4 as our main ingredient, which we reiterate below.

Theorem 1.4 ([LY25])**.** *Let $\mathcal{P} : \mathcal{Q} \times \mathcal{T} \to [-1, +1]$ be a data structure problem with weights, and $\mathcal{D}_\mathcal{Q} \times \mathcal{D}_T$ be a distribution over the queries and inputs such that $\mathbb{E}_T[\mathcal{P}(Q, T)] = 0$ for all $Q \in \mathcal{Q}$ and $|\mathcal{P}(Q, T)| \geq \beta$ for some $\beta > 0$. If there is a data structure D with pre-initialized memory and a cache for $\mathcal{P}$ such that D has at most S updated cells, S_{cac} cells in cache, expected query time t_{tot}, and expected query time into updated cells t_q, then for any $p \in (0, 1)$, there exists a one-way communication protocol $G_{\mathcal{P}_i}$ such that*

$$\overrightarrow{\mathsf{adv}}(G_\mathcal{P}, \mathcal{D}_\mathcal{Q}, \mathcal{D}_T, (8pS + S_{cac})w)$$
$$\geq \exp\left(-O\left(\log(1/p)(t_q + \sqrt{t_{tot}(t_q \log(1/p) + \log(1/\beta))} + \log(1/\beta)\right)\right) - \exp\left(-\Omega(pS)\right)$$

Remark 3.3. *Note that the β in the simulation is the same β in the definition of hard f. Recall that we normalize $f_i(S_Q^i, T_i)$ so that the expectation is zero. We leave it to the reader to verify that the balanced condition $\left|\mathbb{E}_{T_i}\left[f_i(S_Q^i, T_i)\right]\right|$, is exactly the condition required to lower bound the normalized absolute value.*

Proof of Theorem 3.2. Before directly applying Theorem 1.4, we need to filter through a sequence of events and query sets towards the final lower bound.

First, we prove the following claim, which gives a fixed epoch $i \in [\ell/3, 5\ell/6]$, an event E_i^1 and query set $\mathcal{Q}_i$ from $\vec{S}^{-i}, T_{-i}$ that are "good."

Claim 3.4. *There exists an epoch $i \in [\ell/3, 5\ell/6]$, an event E_i^1, and a query set $\mathcal{Q}_i \subset \mathcal{Q}$ depending only on $\vec{S}^{-i}$ and T_{-i} such that:*

$$\Pr_{\vec{S}^{-i}, T_{-i}}[E_i^1] \geq \Omega(g_i), \quad |\mathcal{Q}_i| \geq \frac{g_i m}{2}$$

$$\forall Q \in \mathcal{Q}_i,\ \mathsf{odisc}_C(\Psi_i) \leq n^{-2}, \quad \Pr_{S_Q^i}\left[|\mathbb{E}_{T_i}[f(S_Q, T)]| \geq \frac{1-\beta}{1+\beta}\right] \leq n^{-2}$$

$$\mathbb{E}_{Q, T_i | E_i^1}[\tau_Q] \leq \frac{8 \cdot t_{tot}}{g_i^2 \cdot \ell}, \quad \mathbb{E}_{Q, T_i | E_i^1}[\tau_{tot}] \leq \frac{4 \cdot t_{tot}}{g_i^2}$$

Proof. First, a simple Markov argument shows that there must exist an epoch $i \in [\ell/3, 5\ell/6]$ such that

$$\mathbb{E}_{Q, \vec{S}, T}[\tau_q] \leq 2 \cdot t_{tot}/\ell. \tag{1}$$

Furthermore, due to our assumption on the underlying function f, we have that for each $Q \in \mathcal{Q}$,

$$\Pr_{S_Q^{-i}, T_{-i}}\left[\left(\Pr_{S_Q^i}\left[|\mathbb{E}_{T_i}[f(S_Q, T)]| \geq \frac{1-\beta}{1+\beta}\right] \geq n^{-2}\right) \vee \left(\mathsf{odisc}_C(\Psi_i) \geq n^{-2}\right)\right] \leq 1 - g_i. \tag{2}$$

By Markov's inequality, with probability at least $1 - \frac{1-g_i}{1-(g_i/2)} = \frac{g_i}{2-g_i} \geq \frac{g_i}{2}$ over $\vec{S}^{-i}$ and T_{-i}, there exists $\mathcal{Q}_i \subset \mathcal{Q}$ with $|\mathcal{Q}_i| \geq \frac{g_i m}{2}$ such that for all $Q \in \mathcal{Q}_i$

$$\Pr_{S_Q^i}\left[|\mathbb{E}_{T_i}[f(S_Q, T)]| \geq \frac{1-\beta}{1+\beta}\right] \leq n^{-2}$$

$$\mathsf{odisc}_C(\Psi_i) \leq n^{-2}$$

Denote the event corresponding to this choice of $\vec{S}^{-i}, T_{-i}$ as E_i^1. Then as $\Pr[E_i^1] \geq \frac{g_i}{2}$ and $|\mathcal{Q}_i| \geq \frac{g_i m}{2}$, this implies that

$$\mathbb{E}_{Q \in \mathcal{Q}_i, \vec{S^i}, T_i | E_i^1}[\tau_q] \leq \left(\frac{2}{g_i}\right)^2 \cdot 2 \cdot t_{tot}/\ell$$

$$\mathbb{E}_{Q \in \mathcal{Q}_i, \vec{S^i}, T_i | E_i^1}[\tau_{tot}] \leq \left(\frac{2}{g_i}\right)^2 t_{tot}.$$

which completes the proof of the claim. □

Next, we proceed to the second filter, where we condition on $\vec{S^i}$ as well. Conditioned on E_i^1, we would like to have a sub-event $E_i^2 \subset E_i^1$, and query subset $\mathcal{Q}_i'$ further conditioned on $\vec{S^i}$.

Claim 3.5. *There exists an event E_i^2, query set $\mathcal{Q}_i' \subset \mathcal{Q}_i$ with $|\mathcal{Q}_i'| \geq \left(1 - \frac{10}{n^2}\right)|\mathcal{Q}_i|$ (depending only on $\vec{S}$ and T_{-i}) such that*

$$\Pr_{\vec{S}, T_{-i}}[E_i^2 | E_i^1] \geq 9/10$$

$$\forall Q \in \mathcal{Q}_i',\ \mathsf{odisc}_C(\Psi_i) \leq n^{-2}, \quad |\mathbb{E}_{T_i}[f(S_Q, T)]| \leq \frac{1-\beta}{1+\beta}$$

$$\mathbb{E}_{Q, T_i | E_i^2}[\tau_q] \leq \frac{10 \cdot t_{tot}}{g_i^2 \cdot \ell}, \quad \mathbb{E}_{Q, T_i | E_i^2}[\tau_{tot}] \leq \frac{5 \cdot t_{tot}}{g_i^2}$$

Proof. Now, we switch our attention to choosing $\vec{S^i}$ and the respective $\mathcal{Q}'_i \subset \mathcal{Q}_i$ assuming E_i^1. Then there exists an event E_i^2 depending on the additional $\vec{S^i}$ such that

$$\Pr_{Q\sim\mathcal{Q}_i}\left[|\mathbb{E}_{T_i}\left[f(S_Q,T)\right]| \le \frac{1-\beta}{1+\beta}\right] \ge \left(1-\frac{10}{n^2}\right),$$

that is, a choice of $\vec{S^i}$ such that at most $\frac{10}{n^2}$ of the resulting $f(S_Q,T)$'s are **not** well-balanced. Then by Markov's inequality,

$$1 - \Pr_{\vec{S^i},\vec{S^{-i}},T_{-i}}[E_i^2|E_i^1] \le \frac{n^{-2}}{\frac{10}{n^2}} = 1/10.$$

Assuming E_i^2, $\vec{S^i}$ yields a set $\mathcal{Q}'_i \subset \mathcal{Q}_i$ such that for every $Q \in \mathcal{Q}'_i$,

$$\mathsf{odisc}_C(\Psi_i) \le n^{-2}, \quad |\mathbb{E}_{T_i}\left[f(S_Q,T)\right]| \le \frac{1-\beta}{1+\beta}$$

Note that the first condition only depends on $\vec{S}^{-i}, T_{-i}$, while the second condition depends on $\vec{S}, T_{-i}$. Then over $Q \in \mathcal{Q}'_i$, conditioned on E_i^2, by a simple Markov argument,

$$\mathbb{E}_{Q\in\mathcal{Q}'_i,T_i|E_i^2}\left[\tau_q\right] \le \frac{10}{9}\cdot\frac{1}{1-\frac{10}{n^2}}\left(\frac{2}{g_i}\right)^2\cdot 2\cdot t_{tot}/\ell \le \frac{10\cdot t_{tot}}{g_i^2\ell}$$
$$\mathbb{E}_{Q\in\mathcal{Q}'_i,T_i|E_i^2}\left[\tau_{tot}\right] \le \frac{10}{9}\cdot\frac{1}{1-\frac{10}{n^2}}\cdot\frac{4\cdot t_{tot}}{g_i^2} \le \frac{5\cdot t_{tot}}{g_i^2}$$

□

Now we have the required definition of event E_i^2 and $\mathcal{Q}'_i$. We proceed to apply Theorem 1.4 conditioned on E_i^2 and over the query set $\mathcal{Q}'_i$. Given Claim 3.5, we zoom into the event E_i^2. Given E_i^2, which is determined by $\vec{S}, T_{-i}$, there exists a data structure with updated cells $S \le t_u n_i$ and cache $S_{cac} \le \sum_{i''<i} t_u n_{i''} = t_u\frac{n_i-1}{\gamma-1} \le \frac{2t_u n_i}{\gamma} \le \frac{n_i}{(t_u w)^{\Theta(1)}}$, with expected query time $\frac{5\cdot t_{tot}}{g_i^2}$ and expected query time into updated cells $\frac{10\cdot t_{tot}}{g_i^2\ell}$ over some query set $\mathcal{Q}'_i \subset \mathcal{Q}_i$, where the bound on S_{cac} holds from setting $\gamma = (t_u w)^{\Theta(1)}$. Theorem 1.4 then implies a one-way communication protocol for $f_i(S_Q^i, T_i)$ for any $p \in (0,1)$ with advantage

$$\overrightarrow{\mathsf{adv}}(G_{\overline{f}_i}, \mathcal{D}_{\mathcal{Q}'_i}, \mathcal{D}_{T_i}, (8pS + S_{cac})w)$$
$$\ge \exp\left(-O\left(\log(1/p)\left(t_q + \sqrt{t_{tot}(t_q\log(1/p) + \log(1/\beta))}\right) + \log(1/\beta)\right)\right) - \exp\left(-\Omega(pS)\right) \quad (3)$$

We set $p := \frac{1}{(t_u w)^{\Theta(1)}}$. Then observe that $\ell = \Omega(\log_\gamma(n)) = \Omega\left(\frac{\log n}{\log\log n}\right)$. Plugging in the bounds, we get

$$\beta^{-O(1)}\cdot 2^{-O\left(\log(t_u w)\left(\frac{t_{tot}}{g_i^2\log_\gamma(n)} + \sqrt{\frac{t_{tot}^2\log(t_u w)}{g_i^4\log_\gamma(n)} + \frac{t_{tot}\log(1/\beta)}{g_i^2}}\right)\right)} - \exp\left(-\frac{n_i}{(t_u w)^{\Theta(1)}}\right). \quad (4)$$

Assuming $t_u = \operatorname{poly}\log(n)$, $w = O(\log n)$, $n_i \ge \Omega(n^{1/3})$ due to the epoch being $i \in [\ell/3, 5\ell/6]$, if we assume $\log(1/\beta) \le o(\log^{1/2}(n))$, and $t_{tot} \le o\left(\frac{g_i^2\cdot\log^{3/2}(n)}{(\log\log(n))^2}\right)$, (4) simplifies to

$$(4) \ge \beta^{-O(1)}\cdot n^{-o(1)} \ge n^{-o(1)}.$$

Therefore, there exists a message of length $\frac{n_i}{\operatorname{poly}\log(n)}$ with $n^{-o(1)} = n_i^{-o(1)}$ advantage for $G_{\overline{f}_i}$ □

3.2 Combinatorial Lower Bound

Next, we would like to show that no such one-way protocol exists. Before we delve into the proof of impossibility, we briefly summarize the properties we attained in Section 3.1.

