2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)

An Adaptive Step Toward the Multiphase Conjecture

Young Kun Ko
NYU Courant
New York, USA

Omri Weinstein
Columbia University
New York, USA

***Abstract*—In 2010, Pătraşcu proposed a dynamic set-disjointness problem, known as the *Multiphase problem*, as a candidate for proving *polynomial* lower bounds on the operational time of dynamic data structures. He conjectured that any data structure for the Multiphase problem must make n^ϵ cell-probes in either update or query phases, and showed that this would imply similar *unconditional* lower bounds on many important dynamic data structure problems. There has been almost no progress on this conjecture in the past decade since its introduction. We show an $\tilde{\Omega}(\sqrt{n})$ cell-probe lower bound on the Multiphase problem for data structures with general (adaptive) updates, and queries with unbounded but "layered" adaptivity. This result captures all known set-intersection data structures and significantly strengthens previous Multiphase lower bounds, which only captured non-adaptive data structures.**

Our main technical result is a communication lower bound on a 4-party variant of Pătraşcu's Number-On-Forehead Multiphase game, using information complexity techniques. We then use this result to make progress on understanding the power of *nonlinear* gates in networks computing *linear* operators, a long-standing open problem in circuit complexity and network design: We show that any depth-d circuit that computes a random $m \times n$ linear operator $x \mapsto Ax$ using gates of degree k (width-k DNFs) must have $\Omega(m \cdot n^{1/2(d+k)})$ wires. Finally, we show that a lower bound on Pătraşcu's original NOF game would imply a polynomial wire lower bound ($n^{1+\Omega(1/d)}$) for circuits with *arbitrary* gates computing a random linear operator. This suggests that the NOF conjecture is much stronger than its data structure counterpart.

1. Introduction

Proving unconditional lower bounds on the operational time of dynamic data structures has been a challenge since the introduction of the *cell-probe* model [Yao79]. In this model, the data structure needs to support a sequence of n online updates and queries, where the operational cost is measured only by the number of *memory accesses* ("probes") the data structure makes to its memory, whereas all computations on probed cells are completely free of charge. A natural question to study is the tradeoff between the *update* time t_u and *query* time t_q of the data structure for supporting the underlying dynamic problem. Cell-probe lower bounds provide a compelling answer to this question, as they are purely information-theoretic and independent of implementation constraints, hence apply to any imaginable data structure. Unfortunately, the abstraction of the cell-probe model also comes at a price, and the highest explicit lower bound known to date, on *any* dynamic problem, is merely polylogarithmic ($\max\{t_u, t_q\} \geq \tilde{\Omega}(\lg^2 n)$, see e.g., [Lar12], [LWY18] and references therein). In 2010, Pătraşcu [Pat10] proposed the following dynamic set-disjointness problem, known as the *Multiphase problem*, as a candidate for proving *polynomial* lower bounds on the operational time of dynamic data structures. The problem proceeds in 3 "phases":

- **PI**: Preprocess a collection of $m = \mathrm{poly}(n)$ sets $\vec{S} = S_1, \dots, S_m \subseteq [n]$ in $O(mnt_u)$ time.
- **PII**: A set $T \subseteq [n]$ is revealed, and the data structure can update its memory in $O(nt_u)$ time.
- **PIII**: Given any $i \in [m]$, the data structure must determine if $S_i \cap T \stackrel{?}{=} \emptyset$ in $O(t_q)$-time.

Pătraşcu conjectured that any data structure solving the Multiphase problem must make $\max\{t_u, t_q\} \geq n^\varepsilon$ cell-probes, and showed that such a polynomial lower bound would imply similar polynomial lower bounds on many important dynamic data structure problems, including dynamic reachability in directed graphs and online matrix multiplication (for the broad implications and further context of the Multiphase conjecture within fine-grained complexity, see [Pat10], [HKNS15]). In the same paper, Pătraşcu [Pat10] proposed an approach to prove an unconditional cell-probe lower bound on the Multiphase problem, by reduction to the following 3-party *Number-On-Forehead* (NOF) communication game $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$, henceforth called the *Multiphase Game* :

- Alice receives a collection of sets $\vec{S} = S_1, \dots, S_m \subseteq [n]$ and a random index $i \in_R [m]$.
- Bob receives a set $T \subseteq [n]$ and the index i.
- Charlie receives $\vec{S}$ and T (but not i).