Recap of Previous Section In the previous section, we showed that a "too-good-to-be-true" dynamic data structure yields a "too-good-to-be-true" one-way communication protocol, which had the following property:

- $\mathcal{Q}_i \subset \mathcal{Q}$ which depends only on $\vec{S}^{-i}$ and T_{-i}, with size $\geq g_i m/2$ such that all $Q \in \mathcal{Q}_i$ have:
$$\Pr_{S^i_Q}\left[\left|\mathbb{E}_{T_i}\left[f(S_Q, T)\right]\right| \geq \frac{1-\beta}{1+\beta}\right] \leq n^{-2}$$
$$\mathsf{odisc}_C(\Psi_i) \leq n^{-2};$$
- An event E_i^2 with probability $\Pr_{\vec{S},T_{-i}}[E_i^2|E_i^1] \geq 9/10$ and a set of good queries $\mathcal{Q}'_i \subset \mathcal{Q}_i$ with $|\mathcal{Q}'_i| \geq \left(1 - \frac{10}{n^2}\right)|\mathcal{Q}_i|$ depending only on $\vec{S}, T_{-i}$ where conditioned on E_i^2
$$\overrightarrow{\mathsf{adv}}(G_{\overline{f}_i}, \mathcal{D}_{\mathcal{Q}'_i}, \mathcal{D}_{T_i}, n_i/\mathrm{poly}\log(n)) \geq n^{-o(1)}.$$

We will first pick "good" $\vec{S}^{-i}, T_{-i}$ satisfying E_i^1 as part of public randomness (i.e., we will assume the event E_i^1). Note that $\vec{S}^{-i}, T_{-i}$ reveals zero information about $\vec{S}^i$ and T_i due to our hard distribution being independent. Thus, conditioning on the event will not affect the distribution over $\vec{S}^i$ and T_i.

Main Lower Bound We would like to show that these properties cannot be satisfied simultaneously. To that end, we prove the following combinatorial black-box theorem, which can be used to derive the contradiction.

Theorem 3.6. *If $|M| \leq C = \frac{n_i}{\mathrm{poly}\log n}$, the maximum advantage Bob can attain, conditioned on an event E_i determined by $\vec{S}, T_{-i}$ is*
$$\overrightarrow{\mathsf{adv}}(G_{\overline{f}_i}, \mathcal{D}_{\mathcal{Q}'_i}, \mathcal{D}_{T_i}, C) \leq n^{-\Omega(1)}$$

Before delving into the main proof, we would like to give a brief overview of the proof, and what combinatorial properties are necessary to derive a contradiction. Our main proof strategy roughly follows the technique pioneered in [KW20] and subsequently [Ko25]. We show that a "too-good-to-be-true" one-way message M achieving a large advantage leads to a "too-good-to-be-true" random process Z which achieves (i) small information on T_i; (ii) small information on S^i_Q; and (iii) little correlation between S^i_Q and T_i, while such Z cannot exist, as the underlying f is hard.

We define our "too-good-to-be-true" random process Z as follows (which would be guaranteed by setting $E_i = 1$). Let $\mathcal{P}$ be a non-overlapping random sequence of numbers in $\mathcal{Q}_i$ of length l, where l is a parameter we will decide in the final part of the proof. Note that $\mathcal{Q}_i$ is completely determined by $\vec{S}^{-i}, T_{-i}$. Therefore, $\mathcal{P}$ reveals zero information about $\vec{S}^i$. We denote the number at position j as ρ_j. Then we also take a random number $J \in [l]$ uniformly at random. Let B_Q denote Bob's response for query $Q \in [m]$, and a Boolean random variable G_Q whether $Q \in \mathcal{Q}'_i$. Our Z is then defined as
$$Z := M, \mathcal{P}_{<J}, G_{\rho_{<J}}, B_{\rho_{<J}}, S^i_{\rho_{<J}}, J$$
Then we attach Z along with $\rho_J, G_{\rho_J}, B_{\rho_J}$ and conditioning on $E_i = 1$ to derive the contradiction.

3.2.1 Technical Lemmas

Towards the proof of Theorem 3.6, we prove the necessary technical claims first. The main technical deviation from previous works is on how to deal with the information on $S^i_{\rho_J}$. Just as in [KW20], Z reveals little information about $S^i_{\rho_J}$ in terms of mutual information or KL divergence.

Claim 3.7 (Low information on $S^i_{\rho_J}$)**.**

$$\mathbb{E}_{Z,\rho_J|E_i=1}\left[D(S^i_{\rho_J}|_{Z,\rho_J,E_i=1}||S^i_{\rho_J})\right] \leq \frac{5(C+l)}{m}$$

Proof. Suppose we condition on the event E^1_i. Then we use the analogous direct sum technique from [KW20]. First observe that

$$\begin{aligned} I(S^i_{\rho_J}; Z, \rho_J, E^2_i|E^1_i = 1) &\leq I(S^i_{\rho_J}; M, \mathcal{P}, B_{\rho_{<J}}, G_{\rho_{<J}}, S^i_{\rho_{<J}}, E^2_i, J|E^1_i = 1) \\ &= I(S^i_{\rho_J}; M, E^2_i, B_{\rho_{<J}}|\mathcal{P}, S^i_{\rho_{<J}}, J, E^1_i = 1) \end{aligned}$$

For any fixed $J = j$, note that

$$\begin{aligned} &I(S^i_{\rho_J}; M, E^2_i, B_{\rho_{<j}}, G_{\rho_{<j}}|\mathcal{P}, S^i_{\rho_{<J}}, J = j, E^1_i = 1) \leq \frac{1}{m-j} I(S^i_{-\rho_{<j}}; M, E^2_i, B_{\rho_{<j}} G_{\rho_{<j}}|S^i_{\rho_{<j}}, E^1_i = 1) \\ &\leq \frac{C+2j+1}{m-j} \leq \frac{C+2l+1}{m-l} \leq \frac{3(C+l)}{m}. \end{aligned}$$

Then note that since $\Pr[E^2_i|E^1_i] \geq 9/10$, and

$$\mathbb{E}_{Z,\rho_J,E^2_i|E^1_i=1}\left[D(S^i_{\rho_J}|_{Z,E^2_i}||S^i_{\rho_J})\right] = I(S^i_{\rho_J}; Z, \rho_J, E^2_i|E^1_i = 1),$$

this completes the proof of the claim, by writing $E_i = E^1_i \wedge E^2_i$. □

We will need to translate the bound on $S^i_{\rho_J}$ from Claim 3.7 using the following claim.

Claim 3.8.

$$\mathbb{E}_{Z,\rho_J|G_{\rho_J}=1,E_i=1}\left[\|S^i_{\rho_J}|_{Z,\rho_J,G_{\rho_J}=1,E_i=1} - S^i_{\rho_J}\|_1\right] \leq 5\sqrt{\frac{(C+l)}{m}} + \frac{200}{n^2}$$

Proof. We will upper bound $\|S^i_{\rho_J}|_{Z,\rho_J,G_{\rho_J}=1,E_i=1} - S^i_{\rho_J}\|_1$ using the triangle inequality, namely

$$\begin{aligned} \|S^i_{\rho_J}|_{Z,\rho_J,G_{\rho_J}=1,E_i=1} - S^i_{\rho_J}\|_1 &\leq \|S^i_{\rho_J}|_{Z,\rho_J,G_{\rho_J}=1,E_i=1} - S^i_{\rho_J}|_{Z,\rho_J,E_i=1}\|_1 \quad (5) \\ &\quad + \|S^i_{\rho_J}|_{Z,\rho_J,E_i=1} - S^i_{\rho_J}\|_1 \quad (6) \end{aligned}$$

First, we bound (5). If we consider any fixed Z, ρ_J,

$$\|S^i_{\rho_J}|_{Z,\rho_J,G_{\rho_J}=1,E_i=1} - S^i_{\rho_J}|_{Z,\rho_J,E_i=1}\|_1 \leq 2 \cdot \Pr[G_{\rho_J} = 0|Z, \rho_J, E_i = 1].$$

Then taking expectation over Z, ρ_J

$$\begin{aligned} &\mathbb{E}_{Z,\rho_J|G_{\rho_J}=1,E_i=1}\left[(5)\right] \leq 2 \cdot \mathbb{E}_{Z,\rho_J|G_{\rho_J}=1,E_i=1}\left[\Pr[G_{\rho_J} = 0|Z, \rho_J, E_i = 1]\right] \\ &\leq 2 \cdot \left(\mathbb{E}_{Z,\rho_J|E_i=1}\left[\Pr[G_{\rho_J} = 0|Z, \rho_J, E_i = 1]\right] + 2\Pr[G_{\rho_J} = 0|E_i = 1]\right) \\ &= 6 \cdot \Pr[G_{\rho_J} = 0|E_i = 1] \leq \frac{60}{n^2}. \end{aligned}$$

If we take expectation over (6),

$$\begin{aligned}\mathbb{E}_{Z,\rho_J|G_{\rho_J}=1,E_i=1}\left[(6)\right] &\leq \mathbb{E}_{Z,\rho_J|E_i=1}\left[(6)\right] + 4\Pr[G_{\rho_J}=0|E_i=1] \\ &\leq \mathbb{E}_{Z,\rho_J|E_i=1}\left[\|S^i_{\rho_J}|_{Z,\rho_J,E_i=1} - S^i_{\rho_J}\|_1\right] + \frac{40}{n^2}.\end{aligned}$$

The bound on $\mathbb{E}_{Z,\rho_J|E_i=1}\left[\|S^i_{\rho_J}|_{Z,\rho_J,E_i=1} - S^i_{\rho_J}\|_1\right]$ can be attained immediately from Claim 3.7 due to Pinsker's inequality (Fact 2.6) and Jensen's inequality.

$$\mathbb{E}_{Z,\rho_J|E_i=1}\left[\|S^i_{\rho_J}|_{Z,\rho_J,E_i=1} - S^i_{\rho_J}\|_1\right] \leq 2\sqrt{\mathbb{E}_{Z,\rho_J|E_i=1}\left[D(S^i_{\rho_J}|_{Z,\rho_J,E_i=1}||S^i_{\rho_J})\right]} \leq 2\sqrt{\frac{5(C+l)}{m}}$$

Combining all the bounds, we obtain

$$\mathbb{E}_{Z,\rho_J|G_{\rho_J}=1,E_i=1}\left[\|S^i_{\rho_J}|_{Z,\rho_J,G_{\rho_J}=1,E_i=1} - S^i_{\rho_J}\|_1\right] \leq 5\cdot\sqrt{\frac{(C+l)}{m}} + \frac{100}{n^2}$$

completing the proof of the claim. □

We need the following claim as well to translate the advantage of the protocol to one-way discrepancy of the underlying function,

Claim 3.9. *Let A and B be some random variable which satisfies*

$$\begin{aligned}&\|S^i_{\rho_J}|_{A=a} - S^i_{\rho_J}\|_1 \leq \delta_1 \\ &\mathbb{E}_{S^i_{\rho_J}}\left[\left|\mathbb{E}_{T_i|_{A=a,B=b}}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right] \leq \delta_2\end{aligned}$$

then

$$\mathbb{E}_{S^i_{\rho_J}|_{A=a,B=b}}\left[\left|\mathbb{E}_{T_i|_{A=a,B=b}}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right] \leq \frac{\delta_2+\delta_1}{\Pr[B=b|A=a]}$$

Proof. As $\left|\mathbb{E}_{T_i|_{A=a,B=b}}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right| \leq 1$ for any setting of $S^i_{\rho_J}$,

$$\begin{aligned}\mathbb{E}_{S^i_{\rho_J}|_{A=a}}\left[\left|\mathbb{E}_{T_i|_{A=a,B=b}}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right] &\leq \mathbb{E}_{S^i_{\rho_J}}\left[\left|\mathbb{E}_{T_i|_{A=a,B=b}}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right] + \|S^i_{\rho_J}|_{A=a} - S^i_{\rho_J}\|_1 \\ &\leq \delta_2+\delta_1\end{aligned}$$

Finally, since $\Pr[B=b|A=a] > 0$, and $\left|\mathbb{E}_{T_i|_{A=a,B=b}}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right| \geq 0$, due to a simple Markov argument, we get

$$\mathbb{E}_{S^i_{\rho_J}|_{A=a,B=b}}\left[\left|\mathbb{E}_{T_i|_{A=a,B=b}}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right] \leq \frac{\delta_2+\delta_1}{\Pr[B=b|A=a]}$$

which completes the proof of the claim. □

We prove the final claim of this section, where we argue that $Z, B_{\rho_J}, G_{\rho_J}, C_{\rho_J}, E_i$ introduce little correlation between $S^i_{\rho_J}$ and T_i.

Claim 3.10 (Low Correlation)**.**

$$I(S^i_{\rho_J};T_i|Z,\rho_J,B_{\rho_J},G_{\rho_J}=1,E_i=1) \leq \frac{4C}{l} + \frac{25}{n}$$

Proof. Towards the proof, we introduce an auxiliary event, C_{ρ_J} for the analysis, which is true if a setting of Z, ρ_J is picked such that

$$1 - \Pr[G_{\rho_J}|Z, \rho_J, E_i = 1] < \frac{1}{2}.$$

Then recall that due to our definition of E_i^2 (from Claim 3.5),

$$\mathbb{E}_{Z,\rho_J|E_i=1}\left[1 - \Pr[G_{\rho_J}|Z, \rho_J, E_i = 1]\right] \leq \frac{10}{n^2}.$$

Due to a simple Markov argument

$$1 - \Pr_{Z,\rho_J|E_i=1}[C_{\rho_J}|E_i = 1] < \frac{20}{n^2} \tag{7}$$

With C_{ρ_J} defined, the bound on $I(S^i_{\rho_J}; T_i|Z, \rho_J, B_{\rho_J}, G_{\rho_J} = 1, E_i = 1)$ can be decomposed as

$$\begin{aligned}
&I(S^i_{\rho_J}; T_i|Z, \rho_J, B_{\rho_J}, G_{\rho_J} = 1, E_i = 1) \leq I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J} = 1, E_i = 1) \\
&= \Pr[C_{\rho_J} = 1|G_{\rho_J} = 1, E_i = 1] \cdot I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J} = 1, C_{\rho_J} = 1, E_i = 1) \qquad (8) \\
&+ \Pr[C_{\rho_J} = 0|G_{\rho_J} = 1, E_i = 1] \cdot I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J} = 1, C_{\rho_J} = 0, E_i = 1) \qquad (9)
\end{aligned}$$

Then we bound (8) and (9) separately. For (9), observe that for any fixed Z, ρ_J

$$I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J} = 1, C_{\rho_J} = 0, E_i = 1) \leq n_i$$

while from a simple application of Bayes's rule,

$$\Pr[C_{\rho_J} = 0|G_{\rho_J} = 1, E_i = 1] = \frac{\Pr[C_{\rho_J} = 0 \wedge G_{\rho_J} = 1|E_i = 1]}{\Pr[G_{\rho_J} = 1|E_i = 1]} \leq \frac{\frac{20}{n^2} \cdot \frac{1}{2}}{1 - \frac{10}{n^2}} \leq \frac{20}{n^2}$$

which therefore leads to a bound of

$$(9) \leq \frac{20n}{n^2} = \frac{20}{n}. \tag{10}$$

On the other hand, (8) can be bounded by

$$\begin{aligned}
&I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J} = 1, C_{\rho_J} = 1, E_i = 1) \\
&\leq \frac{I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J}, C_{\rho_J} = 1, E_i = 1)}{\Pr[G_{\rho_J} = 1|Z, \rho_J, C_{\rho_J} = 1, E_i = 1]} \leq 2 \cdot I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J}, C_{\rho_J} = 1, E_i = 1)
\end{aligned}$$

where the bound on $\Pr[G_{\rho_J} = 1|Z, \rho_J, C_{\rho_J} = 1, E_i = 1]$ follows from our definition of C_{ρ_J}. Then $I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J}, C_{\rho_J} = 1, E_i = 1)$ can be bounded as

$$\begin{aligned}
&I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, G_{\rho_J}, C_{\rho_J} = 1, E_i = 1) \\
&\leq I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, C_{\rho_J} = 1, E_i = 1) + \underbrace{H(G_{\rho_J}|Z, \rho_J, C_{\rho_J} = 1, E_i = 1)}_{\leq n^{-1}} \\
&\leq \frac{I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, E_i = 1)}{\Pr[C_{\rho_J} = 1|E_i = 1]} + \frac{1}{n} \leq 2 \cdot I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, E_i = 1) + \frac{1}{n}
\end{aligned}$$

where the second inequality follows from

$$\Pr[C_{\rho_J} = 1|E_i = 1] \cdot I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, C_{\rho_J} = 1, E_i = 1) \leq I(S^i_{\rho_J} B_{\rho_J}; T_i|Z, \rho_J, E_i = 1).$$

Then we bound the $I(S^i_{\rho_J} B_{\rho_J}; T_i | Z, \rho_J, E_i = 1)$ term.

$$\begin{aligned}
&I(S^i_{\rho_J} B_{\rho_J}; T_i | Z, \rho_J, E_i = 1) = I(S^i_{\rho_J} B_{\rho_J}; T_i | M, \rho_{<J}, \rho_J, B_{\rho_{<J}}, S^i_{\rho_{<J}}, J, E_i = 1) \\
&\leq I(S^i_{\rho_J} B_{\rho_J}; T_i | M, \mathcal{P}, B_{\rho_{<J}}, S^i_{\rho_{<J}}, J, E_i = 1) \\
&= \frac{1}{l} \sum_{j=1}^{l} I(S^i_{\rho_j} B_{\rho_j}; T_i | M, E_i = 1, \mathcal{P}, B_{\rho_{<j}}, S^i_{\rho_{<j}}) \\
&= \frac{1}{l} I(S^i_{\mathcal{P}} B_{\mathcal{P}}; T_i | M, E_i = 1, \mathcal{P}) \leq \frac{I(\vec{S}^i, B_{\mathcal{P}}; T_i | M, E_i = 1, \mathcal{P})}{l}
\end{aligned}$$

where the first inequality holds as $\mathcal{P}$ is chosen independently at random, thus conditioning on $\rho_{>J}$ increases mutual information from Fact 2.8. Then the $I(\vec{S}^i, B_{\mathcal{P}}; T_i | M, E_i = 1, \mathcal{P})$ term can be bounded by

$$I(\vec{S}^i, B_{\mathcal{P}}; T_i | M, E_i = 1, \mathcal{P}) = \underbrace{I(\vec{S}^i; T_i | M, E_i = 1, \mathcal{P})}_{\leq C} + \underbrace{I(B_{\mathcal{P}}; T_i | M, \vec{S}^i, E_i = 1, \mathcal{P})}_{=0} \leq C$$

where $I(B_{\mathcal{P}}; T_i | M, E_i = 1, \vec{S}^i, \mathcal{P}) = 0$ as $M, \vec{S}^i$ fully determine B_i, G_i's for all $i \in [m]$, and

$$I(\vec{S}^i; T_i | M, E_i = 1, \mathcal{P}) \leq I(\vec{S}^i; M, T_i | E_i = 1, \mathcal{P}) = \underbrace{I(\vec{S}^i; T_i | E_i = 1, \mathcal{P})}_{=0} + I(\vec{S}^i; M | T_i, E_i = 1, \mathcal{P}) \leq C.$$

as E_i is determined by $\vec{S}, T_{-i}$, and the distribution over T_i remains unchanged. Therefore, we get that

$$I(S^i_{\rho_J} B_{\rho_J}; T_i | Z, \rho_J, G_{\rho_J}, C_{\rho_J} = 1, E_i = 1) \leq \frac{2C}{l} + \frac{1}{n}$$

which then implies the bound on (8) as

$$(8) \leq 2 \left(\frac{2C}{l} + \frac{1}{n} \right) = \frac{4C}{l} + \frac{2}{n}.$$

This completes the proof of the claim as

$$I(S^i_{\rho_J} B_{\rho_J}; T_i | Z, \rho_J, G_{\rho_J} = 1, E_i = 1) \leq \frac{4C}{l} + \frac{2}{n} + \frac{20}{n} \leq \frac{4C}{l} + \frac{25}{n}$$

□

3.2.2 Proof of Theorem 3.6

With all the technical ingredients in place, we are now ready to prove Theorem 3.6, which then contradicts Theorem 3.2.

Theorem 3.6. *If $|M| \leq C = \frac{n_i}{\text{poly} \log n}$, the maximum advantage Bob can attain, conditioned on an event E_i determined by $\vec{S}, T_{-i}$ is*

$$\overrightarrow{\mathsf{adv}}(G_{\overline{f}_i}, \mathcal{D}_{\mathcal{Q}'_i}, \mathcal{D}_{T_i}, C) \leq n^{-\Omega(1)}$$

Proof. Suppose for every setting of $\vec{S}$, conditioned on the event E_i, there exists a one-way message M and guessed value B_Q (depending on M and Q) such that

$$\mathbb{E}_{M, Q \in \mathcal{Q}'_i} \left[\left| \mathbb{E}_{T_i} \left[B_Q \cdot \overline{f}_i(S_Q, T_i) \right] \right| \right] \geq n^{-o(1)}$$

which then implies

$$\mathbb{E}_{\vec{S},M,Q\in\mathcal{Q}'_i|E_i}\left[\left|\mathbb{E}_{T_i}\left[B_Q\cdot\overline{f}_i(S_Q,T_i)\right]\right|\right]\geq n^{-o(1)}.$$

As our setting of random process Z is fully determined by the above $\vec{S},M,Q$ (modulo independent randomness),

$$\mathbb{E}_{Z,B_{\rho_J},S^i_{\rho_J}|E_i,G_{\rho_J},\rho_J}\left[\left|\mathbb{E}_{T_i}\left[B_{\rho_J}\cdot\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\geq n^{-o(1)}. \tag{11}$$

For the rest of the proof, we will show that (11) cannot be the case. Without loss of generality, we will assume the following for any fixed ρ_J and $Z=z$

$$\Pr[B_Q|Z=z,\rho_J=\rho_j,G_{\rho_J},E_i]\geq n^{-0.1} \tag{12}$$

by incurring a cost of only $n^{-0.1}$ in the advantage due to the following observation: With probability $n^{-0.1}$, flip the guessed value. The loss in the advantage then is at most $n^{-0.1}$, which is polynomially small (compared to $n^{-o(1)}$).

Recall that we can write the advantage over random guessing as

$$\mathbb{E}_{Z,\rho_J,B_{\rho_J},S^i_{\rho_J}|G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z,B_{\rho_J},\rho_J,S^i_{\rho_J},G_{\rho_J},E_i}\left[B_{\rho_J}\cdot\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]. \tag{13}$$

Then we can upper bound (13) as

$$\begin{aligned}
(13)&\leq\mathbb{E}_{Z,\rho_J,B_{\rho_J},S^i_{\rho_J}|G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z,B_{\rho_J},\rho_J,S^i_{\rho_J},G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\\
&=\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\mathbb{E}_{S^i_{\rho_J}|Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|S^i_{\rho_J},Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\right]\\
&\leq\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\mathbb{E}_{S^i_{\rho_J}|Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\right] &(14)\\
&+\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\mathbb{E}_{S^i_{\rho_J}|Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\left[\|T_i|_{Z,B_{\rho_J},\rho_J,S^i_{\rho_J},G_{\rho_J},E_i}-T_i|_{Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\|_1\right]\right] &(15)
\end{aligned}$$

Now if we focus on the (15) term, we get

$$\begin{aligned}
(15)&=\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\|S^i_{\rho_J}|_{Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\times T_i|_{Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}-(S^i_{\rho_J},T_i)|_{Z,B_{\rho_J},\rho_J,G_{\rho_J},E_i}\|_1\right]\\
&\leq\sqrt{2\cdot I(S^i_{\rho_J};T_i|Z,\rho_J,B_{\rho_J},G_{\rho_J}=1,E_i=1)}\leq 20\sqrt{\frac{C}{l}+\frac{1}{n}}
\end{aligned} \tag{16}$$

due to Claim 3.10 and Pinsker's inequality (Fact 2.6).