Thus, one can think of i as being on Charlie's fore-

2575-8454/20/$31.00 ©2020 IEEE
DOI 10.1109/FOCS46700.2020.00075

head, T being on Alice's forehead, and $\vec{S}$ being on Bob's forehead. The goal of the players is to determine if $S_i \cap T \stackrel{?}{=} \emptyset$, where communication proceeds in the following way: First, Charlie sends a message ("advice") $U = U(\vec{S}, T)$ *privately to Bob*. Thereafter, Alice and Bob continue to compute $\mathsf{DISJ}_n(S_i, T)$ in the standard 2-party model. Denoting by Π the second stage protocol, Pătraşcu made the following conjecture:

Conjecture 1.1 (Multiphase Game Conjecture, Conjecture 9 of [Pat10])**.** *Any 3-party NOF protocol for the Multiphase game with* $|U| = o(m)$ *bits of advice must have* $|\Pi| > n^{\varepsilon}$ *communication.*

The (naïve) intuition for this conjecture is that, since Charlie's advice is independent of i, it can only provide very little useful information about the interesting subproblem $\mathsf{DISJ}_n(S_i, T)$ (assuming $|U| = o(m)$), and hence Alice and Bob might as well solve the problem in the standard 2-party model. This intuition turns out to be misleading, and in fact when S_i's and T are *correlated*, it is simply false – Chattopadhyay, Edmonds, Ellen and Pitassi [CEEP12] showed a deterministic (2-round) NOF protocol for the Multiphase game with a total of $O(\sqrt{n}\log m) = \tilde{O}(\sqrt{n})$ communication, whereas the 2-party communication complexity of set-disjointness is $\Omega(n)$, even randomized. Surprisingly, they also show that Conjecture 1.1 is equivalent, up to $O(\log m)$ communication factor, for deterministic and randomized protocols. Nevertheless, the intuition behind the Conjecture 1.1 still stands for *product distributions* [CEEP12] (incidentally, the 2-party communication complexity of set-disjointness under product distributions is $\tilde{\Theta}(\sqrt{n})$ [BFS86], [HW07]).

The technical centerpiece of this paper is an $\Omega(\sqrt{n})$ lower bound on the NOF Multiphase game, for (unbounded-round) protocols in which *only the first* message of Alice in Π depends on her entire input $\vec{S} = S_1, \ldots, S_m$ (and i), while in subsequent rounds $j > 1$, Alice's messages can depend only *on* S_i*,* i *and the transcript* $\Pi^{<j}$ (No restriction is placed on Bob and Charlie). Note that Alice's messages in subsequent rounds still heavily depend on *all sets* $\vec{S}$, but only through the *transcript* of $\Pi = \Pi(\vec{S})$ so far (this feature better captures data structures, which can only adapt based on cells probed so far). There is a natural way to view such restricted 3-party NOF protocol in terms of an additional player (Megan, holding $\vec{S}, i$), who, in addition to Charlie's private advice to Bob, can broadcast a single message to *both* Alice and Bob (holding S_i, T respectively), who then continue to communicate in the standard 2-party model (see Figure 1). We define this *4-party NOF* model formally in Section 1.1.

Our main technical result is the following tight lower bound on such NOF protocols:

Theorem 1.2 (Informal)**.** *Let* $m = \omega(n)$*. Any Restricted NOF protocol* $\Gamma = (U, \Pi)$ *for the Multiphase game with* $|U| = o(m)$ *bits of advice must have* $|\Pi| > \Omega(\sqrt{n})$ *communication.*

This lower bound is tight up to logarithmic factors, as the model generalizes the upper bound of [CEEP12] (See full version). This suggests that the NOF model we study is both subtle and powerful. Indeed, while the aforementioned restriction may seem somewhat technical, we show that removing it by allowing as little as *two rounds* of Alice's messages to depend on her entire input $\vec{S}$, would lead to a major breakthrough in circuit lower bounds – see Theorem 1.7 below. Interestingly, the Multiphase conjecture itself does not have this implication, since dynamic data structures only have limited and local access to $\vec{S}$, through the probes ("transcript") of the query algorithm, and hence induce weaker NOF protocols.