Next, we proceed to bound (14) using Claim 3.7, and Claim 3.9. We first consider a fixed $Z=z$ ρ_J and $B_{\rho_J}=b$ conditioned on E_i,G_{ρ_J}. That is,

$$\mathbb{E}_{S^i_{\rho_J}|Z=z,B_{\rho_J}=b,\rho_J,G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z=z,B_{\rho_J}=b,\rho_J,G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right] \tag{17}$$

Now we are ready to use Claim 3.9 to bound (17). Claim 3.9 implies that for any fixed setting of Z,ρ_J and $B_{\rho_J}=b$,

$$\begin{aligned}
&\mathbb{E}_{S^i_{\rho_J}|Z=z,\rho_J,B_{\rho_J}=b,G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J}=b,G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\\
&\leq\frac{\|S^i_{\rho_J}|_{Z=z,\rho_J,E_i,G_{\rho_J}}-S^i_{\rho_J}\|_1+\mathbb{E}_{S^i_{\rho_J}}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J}=b,G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]}{\Pr[B_{\rho_J}=b|Z=z,\rho_J,G_{\rho_J},E_i]}
\end{aligned}$$

Now we would like to take expectation over Z and B_{ρ_J} and ρ_J conditioned on the event G_{ρ_J} and E_i. Decomposing the upper bound for (14), we obtain

$$(14) \leq \mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\frac{\|S^i_{\rho_J}|_{Z=z,\rho_J,E_i,G_{\rho_J}} - S^i_{\rho_J}\|_1}{\Pr[B_{\rho_J}=b|Z=z,\rho_J,G_{\rho_J}=1,E_i]}\right] \tag{18}$$

$$+\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\frac{\mathbb{E}_{S^i_{\rho_J}}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J}=b,G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]}{\Pr[B_{\rho_J}=b|Z=z,\rho_J,G_{\rho_J}=1,E_i]}\right] \tag{19}$$

The bound on (18) follows immediately from Claim 3.8.

$$(18) \leq 5\cdot\sqrt{\frac{(C+l)}{m}} + \frac{100}{n^2}$$

While for (19), rewriting the expectation and applying (12), we get

$$(19) \leq n^{0.1}\cdot\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\mathbb{E}_{S^i_{\rho_J}}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J},G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\right].$$

Now we would like to use the assumption about f_i to upper bound:

$$\begin{aligned}&\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\mathbb{E}_{S^i_{\rho_J}}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J},G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\right]\\ &=\mathbb{E}_{S^i_{\rho_J}}\left[\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J},G_{\rho_J},E_i}\left[\overline{f}_i(S^i_{\rho_J},T_i)\right]\right|\right]\right]\end{aligned} \tag{20}$$

Recall that due to our assumption on f_i, in particular, the low discrepancy condition, we have that for any $\vec{S}, T_{-i}$, $Q\in\mathcal{Q}'$, message M which is at most C bits,

$$\mathbb{E}_{\varsigma^i_{\rho_J}\sim\mathcal{U}}\left[\mathbb{E}_{M|\vec{S},T_{-i}}\left[\left|\mathbb{E}_{T_i|M}\left[\Psi_i(\varsigma^i_{\rho_J},T_i)\right]\right|\right]\right] \leq n^{-2}. \tag{21}$$

Recall that

$$Z := M, \mathcal{P}_{<J}, G_{\rho_{<J}}, B_{\rho_{<J}}, S^i_{\rho_{<J}}, J.$$

(21) then implies (by taking the expectation over $\vec{S^i}$), as M and $\vec{S^i}$ fully determine all the random variables, that for any fixed $\varsigma^i_{\rho_J}$ (chosen independently, thus no relationship to $\vec{S^i}$)

$$\begin{aligned}&\mathbb{E}_{Z,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J}=b,G_{\rho_J},E_i}\left[\overline{f}_i(\varsigma^i_{\rho_J},T_i)\right]\right|\right]\\ &=\mathbb{E}_{M,\mathcal{P}_{<J},G_{\rho_{<J}},B_{\rho_{<J}},S^i_{\rho_{<J}},J,\rho_J,B_{\rho_J}|G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|Z=z,\rho_J,B_{\rho_J}=b,G_{\rho_J},E_i}\left[\overline{f}_i(\varsigma^i_{\rho_J},T_i)\right]\right|\right]\\ &\leq\mathbb{E}_{M,\vec{S^i},J,\rho_J|G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|M,\vec{S^i},J,\rho_J,G_{\rho_J},E_i}\left[\overline{f}_i(\varsigma^i_{\rho_J},T_i)\right]\right|\right]\\ &=\mathbb{E}_{\vec{S^i},J,\rho_J|G_{\rho_J},E_i}\mathbb{E}_{M|\vec{S^i},J,\rho_J,G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|M,\vec{S^i},J,\rho_J,G_{\rho_J},E_i}\left[\overline{f}_i(\varsigma^i_{\rho_J},T_i)\right]\right|\right].\end{aligned}$$

We remark that $\vec{S^i}, J, \rho_J, G_{\rho_J}, E_i$ are all independent of T_i, therefore $M|\vec{S^i}, J, \rho_J, G_{\rho_J}, E_i$ is a C-bit message about T_i. Then taking expectation over $\varsigma^i_{\rho_J}\sim\mathcal{U}$, we get

$$\begin{aligned}&\mathbb{E}_{\varsigma^i_{\rho_J}\sim\mathcal{U}}\left[\mathbb{E}_{\vec{S^i},J,\rho_J|G_{\rho_J},E_i}\mathbb{E}_{M|\vec{S^i},J,\rho_J,G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|M,\vec{S^i},J,\rho_J,G_{\rho_J},E_i}\left[\overline{f}_i(\varsigma^i_{\rho_J},T_i)\right]\right|\right]\right]\\ &=\mathbb{E}_{\vec{S^i},J,\rho_J|G_{\rho_J},E_i}\left[\underbrace{\mathbb{E}_{\varsigma^i_{\rho_J}\sim\mathcal{U}}\left[\mathbb{E}_{M|\vec{S^i},J,\rho_J,G_{\rho_J},E_i}\left[\left|\mathbb{E}_{T_i|M,\vec{S^i},J,\rho_J,G_{\rho_J},E_i}\left[\overline{f}_i(\varsigma^i_{\rho_J},T_i)\right]\right|\right]\right]}_{\leq n^{-2}\text{ via (21)}}\right]\leq n^{-2}\end{aligned}$$

This in turn implies a bound on (19),

$$(19) \leq n^{0.1} \cdot n^{-2} = n^{-1.9}.$$

Finally we get the bound on (14) assuming $\frac{(C+l)}{m}, \frac{C}{l} = n^{-\Omega(1)}$, which is true as $m = \Omega(n^{0.99})$ and by choosing $l = n^{0.9}$, as $n_i \leq n^{0.8}$ (since the number of blocks $n_i \leq n/k$ and $i \leq 5\ell/6$). This results in

$$(14) \leq n^{-1.9} + 5 \cdot \sqrt{\frac{(C+l)}{m}} + \frac{100}{n^2} \leq n^{-\Omega(1)}$$

which then implies our main bound on (13) as

$$(13) \leq n^{-\Omega(1)} + 20\sqrt{\frac{C}{l} + \frac{1}{n}} \leq n^{-\Omega(1)}.$$

a contradiction.

□

4 Applications

In this section, we list some applications of our black-box theorem. Note that we only need to show that some particular function f is "hard." We introduce some additional preliminaries towards that goal.

4.1 Preliminaries for Min-Entropy

We will need the following properties of min-entropy for the proof of one-way communication lower bounds.

Definition 4.1. *We define the renyi entropy $H_2(A)$ and min-entropy $H_\infty(A)$ as*

$$H_2(A) := -\log\left(\sum_a \Pr[A = a]^2\right)$$
$$H_\infty(A) := -\log\left(\max_a \Pr[A = a]\right)$$

Fact 4.2 (Renyi Entropy)**.** *Let A be a random variable. Then*

$$H(A) \geq H_2(A) \geq H_\infty(A)$$

In particular, for any fixed b we have

$$H_2(A|B = b) \geq H_\infty(A|B = b)$$

We use the following lemma on "average" min-entropy.

Definition 4.3 (Average Min-Entropy)**.**

$$\widetilde{H}_\infty(A|B) = -\log\left(\mathbb{E}_{b\sim B}\left[\max_a \Pr[A = a|B = b]\right]\right) = -\log\left(\mathbb{E}_{b\sim B}\left[2^{-H_\infty(A|B=b)}\right]\right)$$

Lemma 4.4 (Lemma 2.2 of [DORS08])**.** *Let A, B be random variables. Then if B has at most 2^λ possible values, then*

$$\widetilde{H}_\infty(A|B) \geq \widetilde{H}_\infty(A, B) - \lambda \geq H_\infty(A) - \lambda.$$

Claim 4.5.

$$\widetilde{H}_\infty(A|B,C) \le \widetilde{H}_\infty(A|B)$$

Proof. We first proceed with showing the following inequality,

$$\mathbb{E}_{c\sim C|_{B=b}}\left[\max_a \Pr[A=a|B=b,C=c]\right] \ge \max_a \Pr[A=a|B=b]. \tag{22}$$

Let $a^* := \arg\max_a \Pr[A=a|B=b]$. Then

$$\Pr[A=a^*|B=b] = \mathbb{E}_{c\sim C|_{B=b}}\left[\Pr[A=a^*|B=b,C=c]\right] \le \mathbb{E}_{c\sim C|_{B=b}}\left[\max_a \Pr[A=a|B=b,C=c]\right]$$

With (22) established, we are ready to prove the claim. Recall that

$$\widetilde{H}_\infty(A|B,C) := -\log\left(\mathbb{E}_{b,c\sim B,C}\left[\max_a \Pr[A=a|B=b,C=c]\right]\right)$$

Taking expectation over B on both side of (22), we have

$$\mathbb{E}_{b,c\sim B,C}\left[\max_a \Pr[A=a|B=b,C=c]\right] \ge \mathbb{E}_{b\sim B}\left[\max_a \Pr[A=a|B=b]\right]$$

Therefore, we get

$$\begin{aligned}\widetilde{H}_\infty(A|B,C) &= -\log\left(\mathbb{E}_{b,c\sim B,C}\left[\max_a \Pr[A=a|B=b,C=c]\right]\right)\\ &\le -\log\left(\mathbb{E}_{b\sim B}\left[\max_a \Pr[A=a|B=b]\right]\right) = \widetilde{H}_\infty(A|B)\end{aligned}$$

□

We can also bound the support size of the distribution using the min-entropy.

Claim 4.6. *Let the distribution on a random variable $X \in \{0,1\}^n$ conditioned on some random variable $M=m$ be supported over the set $T_m \subset \{0,1\}^n$. Then $|T_m| \ge 2^{H_\infty(X|M=m)}$*

Proof. Recall that

$$H_\infty(X|M=m) := -\log \max_x \left(\Pr[X=x|M=m]\right).$$

Therefore, we have

$$2^{-H_\infty(X|M=m)} \cdot |T_m| \ge \sum_{x\in T_m} \Pr[X=x|M=m] = 1,$$

which completes the proof of the claim. □

The main technical issue with using min-entropy instead of KL-divergence is that **chain-rule does not hold for min-entropy**. Nevertheless, we have the following strong chain-rule if we are willing to "spoil" a few bits.