1.1. NOF Communication Models

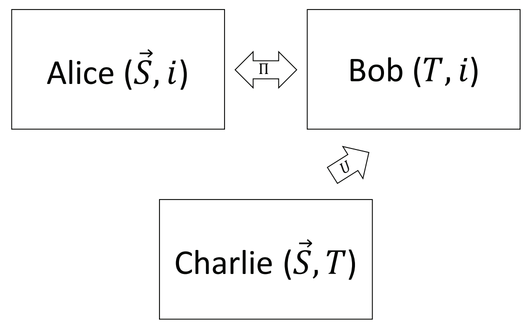

Figure 1. 3-party NOF Communication

In Pătraşcu's NOF Multiphase Game SEL_f^m, there are 3 players with the following information on their foreheads: **Charlie:** an index $i \in [m]$; **Bob:** a collection of sets $\vec{S} := S_1, \ldots, S_m \subseteq [n]$; **Alice:** a set $T \subseteq [n]$. I.e., Charlie has access to both $\vec{S}$ and T, but not to i. Alice has access to $\vec{S}$ and i, and Bob has access to T and i. The goal is to compute

$$\mathsf{SEL}_f^m(\vec{S}, T, i) := f(S_i, T).$$

The communication proceeds as follows: In the first stage of the game, Charlie sends a message ("*advice*") $U = U(\vec{S}, T)$ *privately to Bob*. In the second stage, Alice and Bob continue to communicate in the standard 2-party settting to compute $f(S_i, T)$ (see Figure 1). We denote such protocol by $\Gamma := (U, \Pi_i)$ where Π_i is the second stage transcript, assuming the index of the interesting subproblem is i.

Unfortunately, lower bounds for general protocols in Pătraşcu's 3-party NOF model seem beyond the reach of current techniques, as we show in Section 3 that Conjecture 1.1, even for *3-round* protocols, would resolve a major open problem in circuit complexity. Fortunately, for dynamic data structure applications, weaker versions of the NOF model suffice (this is indeed one of the main messages of this paper).

We consider the following restricted class of protocols. We say that $\Gamma = (U, \Pi)$ is a **restricted** NOF protocol if Alice is the first speaker in Π (in the second stage of the game) and only her *first* message Π_i^1 to Bob depends on her entire input $\vec{S}$ and i, whereas in subsequent rounds, Alice's messages Π^τ may depend only on S_i, i and the history of the transcript $\Pi^{<\tau}$ with Bob. Note that the latter means that Alice and Bob's subsequent messages can still heavily depend on $S_{-i} := \{S_j | j \neq i\}$, but only through the transcript (this feature better captures data structures, since the query algorithm can only adapt based on the information in cells it already probed, and not the entire memory).

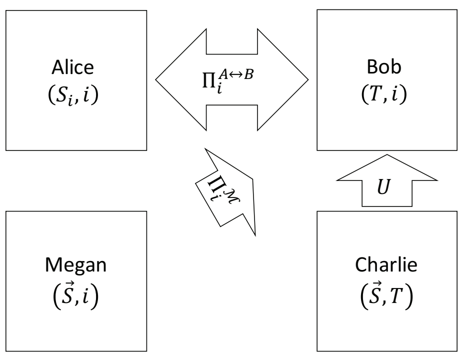

Figure 2. 4-party NOF Communication

An equivalent 4-party NOF model. Restricted 3-party NOF protocols are more naturally described by the following 4-party NOF model. **Alice** has access only to S_i and i, **Bob** has access to T and i. **Charlie** has access to $\vec{S}$ and T, but no access to i. **Megan** has access to $\vec{S}$ and i. In the first stage of the protocol Γ, in addition to Charlie's *private* message to Bob $U(\vec{S}, T)$, Megan can *broadcast* a message $\Pi^{\mathcal{M}} = \Pi^{\mathcal{M}}(\vec{S}, i)$ to both Alice and Bob. Thereafter, Alice and Bob proceed to communicate in the 2-party model as before, denoted $\Pi^{A\leftrightarrow B}$. See Figure 2. We denote 4-party protocols by $\Gamma = (U, \Pi)$ where $\Pi := (\Pi^{\mathcal{M}}, \Pi^{A\leftrightarrow B})$. We write $\Pi_i := (\Pi_i^{\mathcal{M}}, \Pi_i^{A\leftrightarrow B})$ to denote the transcript of Π when the index of the interesting subproblem is $i \in [m]$.

It is straightforward to see that Restricted 3-party NOF protocols for the Multiphase Game are equivalent to (unrestricted) 4-party protocols (by setting Alice's first message as Megan's message $\Pi_i^{\mathcal{M}}$, and Charlie remains unchanged). As such, our main technical theorem (Theorem 1.2) can be restated as follows.

Theorem 1.3 (4-party NOF Lower Bound)**.** *Let $m > \omega(n)$. For any 4-party NOF protocol $\Gamma = (U, \Pi)$ that solves $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$ with $|U| < o(m)$, there exists $i \in [m]$ such that $|\Pi_i| > \Omega(\sqrt{n})$.*

1.2. Implications to dynamic data structure lower bounds

In contrast to the *static* cell-probe model, *adaptivity* plays a dramatic role when it comes to dynamic data structures. In [BL15], Brody and Larsen consider a variation of the Multiphase problem with $(\lg n)$-bit updates (i.e., the 2nd phase set is of cardinality $|T| = 1$), and show that any dynamic data structure whose query algorithm is *non-adaptive*[1] must make $\max\{t_u, t_q\} \geq \Omega(n/w)$ cell-probes when the word-size is w bits. Nevertheless, such small-update problems have a trivial ($t_q = O(1)$) adaptive upper bound and therefore are less compelling from the prospect of making progress on general lower bounds. (By contrast, proving polynomial cell-probe lower bounds for dynamic problems with *large* $\mathrm{poly}(n)$-bit updates, like Multiphase, even against non-adaptive query algorithms, already seems beyond the reach of current techniques[2]).

We prove a polynomial lower bound on the Multiphase problem, against a much stronger class of data structures, which we call *semi-adaptive*, defined as follows:

Definition 1.4 (Semi-Adaptive Data Structures)**.** *Let D be a dynamic data structure for the Multiphase problem with general (adaptive) updates. Let $\mathcal{M}(\vec{S})$ denote the memory state of D after the preprocessing Phase I, and let $\Delta(\mathcal{M}, T)$ denote the set of ($\leq |T| \cdot t_u$) cells updated in Phase II. D is* semi-adaptive *if its query algorithm in Phase III operates in the following stages ("layers"):*

- *Given the query $i \in [m]$, D may first read S_i* free of charge.
- *D (adaptively) reads at most t_1 cells from $\mathcal{M}$.*
- *D (adaptively) reads at most t_2 cells from $\Delta(\mathcal{M}, T)$, and returns the answer $S_i \cap T \stackrel{?}{=} \emptyset$.*

The update time of D is t_u, and the query time is $t_q := t_1 + t_2$.

Thus, the query algorithm has unbounded but "layered" adaptivity in Phase III, in the sense that the model allows only a single alternation between the two layers of memory cells, $\mathcal{M}$ and $\Delta(\mathcal{M}, T)$. While this restriction may seem somewhat technical, all known set-intersection data structures are special cases of the semi-adaptive model (see [DLOM00], [BK02], [BY04], [BPP07], [CP10], [KPP15] and references therein). But more importantly, in this model there is actually a

[1]An algorithm is non-adaptive, if the addresses of probed memory cells are predetermined by the query itself, and do not depend on content of cells probed along the way.

[2]While intuitively larger updates $|T| = \mathrm{poly}(n)$ only make the problem harder and should therefore only be easier to prove lower bounds against, the *total* update time of the data structure in Phase II is also proportional to $|T|$ and hence the data structure has potentially much more power as it can "amortize" its operations. This is why encoding-style arguments fail for large updates (enumerating all $\binom{n}{|T|} = \exp(n)$ possible updates is prohibitive).

nontrivial *upper bound* for the Multiphase problem – A semi-adaptive data structure solving it in $t_u = t_q = O(\sqrt{n} \lg m)$ time (based on [CEEP12]'s communication protocol, see full version), indicating that the model is powerful. We remark that even though the set of modified cells $\Delta(\mathcal{M}, T)$ may be *unknown* to D at query time, it is easy to implement a semi-adaptive data structure by maintaining $\Delta(\mathcal{M}, T)$ in a dynamic dictionary [MPP05] that checks membership of cells in Δ, and returns $\perp$ if the cell is $\notin \Delta(\mathcal{M}, T)$.

Our main result is an essentially tight $\tilde{\Omega}(\sqrt{n})$ cell-probe lower bound on the Multiphase problem against semi-adaptive data structures. This follows from Theorem 1.2 by a simple variation of the reduction in [Pat10]:

Theorem 1.5 (Multiphase Lower Bound for Semi-Adaptive Data Structures)**.** *Let $m > \omega(n)$. Any semi-adaptive data structure that solves the Multiphase problem, must have either $n \cdot t_u \geq \Omega(m/w)$ or $t_q \geq \Omega(\sqrt{n}/w)$, in the dynamic cell-probe model with word size w.*