Theorem 4.7 ([Sko19])**.** *Let $\mathcal{X}$ be a fixed alphabet, and $X=(X_1,\ldots,X_t)$ be a sequence of (possibly correlated) random variables each over $\mathcal{X}$, equipped with a distribution μ. Then for any $\varepsilon \in (0,1)$ and $\delta > 0$, there exists a collection $\mathcal{B}$ of disjoint sets on $\mathcal{X}^t$ such that*

- *$\mathcal{B}$ can be indexed by a small number of bits, namely*

$$\log|\mathcal{B}| = t \cdot O\left(\log\log|\mathcal{X}| + \log\log\varepsilon^{-1} + \log(t/\delta)\right)$$

- $\mathcal{B}$ *almost covers the domain*
$$\sum_{B\in\mathcal{B}} \mu(B) \geq 1-\varepsilon$$
- *Conditioned on $\mathcal{B}$, block distributions $X_i|X_{<i}$ are nearly flat. That is*
$$\forall x, x' \in B, \quad 2^{-O(\delta)} \leq \frac{\mu(x_i|x_{<i})}{\mu(x'_i|x'_{<i})} \leq 2^{O(\delta)}$$
- *For every $B \in \mathcal{B}$, for every index $i \in [t]$, and for every set $I \subset [i-1]$, we have*
  1. *The chain-rule for min-entropy*
  $$H_\infty(X_i|X_I, B) + H_\infty(X_I|B) = H_\infty(X_i, X_I|B) \pm O(\delta)$$
  2. *The average and worst-case min-entropy almost match*
  $$\widetilde{H}_\infty(X_i|X_I, B) = H_\infty(X_i|X_I, B) \pm O(\delta)$$

The corollary of Theorem 4.7 that we will need is the following

Corollary 4.8. *For any fixed $B \in \mathcal{B}$*
$$\sum_{i=1}^{t} \widetilde{H}_\infty(X_i|X_{<i}, B) = H_\infty(X|B) \pm t \cdot O(\delta).$$
In other words, if we take a uniform distribution over $[t]$,
$$\mathbb{E}_{i\in[t]}\left[\widetilde{H}_\infty(X_i|X_{<i}, B)\right] = \frac{H_\infty(X|B)}{t} \pm O(\delta)$$

Proof. We prove by induction on the number of summands. Consider the base case. If there is a single summand,
$$\widetilde{H}_\infty(X_1|B) = H_\infty(X_1|B)$$
thus this is trivially true. Now assume (as inductive hypothesis) that
$$\sum_{i=1}^{j-1} \widetilde{H}_\infty(X_i|X_{<i}, B) = H_\infty(X_1, \ldots, X_{j-1}|B) \pm (j-1) \cdot O(\delta).$$
Then for $H_\infty(X_1, \ldots, X_j|B)$, Theorem 4.7 implies that
$$\begin{aligned} H_\infty(X_1, \ldots, X_j|B) &= \widetilde{H}_\infty(X_j|X_{<j}, B) + H_\infty(X_{<j}|B) \pm O(\delta) \\ &= \widetilde{H}_\infty(X_j|X_{<j}, B) + \sum_{i=1}^{j-1} \widetilde{H}_\infty(X_i|X_{<i}, B) \pm j \cdot O(\delta) \end{aligned}$$
where the second equality holds by the inductive hypothesis. This then completes the proof. □

For our application, we will take $\varepsilon = n^{-\Omega(1)}$. By setting ε as such, even if we condition on being inside the support of $\mathcal{B}$, the advantage of the one-way protocol is not affected.

We will use the following corollary/variant of KKL Theorem. We follow the standard notations from [O'D14], to which we refer the reader for Fourier Analysis on Boolean functions.

Lemma 4.9 (KKL-Theorem)**.** *Let $T \subset \{0,1\}^n$, and $\delta \in [0,1]$. Let f be a function supported on T such that for all $x \in T$, $|f(x)| \leq 1+\varepsilon$. Suppose we write $\widetilde{f}(S) := \frac{\widehat{f}(S)2^n}{|T|}$. Then*

$$\sum_{S \subset [n]} \delta^{|S|} \widetilde{f}(S)^2 \leq (1+\varepsilon)^2 \left(\frac{2^n}{|T|}\right)^{2\delta}$$

Proof. Follows from the usual proof of KKL-Theorem using Hypercontractivity. Let $p = 1+\delta$, and $\rho = \sqrt{p-1}$. Then

$$\|f\|_p^2 = \left(\sum_{x \in T} f(x)^p\right)^{2/p} \leq \left((1+\varepsilon)^p \cdot \frac{|T|}{2^n}\right)^{2/p} = (1+\varepsilon)^2 \cdot \left(\frac{|T|}{2^n}\right)^{2/p}$$
$$\|T_\rho(f)\|_2^2 = \sum_{S \subset [n]} \delta^{|S|} \widehat{f}(S)^2$$

Dividing both sides by $(|T|/2^n)^2$, and using the well-known Hypercontractivity which states $\|T_\rho(f)\|_2^2 \leq \|f\|_p^2$, we obtain

$$\sum_{S \subset [n]} \delta^{|S|} \widetilde{f}(S)^2 \leq \left(\frac{|T|}{2^n}\right)^{-2\delta}$$

which completes the proof. □

4.2 Super-Logarithmic Lower Bound for the Multiphase Problem

We simply need to show that inner product (mod 2) is "hard" as in Definition 1.2.

Lemma 4.10. IP *satisfies Definition 1.2.*

Proof. For any fixing of S_Q^{-i}, T_{-i}, indeed the function is balanced with high probability. As long as S_Q^i is not all zero string. The function is perfectly balanced. The probability of such event is at most $2^{-n_i} \leq 2^{-n^{1/4}}$, thereby satisfying the balanced condition.

It remains to show that f_i has low discrepancy. f_i also has low discrepancy due to the following observation. The corresponding Ψ_i is the Hadamard matrix. Then we can use the following well-known lemma to bound the discrepancy.

Fact 4.11 (Lindsey's Lemma)**.** *Let H be a Hadamard matrix. Let P and Q be distributions. Then*

$$P^T H Q \leq \|P\|_2 \|Q\|_2 \cdot 2^{n/2}$$

Fact 4.11 then implies that T_i under the distribution $T_i|_{M=m}$ (the prior on T_i conditioned on the message $M=m$) has

$$\mathsf{odisc}_{T_i|_{M=m}}(\Psi_i) = v_m^T \cdot H_{n_i} \cdot T_i|_{M=m} \leq 2^{n_i/2} \cdot \underbrace{\|v_m\|_2}_{=2^{-n_i/2}} \cdot \underbrace{\|T_i|_{M=m}\|_2}_{\leq 2^{-\frac{H_\infty(T_i|M=m)}{2}}} \leq 2^{-\frac{H_\infty(T_i|M=m)}{2}}$$

where $v_m \in \{\pm 2^{-n_i}\}^{2^{n_i}}$ is a vector that matches the sign per coordinate of $H_{n_i} \cdot T_i|_{M=m}$, justifying the norm of such vector.

Since M is of length at most $c = \frac{n_i}{\mathrm{poly}\log(n)}$ and recall that we have defined

$$\mathsf{odisc}_c(\Psi_i) := \max_{M:|M|<c} \mathbb{E}_M \left[\mathsf{odisc}_{T_i|_M}(\Psi_i)\right]$$

$$\mathbb{E}_M\left[\mathsf{odisc}_{T_i|M}(\Psi_i)\right] \le \mathbb{E}_M\left[2^{-\frac{H_\infty(T_i|M=m)}{2}}\right] \le \sqrt{\mathbb{E}_M\left[2^{-H_\infty(T_i|M=m)}\right]} = 2^{-\frac{\widetilde{H}_\infty(T_i|M)}{2}}$$

where the second inequality holds from Jensen's inequality (on $\sqrt{x}$) and the equality holds from the definition of average min-entropy. Lemma 4.4 then implies that $\widetilde{H}_\infty(T_i|M) \ge n_i - c$, which in turn implies

$$\mathsf{odisc}_c(\Psi_i) \le 2^{-\Omega(n_i)},$$

thereby satisfying the low discrepancy if $c = \frac{n_i}{\mathrm{poly}\log(n)}$. □

Remark 4.12. *Note that the inner product is way stronger than what is necessary for our proof. Nevertheless, this gives the state-of-the-art bound for the Multiphase Problem [Pat10].*

4.3 A Simple Lifting Theorem

Lifting technique refers to translating a lower bound for a weaker model to a lower bound for a stronger general model. A classic example of the lifting theorem is Sherstov's Pattern Matrix method [She11] which translates approximate degree lower bounds (a structural lower bound for a weaker model) into approximate-rank and communication lower bounds.

Unfortunately, we cannot make a full translation, as we require the underlying hard distribution to be **a product distribution**, instead of a general distribution. In order to completely translate between the some algebraic properties of ψ such as its approximate degree (See [BT22] and references therein), to its communication complexity, we need to consider all possible distribution, instead of just product distribution.

Nevertheless, the following lemma from [She11] can be used to create hard functions. We would like to lift the function using the following composition, $f = \psi \circ g_k^n$ where g (the so-called inner gadget) is

$$g((j,b),(x_{i+1},\dots,x_{i+k})) = b \oplus x_{i+j}.$$

Now the task in our one-way communication is the following. Alice is given $\varsigma \in ([k] \times \{\pm 1\})^n$, Bob is given $(X_1,\dots,X_n) \in \left(\{\pm 1\}^k\right)^n$. Bob sends a one-way message for Alice to compute

$$f(\varsigma, X) = \psi(g(\varsigma_1, X_1),\dots,g(\varsigma_n, X_n)).$$

We consider the composition, $\psi = \varphi(\psi_1,\dots,\psi_\ell)$. The only property we need about φ is that when ψ_{-i}'s are fixed, ψ value is still undetermined. That is $\psi_i' := \psi|_{\psi_{-i}}$ satisfies the "hard" condition. For example, it is relatively straightforward to see that if ψ_i is balanced, then $\bigoplus_{i=1}^\ell \psi_i$ is also balanced. For $\bigvee_{i=1}^\ell \psi_i$, we need to have ψ_i biased to say $\min\{\Pr[\psi_i = +1], \Pr[\psi_i = -1]\} \ge \ell^{-1}$, the trick introduced in [LY25].

A straightforward connection can be achieved using the following lemma, which can be viewed as a generalization of Fact 4.11.

Lemma 4.13 ([She11])**.**

$$\|\Psi_i\| = \sqrt{2^{kn}(2k)^n} \max_{S\subset[n]} \left(\left|\hat{\psi}(S)\right| \cdot k^{-|S|/2}\right)$$

Directly applying the bound by [She11] and the argument introduced in Section 4 to a uniform distribution over ς, we get

$$\mathsf{odisc}_c(\Psi_i) \le 2^{c/2} \cdot \max_{S\subset[n]} \left(\left|\hat{\psi}(S)\right| \cdot k^{-|S|/2}\right)$$

connecting the spectral property of the underlying ψ to its odisc. Note that this is matching (i.e. with Section 4) when ψ is a parity function.

Remark 4.14. *In standard communication complexity setting, Lemma 4.13 can be used to obtain a close relation between algebraic properties of ψ and the communication complexity of f (i.e. so-called lifting technique). Unfortunately, this is not the case for our application. We need a lower bound against a product distribution between Alice's input and Bob's input. But the lower bound achieved through connections to algebraic properties of ψ is not necessarily for a product distribution.*

5 Generalization of 0-XOR

As a further illustration of our method, we consider a generalization of a hard function (so-called 0-XOR) used in [LY25].

0-XOR Function Fix an epoch $i \in [\ell]$. We have the following function. If n_i, the size of epoch i, is small (say less than $\log^2(n)$), we have $f_i := 0$. For any other epochs, we consider the following function. We divide the updates to say $\log \ell$ columns and $R = O(\log n)$-rows. For fixed column d, we select one entry per row, denoted as $\varsigma_i[d]$. Then we take AND over the columns. Therefore our function in question becomes

$$f_i(\varsigma_i, T_i) := \mathsf{AND}^D_{d=1}\left(\langle \varsigma_i[d], T_i\rangle\right).$$

Then we take OR over the epochs $i \in [\ell]$. Note that regardless of which ς_i we select, if T_i's are distributed i.i.d. all bits being $\mathcal{B}_{1/2}$, f_i is 1 with probability $2^{-\ell} = \frac{1}{\ell}$. This ensures that our function $\mathsf{OR}(f_1, \dots f_{i-1}, f_{i+1}, f_\ell)$ is 1 with probability at most $1/2$ for any $i \in [\ell/2, \ell]$. In particular, the function is indeed balanced as from Definition 1.2.

5.1 Reinterpretation of [LY25]

[LY25] uses a distribution over $\varsigma_i[j]$'s which are highly correlated one another, stemming from the fact that the choice of ς_i derives from a distribution on the underlying graph (the butterfly graph).

This leads to the distribution over ς_i used in [LY25] not having a small odisc, as shown in Section 5.2 of [LY25]. This occurs due to odisc-quantity highly depending on the distribution over ς, while they need a specific distribution over ς due to their dependence on the underlying Butterfly Graph and the reduction to Graph Connectivity.

Instead, they show k-XOR version of the underlying problem is hard. In fact, one could then phrase their technical communication lower bound (on so-called Meta-queries) in Section 5.2 as the following in our notation.

Lemma 5.1 (Lemma 5.4 of [LY25])**.** *There exists a distribution $\mathcal{M}$ (which is a uniform distribution over its support of size $n^{\Omega(k)}$) on k-tuples $(\varsigma_i^1, \dots, \varsigma_i^k)$ with $k = n/\mathrm{poly}\log(n)$ such that if $|M| < o(n/\mathrm{poly}\log(n))$, then the advantage for outputting $\oplus_{j=1}^k f_i(\varsigma_i^j, T_i)$ is at most $2^{-\Omega(k \log n)}$.*

Observe that $\oplus_{j=1}^k f_i(\varsigma_i^j, T_i)$ is hard (as in Definition 1.2) with the parameter $n^{-\Omega(k)}$ instead of n^{-2}. Since a too-good-to-be true dynamic data structure (with query time t_{tot} and update time $t_u = \mathrm{poly}\log(n)$) for 0-XOR implies a data structure with query time $k \cdot t_{tot}$, this implies a one-way message M of length $|M| < o(n/\mathrm{poly}\log(n))$ with advantage $n^{-o(k)}$, a contradiction.

Though we omit the details as this is simply a reinterpretation, and do not imply any new result, our argument can further be extended to the settings where we select some subset of k-tuples (of size $n^{\Omega(k)}$) from the support of $\mathcal{M}$, instead of selecting all elements of $\mathcal{M}$ as the query set.

5.2 Independent Instances of 0-XOR

Our goal in this section to exhibit another useful example of our method and show that a generalization of 0-XOR when the queries are independently formed is "hard."

The main technical challenge here is that other than IP considered in Lemma 4.10, which more-or-less follows from the quasi-random property of the underlying Hadamard matrix (Fact 4.11), there does not seem to be a general technique for bounding the ε-advantage one-way protocols with the lower bound of the form $\Omega(n/\operatorname{poly}\log(1/\varepsilon))$ under a **product distribution** as far as we are aware of (See [Wat20] and references therein). And such a bound is what is necessary to satisfy Definition 1.2.

In this section, we develop novel technical tools to show 0-XOR function is "hard" using the average min-entropy and its relation to KKL Theorem, which can be of independent interest. We suspect connections to lower bounds in other models of computation such as streaming (See Chapter 2 of [Rou16]).

Theorem 5.2. *Suppose $\varsigma_i[j]$'s are selected independently at random. There exists some $D = \Omega(\log n)$ such that $f = \bigvee_{i=1}^{\ell} f_i$ with $f_i := \mathsf{AND}_{d=1}^{D}(\langle \varsigma_i[d], T_i\rangle)$ is hard as in Definition 1.2.*

Observe that balancedness of Definition 1.2 is given for free due to the choice of ς_i's. All it remains to show is that for any M of length $< n_i/\operatorname{poly}\log(n)$,

$$\mathbb{E}_M\left[\mathsf{odisc}_{T_i|_{M=m}}(\Psi_i)\right] \le n^{-2}.$$

Here is the rough outline of the proof.

1. By increasing the length of the message by attaching a message from $\mathcal{B}$, "spoiling" few bits, that is giving up on some $2^{-n^{1/3}}$ fraction of the input, and dividing the updates into blocks of size $K = \log^k(n)$, we can make the distribution easy for min-entropy (i.e. Theorem 4.7). This allows us to use (i) near chain-rule; (ii) KKL Theorem when conditioned on the message, two crucial components of the proof. We consider a set of "good" messages M and $\mathcal{B}$, which again constitute all but exponential fraction of the input. (See Lemma 5.4)

2. For every choice of $d \in [D]$, using the chain-rule, we argue that since the min-entropy on parts of the input must be large, the min-entropy (over a random choice of blocks) must be large. We show a stronger statement which states that with all but $n^{-\Omega(1)}$ choice of blocks, the min-entropy must be large. (See Lemma 5.5)

3. Using a variant of KKL Theorem in our context, we show that a large min-entropy over some small set of coordinates translates to small average bias. (See Lemma 5.7)

4. Iteratively apply step 2 and step 3 per $d \in [D]$. Via union bound over $d \in [D]$, this shows that for all but $n^{-\Omega(1)}$ choice of $\varsigma_i[d]$'s, the advantage must be $n^{-\Omega(1)}$ when conditioned on M and $\mathcal{B}$.

Notations We introduce the following notations for the proof. As we will divide the updates into blocks of size K each, these blocks will be arranged so that D columns and R rows will have integer number of blocks. Recall that $\varsigma_i[d]$ corresponds to a vector which chooses exactly one entry per row. Denote $T_i^{(d,r,j)}$ as j-th block of $K := \log^k(n)$ updates associated with the column d, row r. There will be D rows, R columns. So there are $\frac{n_i}{DRK}$ many possible blocks for the index j. We denote $T_i^{(d,r)}$ as all the blocks in column d, row r. Then we denote $T_i^{(d)}$ as all the blocks in column d.

5.2.1 Constructing $\mathcal{B}$ for min-entropy

Without loss of generality, we can assume that our message M is "good," in a following sense.

Proposition 5.3. *Write* $T_i := \left\{T_i^{(d,r,j)}\right\}_{d\in[D],r\in[R],j\in[\frac{n_i}{DRK}]}$. *Then there exists partitioning of* $T_i|_{M=m}$, *say* $\mathcal{B}$ *of size* $\log|\mathcal{B}| = O(\frac{n_i \log n}{K} + \frac{kn_i \log\log n}{K})$ *such that*

- $\sum_{B\in\mathcal{B}} \Pr[B|M=m] \geq 1 - \exp(n^{-\Omega(1)})$, *that is the partition covers most of* T_i.
- *Conditioned on* $B \in \mathcal{B}$, *block distributions* $T_i^j|T_i^{<j}$ *are nearly flat. That is*

$$\forall T_i, T_i' \in B, \quad 2^{-O(n^{-1})} \leq \frac{\Pr[T_i^j|T_i^{<j}, M=m, B]}{\Pr[T_i'^j|T_i'^{<j}, M=m, B]} \leq 2^{O(n^{-1})}$$

- *Suppose we give lexicographic ordering to* (d,r,j)*'s. For every* $B \in \mathcal{B}$, *for every index* (d,r,j), *and for every set* J *whose elements are all less than* (d,r,j), *we have*

  1. *The chain-rule for min-entropy*

$$\begin{aligned}&\min_t\{H_\infty(T_i^{(d,r,j)}|T_i^J = t, B, M=m)\} + H_\infty(T_i^J|B, M=m)\\ &= H_\infty(T_i^{(d,r,j)}, T_i^J|B, M=m) \pm n^{-1}\end{aligned}$$

  2. *The average and worst-case min-entropy almost match*

$$\widetilde{H}_\infty(T_i^{(d,r,j)}|T_i^J, B, M=m) = \min_t\{H_\infty(T_i^{(d,r,j)}|T_i^J = t, B, M=m)\} \pm n^{-1}$$

Proof. This is a direct corollary of Theorem 4.7 in our setting. By setting $\delta = n^{-1}, \varepsilon = n^{-\Omega(1)}$, and $\log|\mathcal{X}| = \log^k(n)$ on the distribution of T_i conditioned on $M=m$, Theorem 4.7 implies the existence of such $\mathcal{B}$ with the claimed parameters. □

Note that we can also make $\log|\mathcal{B}| = n_i/\mathrm{poly}\log(n)$ via adjusting the block size parameter K to match with the length of the message M. Thereby, we can attach the partitioning B along with the message $M=m$ as our one-way message from Bob. Without loss of generality, we will denote B_0 as the "spoiled" part, that is a set of T_i not covered by $\mathcal{B}$.

This implies the following lemma in our context, which states that the average min-entropy must be large when conditioned on $M=m$ and B.

Lemma 5.4. *Recall that* $|M| + |\mathcal{B}| = O(n_i \log n/K)$.

$$\Pr_{M,\mathcal{B}}\left[\exists d\in[D], \quad \widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B\right) \leq \frac{n_i}{D} - 2(|M|+|\mathcal{B}|)\right] \leq 2^{-n^{1/3}}$$

Proof. Due to Lemma 4.4, and the independence of $T_i^{(d)}$ and $T_i^{(<d)}$ (without conditioning on M and $\mathcal{B}$) we know that for any fixed $T_i^{(<d)} = t_i^{(<d)}$

$$\begin{aligned}\widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)} = t_i^{(<d)}, M, \mathcal{B}\right) &= -\log \mathbb{E}_{M,\mathcal{B},T_i^{(<d)}=t_i^{(<d)}}\left[2^{-H_\infty(T_i^{(d)}|T_i^{(<d)}=t_i^{(<d)}, M=m, B)}\right]\\ &\geq \frac{n_i}{D} - (|M|+|\mathcal{B}|)\end{aligned}$$

which in turn implies

$$\mathbb{E}_{M,\mathcal{B},T_i^{(<d)}}\left[2^{-H_\infty(T_i^{(d)}|T_i^{(<d)}=t_i^{(<d)},M=m,B)}\right] \leq 2^{-\frac{n_i}{D}+(|M|+|\mathcal{B}|)}$$

Due to a simple Markov's inequality, for a fixed $d \in [D]$

$$\Pr_{M,\mathcal{B}}\left[\mathbb{E}_{T_i^{(<d)}|M=m,\mathcal{B}=B}\left[2^{-H_\infty(T_i^{(d)}|T_i^{(<d)}=t_i^{(<d)},M=m,B)}\right] \geq 2^{-\frac{n_i}{D}+2(|M|+|\mathcal{B}|)}\right] \leq 2^{-(|M|+|\mathcal{B}|)} \tag{23}$$

Then applying union bound over $d \in [D]$ with (23),

$$\Pr_{M,\mathcal{B}}\left[\exists d \in [D],\ \ \widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B\right) \leq \frac{n_i}{D} - 2(|M|+|\mathcal{B}|)\right] \leq D \cdot 2^{-(|M|+|\mathcal{B}|)} \leq 2^{-n^{1/3}}$$

□

We will denote the set of message and partition pair $M = m, B$ induced from Lemma 5.4 as $\mathcal{G}_{m,B}$. That is define

$$\mathcal{G}_{m,B} := \left\{(m,B) | \forall d \in [D],\ \ \widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B\right) \geq \frac{n_i}{D} - 2(|M|+|\mathcal{B}|)\right\}. \tag{24}$$

For the rest of the inductive arguments, we will only consider (m, B) pairs from $\mathcal{G}_{m,B}$.

5.2.2 Large Min-Entropy per random choice of Blocks

For brevity denote O_d as the event $\mathsf{AND}_{j=1}^{d-1}\left(\langle \varsigma_i[j], T_i\rangle\right) = 1$. As inductive hypothesis, we assume that we have chosen $\varsigma_i[j]$'s for $j \in [d-1]$ such that

$$\Pr[O_d | M=m, B] \in \left[\left(\frac{1-n^{-\Omega(1)}}{2}\right)^{d-1}, \left(\frac{1+n^{-\Omega(1)}}{2}\right)^{d-1}\right] \tag{25}$$

Note that as a base case, when $d = 1$, the statement is trivially true, by assuming AND over null arguments to be 1.

Lemma 5.5. *Suppose $(m, B) \in \mathcal{G}_{m,B}$ and as inductive hypothesis, assume (25) holds for $d \in [D]$.*

$$\Pr_{j_1,\ldots,j_R \in \mathcal{U}\left[\frac{n_i}{DRK}\right]}\left[\widetilde{H}_\infty\left(\left\{T_i^{(d,r,j_r)}\right\}_{r=1}^R | T_i^{(<d)}, M=m, B, O_d\right) \leq (1-\gamma)KR\right] \leq \exp(-\Omega(\gamma^2 R))$$

Proof. We consider the following random variables (over the random choice of j_r's) to apply McDiarmid's Inequality by considering the real-valued function

$$h(j_1,\ldots,j_R) := \widetilde{H}_\infty\left(\left\{T_i^{(d,r,j_r)}\right\}_{r=1}^R | T_i^{(<d)}, M=m, B, O_d\right).$$

For completeness, we state the McDiarmid's Inequality here.

Fact 5.6 (McDiarmid's Inequality)**.** *Let $f : \mathcal{X}_1 \times \ldots \times \mathcal{X}_n \to \mathbb{R}$ satisfy the following property: there exists $c_1, \ldots, c_n$ such that for all $x_1 \in \mathcal{X}_1, \ldots, x_n \in \mathcal{X}_n$*

$$\sup_{x_i' \in \mathcal{X}_i} \left|f(x_1,\ldots,x_i,\ldots,x_n) - f(x_1,\ldots,x_i',\ldots,x_n)\right| \leq c_i$$

Then if $X_1, \ldots, X_n$ are chosen independently at random from $\mathcal{X}_1, \ldots, \mathcal{X}_n$ respectively,

$$\Pr\left[f(X_1,\ldots,X_n) - \mathbb{E}_{X_1,\ldots,X_n}[f(X_1,\ldots,X_n)] \leq -\varepsilon\right] \leq \exp\left(\frac{2\varepsilon^2}{\sum_{i=1}^n c_i^2}\right)$$

Note that due to the chain rule property, replacing a single coordinate can only lead to a difference of at most $K \pm n^{-1}$ in h. Therefore Fact 5.6 directly implies

$$\Pr_{j_1,\ldots,j_R \in \mathcal{U}}[h(j_1,\ldots,j_r) \leq \mathbb{E}[h] - \alpha] \leq \exp\left(-\frac{2\alpha^2}{R(K+n^{-1})^2}\right). \tag{26}$$

To transform (26) to a desired form, we calculate the expected value of h. Observe that we can decompose h (upto n^{-1} error) as

$$h(j_1,\ldots,j_R) = \sum_{r=1}^{R}\left(\underbrace{\widetilde{H}_\infty\left(T_i^{(d,r,j_r)}|T_i^{(d,<r,j_{<r})}, T_i^{(<d)}, M=m, B, O_d\right)}_{:=h_r(j_1,\ldots,j_R)} \pm \frac{1}{n}\right) \tag{27}$$

Observe that h_r only depends on $j_1,\ldots,j_r$. Consider a fixed r. For any choice of $j_1,\ldots,j_{r-1}$,

$$\begin{aligned}
\mathbb{E}_{j_r}[h_r] &= \frac{DRK}{n_i}\sum_{j_r}\widetilde{H}_\infty\left(T_i^{(d,r,j_r)}|T_i^{(d,<r,j_{<r})}, T_i^{(<d)}, M=m, B, O_d\right) \\
&\geq \frac{DRK}{n_i}\widetilde{H}_\infty\left(T_i^{(d,r)}|T_i^{(d,<r,j_{<r})}, T_i^{(<d)}, M=m, B, O_d\right) \pm \frac{1}{n} \\
&\geq \frac{DRK}{n_i}\widetilde{H}_\infty\left(T_i^{(d,r)}|T_i^{(d,<r)}, T_i^{(<d)}, M=m, B, O_d\right) \pm \frac{1}{n}
\end{aligned}$$

where the first inequality follows from Claim 4.5 and the chain-rule property induced by $\mathcal{B}$. The second inequality follows from Claim 4.5.

$$\begin{aligned}
\mathbb{E}[h] = \sum_{r=1}^{R}\mathbb{E}[h_r] &\geq \frac{DRK}{n_i}\sum_{r=1}^{R}\left(\widetilde{H}_\infty\left(T_i^{(d,r)}|T_i^{(d,<r)}, T_i^{(<d)}, M=m, B, O_d\right) - 2n^{-1}\right) \\
&\geq \frac{DRK}{n_i}\left(\widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B, O_d\right) - 4Rn^{-1}\right) \\
&\geq \frac{DRK}{n_i}\left(\widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B\right) - 4Rn^{-1} - \log\frac{1}{\Pr[O_d|M=m,B]}\right)
\end{aligned} \tag{28}$$

where the last bound holds due to the definition of O_d and average min-entropy, which implies

$$\Pr[O_d|M=m,B]\cdot 2^{-\widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B, O_d\right)} \leq 2^{-\widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B\right)}.$$

Recall that due to our assumption that $(m,B) \in \mathcal{G}_{m,B}$,

$$\widetilde{H}_\infty\left(T_i^{(d)}|T_i^{(<d)}, M=m, B\right) = \frac{n_i}{D} - 2(|M|+|\mathcal{B}|) \geq \frac{n_i}{D} - O(n_i \log n/K)$$

as well, then

$$\mathbb{E}[h] \geq RK\underbrace{-O(DR\log n) - \frac{4DR^2K}{n\cdot n_i} - \frac{DRK}{n_i}\cdot\log\frac{1}{\Pr[O_d|M=m,B]}}_{-o(RK)}.$$

This implies that

$$\begin{aligned}
&\Pr_{j_1,\ldots,j_R\in\mathcal{U}}[h(j_1,\ldots,j_r) < RK - \gamma RK] \\
&\leq \Pr_{j_1,\ldots,j_R\in\mathcal{U}}[h(j_1,\ldots,j_r) < \mathbb{E}[h] - (\gamma+o(1))RK] \leq \exp\left(-\frac{3}{2}\gamma^2 R\right)
\end{aligned}$$

completing the proof of the lemma. □

5.2.3 Large Min-Entropy Implies Small Average Bias

We would like to then show that assuming $M = m, B, O_d$, the "average" bias over chosen coordinates are small using Lemma 4.9.

Suppose we take the set $\varsigma_i[d]$ as those induced by selecting one coordinate per row r, but restricted to be in j_r block of the input. Then there are K^R possible choices of such $\varsigma_i[d]$. Consider a uniform distribution over such $\varsigma_i[d]$ as $\mathcal{U}_{j_1,\ldots,j_R}$. A simple observation here is that the original uniform distribution $\mathcal{U}$ over $\varsigma_i[d]$ (i.e. uniformly selecting one entry per row at random) can be decomposed as

$$\mathbb{E}_{j_1,\ldots,j_R \in \mathcal{U}\left[\frac{n_i}{DRK}\right]} \left[\mathcal{U}_{j_1,\ldots,j_R}\right] = \mathcal{U}$$

Now suppose we consider $(j_1, \ldots, j_R)$ that satisfies Lemma 5.5. Denote such $(j_1, \ldots, j_R)$ as $\mathcal{J}_d$. That is

$$\mathcal{J}_d := \left\{ (j_1, \ldots, j_R) | \widetilde{H}_\infty \left(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R | T_i^{(<d)}, M = m, B, O_d \right) \geq (1-\gamma)KR \right\} \tag{29}$$

We can show the following lemma for $(j_1, \ldots, j_R) \in \mathcal{J}_d$ using Lemma 4.9.

Lemma 5.7. *Assume the premise of Lemma 5.5. Suppose $(j_1, \ldots, j_R) \in \mathcal{J}_d$. Then*

$$\mathbb{E}_{\varsigma_i[d] \sim \mathcal{U}_{j_1,\ldots,j_R}} \left[\mathbb{E} \left[\chi_S(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R) | M = m, B, O_d \right]^2 \right] \leq 2^{R \log(4\gamma)}$$

Proof. By the property guaranteed by Theorem 4.7, Proposition 5.3, we know that

$$\forall t, t' \in \mathsf{supp}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R |_{M=m,B,O_d}), \quad \frac{\Pr\left[\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R = t | M = m, B, O_d \right]}{\Pr\left[\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R = t' | M = m, B, O_d \right]} \in 1 \pm O(Rn^{-1}) \tag{30}$$

Thus, if we simply consider the function $g : \left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R \to \mathbb{R}^{\geq 0}$ as

$$g(x) := \Pr\left[\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R = x | M = m, B, O_d \right] \cdot \left| \mathsf{supp}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R |_{M=m,B,O_d}) \right|,$$

as (30) implies that

$$\Pr\left[\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R = x | M = m, B, O_d \right] \leq \frac{1 + O(Rn^{-1})}{\left| \mathsf{supp}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R |_{M=m,B,O_d}) \right|}$$

we have the guarantee that for any x in the support of g,

$$g(x) \leq 1 + O(Rn^{-1}).$$

Then Lemma 4.9 implies that for any $\delta > 0$,

$$\sum_{S \subset [RK]} \delta^{|S|} \cdot \widetilde{g}(S)^2 \leq (1 + O(Rn^{-1}))^2 \left(\frac{2^{RK}}{\left| \mathsf{supp}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^R |_{M=m,B,O_d}) \right|} \right)^{2\delta}. \tag{31}$$

Observe that $\widetilde{g}(S)$ term is exactly

$$\widetilde{g}(S) = \frac{1}{2^{RK}} \sum_{x \in \{0,1\}^{RK}} g(x) \cdot \chi_S(x) \frac{2^{RK}}{\left| \mathsf{supp}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R} |_{M=m,B,O_d}) \right|}$$
$$= \mathbb{E}\left[\chi_S(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R}) | M = m, B, O_d \right],$$

while we can bound

$$\frac{2^{RK}}{\left| \mathsf{supp}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R} |_{M=m,B,O_d}) \right|} \leq 2^{(RK - H_\infty(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R} | M=m,B,O_d))}$$

due to Claim 4.6. As $(j_1, \ldots, j_R) \in \mathcal{J}_d$, and due to Claim 4.5, we can rewrite (31) as

$$\sum_{S \subset [RK]} \delta^{|S|} \mathbb{E}\left[\chi_S(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R}) | M = m, B, O_d \right]^2$$
$$\leq (1 + O(Rn^{-1}))^2 2^{2\delta \cdot \gamma RK}$$

Select δ as $\frac{1}{2\gamma K}$. Then by considering $\varsigma_i[d] \sim \mathcal{U}_{j_1,\ldots,j_R}$,

$$\mathbb{E}_{\varsigma_i[d] \sim \mathcal{U}_{j_1,\ldots,j_R}} \left[\mathbb{E}\left[\chi_{\varsigma_i[d]}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R}) | M = m, B, O_d \right]^2 \right] \leq 2^{1 + R + R\log(2\gamma K) - R\log K} = 2^{1 + R\log(4\gamma)}.$$

□

A simple Markov's inequality implies the following corollary, which we will use towards our main proof.

Corollary 5.8. *Assume the premise of Lemma 5.5 and $(j_1, \ldots, j_R) \in \mathcal{J}_d$. Then*

$$\Pr_{\varsigma_i[d] \sim \mathcal{U}_{j_1,\ldots,j_R}} \left[|2\Pr[O_{d+1} | M = m, B, O_d] - 1| \geq 2^{\frac{R\log(8\gamma)}{2}} \right] \leq 2^{1-2R}$$

Proof. Observe that

$$|2\Pr[O_{d+1} | M = m, B, O_d] - 1| = \left| \mathbb{E}\left[\chi_{\varsigma_i[d]}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R}) | M = m, B, O_d \right] \right|$$

and that Lemma 5.7 along with Markov's inequality implies

$$\Pr_{\varsigma_i[d] \sim \mathcal{U}_{j_1,\ldots,j_R}} \left[\mathbb{E}\left[\chi_{\varsigma_i[d]}(\left\{ T_i^{(d,r,j_r)} \right\}_{r=1}^{R}) | M = m, B, O_d \right]^2 \geq 2^{R\log(8\gamma)} \right] \leq \frac{2^{1+R\log(4\gamma)}}{2^{R\log(8\gamma)}} = 2^{1-2R}$$

□

Taking γ to be a sufficiently small constant, while taking large enough $R = \Theta(\log n)$, we obtain the bounds for Lemma 5.5 and Corollary 5.8 with

$$\exp(-1.5\gamma^2 R) \leq n^{-100}, \quad 2^{\frac{R\log(8\gamma)}{2}} \leq o(n^{-2}), \quad 2^{1-2R} \leq n^{-100}.$$

5.2.4 Combining All

Applying Lemma 5.5 and Lemma 5.7 iteratively leads to the following lemma.

Lemma 5.9.

$$\Pr_{\varsigma_i[1],\ldots,\varsigma_i[D]\sim\mathcal{U}^D}\left[\Pr_{T_i|M=m,B}[O_D|M=m,B]\notin\left[\left(\frac{1-n^{-2}}{2}\right)^D,\left(\frac{1+n^{-2}}{2}\right)^D\right]\right]\leq n^{-99}$$

Proof. Let

$$\mathcal{F}_d:=\left\{\varsigma_i[d]:\Pr_{T_i|M=m,B}[O_{d+1}|M=m,B,O_d]\in\left[\left(\frac{1-n^{-2}}{2}\right),\left(\frac{1+n^{-2}}{2}\right)\right]\right\}$$

If we choose $\varsigma_i[1]\in\mathcal{F}_1,\ldots,\varsigma_i[d-1]\in\mathcal{F}_{d-1}$,

$$\Pr_{T_i|M=m,B}[O_d|M=m,B]\in\left[\left(\frac{1-n^{-2}}{2}\right)^{d-1},\left(\frac{1+n^{-2}}{2}\right)^{d-1}\right].$$

Then for such choice of $\varsigma_i[1],\ldots,\varsigma_i[d-1]$, Lemma 5.5 and Corollary 5.8 implies that

$$1-\Pr_{\varsigma_i[d]}[\mathcal{F}_d]\leq 2n^{-100}$$

and furthermore, choosing $\varsigma_i[d]$ from $\mathcal{F}_d$ would further give

$$\Pr_{T_i|M=m,B}[O_{d+1}|M=m,B]\in\left[\left(\frac{1-n^{-2}}{2}\right)^{d},\left(\frac{1+n^{-2}}{2}\right)^{d}\right].$$

Iteratively applying the argument, if $\forall d\in[D]$, $\varsigma_i[d]\in\mathcal{F}_d$,

$$\Pr_{T_i|M=m,B}[O_D|M=m,B]\in\left[\left(\frac{1-n^{-2}}{2}\right)^{D},\left(\frac{1+n^{-2}}{2}\right)^{D}\right].$$

As each $\varsigma_i[d]$'s are chosen independently at random, the probability of

$$\Pr_{\varsigma_i[1],\ldots,\varsigma_i[D]}[\forall d\in[D],\varsigma_i[d]\in\mathcal{F}_d]\geq 1-2Dn^{-100}\geq 1-n^{-99}$$

as D is some linear factor of $\log n$, which completes the proof of the lemma. □

Now we are ready to complete the proof of Theorem 5.2 with Lemma 5.9.

Proof of Theorem 5.2.

We would like to show that Ψ_i induced by f_i and underlying f satisfies Definition 1.2. For any setting of ς_i's over $i\in[\ell]$,

$$\Pr_T[\mathsf{OR}(f_1(T_1),\ldots f_{i-1}(T_{i-1}),f_{i+1}(T_{i+1}),f_\ell(T_\ell))=+1]=1-\left(1-2^{-D}\right)^\ell.$$

The choice of D was ensured to guarantee that the above quantity is $\Omega(1)$, that is with probability $g_i\geq\Omega(1)$, $f_i(T_i)$ matters. Furthermore, for any setting of ς_i,

$$\Pr_{T_i}[f_i(T_i)=+1]=2^{-D}=\Theta(\ell^{-1}),$$

making the function balanced with the balanced parameter $\beta = \Theta(\ell^{-1})$. This is good enough as $\beta \geq 2^{-o(\sqrt{\log n})}$. Therefore, our f satisfies the first condition of Definition 1.2.

Now observe that due to normalization

$$\Psi_i(\varsigma_i, T_i) = \begin{cases} +1 & \text{if } f_i(\varsigma_i, T_i) = +1 \\ -\frac{1}{1-2^{-D}} + 1 = -\frac{2^{-D}}{1-2^{-D}} & \text{otherwise} \end{cases}$$

We would like to show that the above matrix Ψ_i has small odisc for $|M| = c \leq n/\mathrm{poly}\log(n)$ where

$$\mathsf{odisc}_{T_i|_{M=m}}(\Psi_i) := \mathbb{E}_{\varsigma_i}\left|\sum_{T_i} \Psi_i(\varsigma_i, T_i) \cdot \Pr[T_i|M=m]\right|$$
$$\mathsf{odisc}_c(\Psi_i) := \max_{|M|\leq c} \mathbb{E}_M\left[\mathsf{odisc}_{T_i|_{M=m}}(\Psi_i)\right]$$

Recall that for any M, we can create a further partitioning $\mathcal{B}$, which would then

$$\mathbb{E}_{B\in\mathcal{B}}\left[\mathsf{odisc}_{T_i|_{M=m,B}}(\Psi_i)\right] = \mathbb{E}_{B\in\mathcal{B}}\left[\mathbb{E}_{\varsigma_i}\left|\sum_{T_i} \Psi_i(\varsigma_i, T_i) \cdot \Pr[T_i|M=m,B]\right|\right]$$
$$\geq \mathbb{E}_{\varsigma_i}\left|\mathbb{E}_{B\in\mathcal{B}}\left[\sum_{T_i} \Psi_i(\varsigma_i, T_i) \cdot \Pr[T_i|M=m,B]\right]\right| \geq \mathsf{odisc}_{T_i|_{M=m}}(\Psi_i) - \exp(n^{-\Omega(1)})$$

where the last bound follows from

$$\sum_{T_i} |\mathbb{E}_{B\in\mathcal{B}}\left[\Pr[T_i|M=m,B]\right] - \Pr[T_i|M=m]| \leq \exp(n^{-\Omega(1)})$$

due to the property of $\mathcal{B}$.

Our goal is then to bound $\mathbb{E}_M\left[\mathbb{E}_{B\in\mathcal{B}}\left[\mathsf{odisc}_{T_i|_{M=m,B}}(\Psi_i)\right]\right]$ as

$$\mathbb{E}_M\left[\mathsf{odisc}_{T_i|_{M=m}}(\Psi_i)\right] \leq \mathbb{E}_M\left[\mathbb{E}_{B\in\mathcal{B}}\left[\mathsf{odisc}_{T_i|_{M=m,B}}(\Psi_i)\right]\right] + \exp(n^{-\Omega(1)})$$

We consider $\mathsf{odisc}_{T_i|_{M=m,B}}(\Psi_i)$ for some fixed $M = m, B$. Suppose $M = m, B$ satisfies

$$\forall d \in [D], \quad \widetilde{H}_\infty\left(T_i^{(d)} | T_i^{(<d)}, M=m, B\right) \leq \frac{n_i}{D} - 2(|M| + |\mathcal{B}|). \tag{32}$$

Lemma 5.9 implies that for such choice of $M = m, B$,

$$\Pr_{\varsigma_i[1],\ldots,\varsigma_i[D]}\left[\Pr_{T_i|M=m,B}[O_D|M=m,B] \notin \left[\left(\frac{1-n^{-2}}{2}\right)^D, \left(\frac{1+n^{-2}}{2}\right)^D\right]\right] \leq n^{-99} \tag{33}$$

As per choice of $\varsigma_i[1], \ldots, \varsigma_i[D]$, the advantage can be written as

$$\left|\Pr_{T_i|M=m,B}[O_D|M=m,B] - \frac{2^{-D}}{1-2^{-D}}(1 - \Pr_{T_i|M=m,B}[O_D|M=m,B])\right|$$
$$= \frac{1}{1-2^{-D}}\left|\Pr_{T_i|M=m,B}[O_D|M=m,B] - 2^{-D}\right| \tag{34}$$

(33) implies that with probability all but n^{-99} over $\varsigma_i[1], \ldots, \varsigma_i[D]$, (34) is then at most

$$(34) \leq \frac{2^{-D}}{1-2^{-D}} \left(\left(1+n^{-2}\right)^D - 1 \right) \leq o(n^{-2})$$

which then concludes that if $M = m, B$ satisfies (32),

$$\mathsf{odisc}_{T_i|_{M=m,B}}(\Psi_i) \leq o(n^{-2})$$

Lemma 5.4 then implies that the fraction of $M = m, B$ that does not satisfy (32) is at most $2^{-n^{1/3}}$. Therefore,

$$\mathbb{E}_M \left[\mathsf{odisc}_{T_i|_{M=m}}(\Psi_i) \right] \leq \mathbb{E}_M \left[\mathbb{E}_{B \in \mathcal{B}} \left[\mathsf{odisc}_{T_i|_{M=m,B}}(\Psi_i) \right] \right] + \exp(n^{-\Omega(1)}) \leq o(n^{-2})$$

□

References

[BL15] Joshua Brody and Kasper Gren Larsen. Adapt or Die: Polynomial Lower Bounds for Non-Adaptive Dynamic Data Structures. *THEORY OF COMPUTING*, 11:19, 2015.

[BT22] Mark Bun and Justin Thaler. *Approximate degree in classical and quantum computing.* Number volume 15, issue 3-4 (2022) in Foundations and trends® in theoretical computer science. now, Boston Delft, 2022.

[CEEP12] Arkadev Chattopadhyay, Jeff Edmonds, Faith Ellen, and Toniann Pitassi. A Little Advice Can Be Very Helpful. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 615–625. Society for Industrial and Applied Mathematics, January 2012.

[CGL15] Raphael Clifford, A. Grønlund, and K. G. Larsen. New Unconditional Hardness Results for Dynamic and Online Problems. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1089–1107, October 2015.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing).* Wiley-Interscience, New York, NY, USA, 2006.

[DL20] Pavel Dvořák and Bruno Loff. Lower Bounds for Semi-adaptive Data Structures via Corruption. *LIPIcs, Volume 182, FSTTCS 2020*, 182:20:1–20:15, 2020. Artwork Size: 15 pages, 547135 bytes ISBN: 9783959771740 Medium: application/pdf Publisher: Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[DORS08] Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. *SIAM Journal on Computing*, 38(1):97–139, January 2008.

[FS89] Michael L. Fredman and Michael E. Saks. The Cell Probe Complexity of Dynamic Data Structures. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, pages 345–354, 1989.

[JS10] S. Jukna and G. Schnitger. Circuits with arbitrary gates for random operators, 2010. Version Number: 1.

[KNR95] Ilan Kremer, Noam Nisan, and Dana Ron. On randomized one-round communication complexity. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing - STOC '95*, pages 596–605, Las Vegas, Nevada, United States, 1995. ACM Press.

[Ko25] Young Kun Ko. Lower Bounds for Linear Operators, 2025. Version Number: 1.

[KW20] Young Kun Ko and Omri Weinstein. An Adaptive Step Toward the Multiphase Conjecture. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 752–761, Durham, NC, USA, November 2020. IEEE.

[Lar12] Kasper Green Larsen. The Cell Probe Complexity of Dynamic Range Counting. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 85–94, New York, NY, USA, 2012. ACM. event-place: New York, New York, USA.

[LWY20] Kasper Green Larsen, Omri Weinstein, and Huacheng Yu. Crossing the Logarithmic Barrier for Dynamic Boolean Data Structure Lower Bounds. *SIAM Journal on Computing*, 49(5):STOC18–323–STOC18–367, January 2020.

[LY25] Kasper Green Larsen and Huacheng Yu. Super-Logarithmic Lower Bounds for Dynamic Graph Problems. *SIAM Journal on Computing*, pages FOCS23–42–FOCS23–69, February 2025.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 1 edition, June 2014.

[Pat10] Mihai Patrascu. Towards polynomial lower bounds for dynamic problems. In Leonard J. Schulman, editor, *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 603–610. ACM, 2010.

[PD04] Mihai Pătraşcu and Erik D. Demaine. Lower bounds for dynamic connectivity. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 546–553, 2004.

[PD06] Mihai Patrascu and Erik D. Demaine. Logarithmic Lower Bounds in the Cell-Probe Model. *SIAM J. Comput.*, 35(4):932–963, April 2006.

[PTW10] Rina Panigrahy, Kunal Talwar, and Udi Wieder. Lower Bounds on Near Neighbor Search via Metric Expansion. In *FOCS 2010*, pages 805–814, 2010.

[Pă11] Mihai Pătraşcu. Unifying the Landscape of Cell-Probe Lower Bounds. *SIAM J. Comput.*, 40(3):827–847, 2011.

[Rou16] Tim Roughgarden. Communication Complexity (for Algorithm Designers). *Foundations and Trends® in Theoretical Computer Science*, 11(3-4):217–404, 2016.

[RY18] Anup Rao and Amir Yehudayoff. *Communication Complexity and Applications*. 2018.

[She11] Alexander A. Sherstov. The Pattern Matrix Method. *SIAM Journal on Computing*, 40(6):1969–2000, January 2011.

[Sie04] Alan Siegel. On Universal Classes of Extremely Random Constant-Time Hash Functions. *SIAM J. Comput.*, 33(3):505–543, 2004.

[Sko19] Maciej Skorski. Strong Chain Rules for Min-Entropy under Few Bits Spoiled. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 1122–1126, Paris, France, July 2019. IEEE.

[Tho13] Mikkel Thorup. Mihai PăTraşCu: Obituary and Open Problems. *SIGACT News*, 44(1):110–114, March 2013.

[Wat20] Thomas Watson. Communication Complexity with Small Advantage. *computational complexity*, 29(1):2, June 2020.

[Yao81] Andrew Chi-Chih Yao. Should Tables Be Sorted? *J. ACM*, 28(3):615–628, July 1981.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il