1.3. Implications to Network and Circuit Lower Bounds

Can *non-linear* gates help compute linear operators? A long-standing open problem in networks and circuit complexity is whether *non-linear* gates can help computing linear operators ([Lup56], [JS10], [Dru12]). Specifically, the challenge is to prove a polynomial lower bound on the number of *wires* of constant-depth circuits with *arbitrary gates* for computing any $m \times n$ linear operator $x \mapsto Ax$ [Juk12]. A random matrix $A \in \{0,1\}^{m \times n}$ easily gives a polynomial $\Omega(mn/\lg m)$ lower bound against *linear circuits* [Lup56], [JS10] (this restricted the interest to finding *explicit* hard matrices A, see [Val77]). In contrast, for general circuits, the highest lower bound on the number of wires, even for computing a *random* matrix A, is near-linear [Dru12], [GHK$^+$13]. Indeed, the current state of affairs cannot even rule out the possibility that nonlinear networks with $O(m \cdot \text{poly}\log n)$ wires suffice to compute **all** $m \times n$ linear operators [Dru12]. As such, a perplexing open question is whether one can prove the existence of a matrix A which requires a polynomial $m^{1+\varepsilon}$ number of wires for some constant $\varepsilon > 0$ when $m = \text{poly}(n)$.

Motivated by this question, we study an intermediate model of non-linear circuits, where each gate computes a *degree-k polynomial* on its input wires (more precisely, a width-k DNF) and may have unbounded fan-in and arbitrary depth. Note that even in this intermediate model, proving existential lower bounds against linear operators falls short of a counting argument: There are only 2^{n^2} possible $n \times n$ linear operators, but at least $\sim 2^{n^k}$ possible gates/functions of degree k (width-k DNFs) on n inputs[3], hence the counting argument breaks even for $k = 3$ (!).

We use Theorem 1.2 to prove that *most* linear operators require a polynomial number of wires, unless the network computing them is using highly nonlinear gates ($k = \omega(1)$). More formally:

Theorem 1.6. *There are linear operators $A \in \{0,1\}^{m \times n}$ such that any depth-d circuit with width-k DNF gates computing Ax must have $W \geq \Omega\left(m \cdot n^{\frac{1}{2(d+k)}}\right)$ wires.*

Finally, building on a recent reduction of Viola [Vio18], we show that Pătraşcu's NOF Conjecture 1.1, even for *3-round protocols*, would prove a polynomial wire lower bound against networks with *arbitrary* gates (i.e., $k = n$), resolving this longstanding open question. This indicates that Conjecture 1.1 may be much stronger than the Multiphase conjecture itself.

Theorem 1.7 (NOF Game Implies Circuit Lower Bounds)**.** *Suppose Conjecture 1.1 holds, even for 3-round protocols. Then for $m = \omega(n)$, there exists a linear* [4] *operator $A \in \{0,1\}^{m \times n}$ such that any depth-d circuit computing $x \mapsto Ax$ (with* arbitrary gates and unbounded fan-in*) requires $n^{1+\Omega(\varepsilon/d)}$ wires. In particular, if $d = 2$, the conjecture implies that computing Ax for some A requires $n^{1+\frac{\varepsilon}{2}-o(1)}$ wires.*

Comparison to previous work. The aforementioned work of Brody and Larsen [BL15] proves essentially optimal ($\Omega(n/w)$) dynamic lower bounds on variations of the Multiphase problem, when either the update or query algorithms are *nonadaptive* (or in fact "memoryless" in the former, which is an even stronger restriction). Proving lower bounds in the semi-adaptive model is a different ballgame, as the $\tilde{O}(\sqrt{n}/w)$ upper bound suggests [CEEP12]. We also remark that [BL15] were the first to observe a (similar but different) connection between nonadaptive data structures and depth-2 circuit lower bounds.

A more recent result of Clifford et. al [CGL15] proves a "threshold" cell-probe lower bound on *general* dynamic data structures solving the Multiphase problem, asserting that fast queries $t_q = o(\log m/\log n)$ require very high $t_u > m^{1-o(1)}$ update time. This result does not rule out data structures with $t_u = t_q = \text{poly}\log(n)$ time for the Multiphase problem (For general data structures, neither does ours).

As far as the Multiphase NOF Game, the aforementioned work of Chattopadhyay et. al [CEEP12] proves a tight $\tilde{\Theta}(\sqrt{n})$ communication lower bound against so-called *"1.5-round"* protocols, in which Bob's message

[3]Indeed, even though the *degree* of each gate is bounded (k), it may have *unbounded* fan-in.

[4]Over the boolean Semiring, i.e., where addition are replaced with $\vee$ and multiplication are replaced with $\wedge$. We note that there is evidence that computing Ax over the boolean Semiring is easier than over $\mathbb{F}_2$ [CGL15], hence in that sense our lower bound is stronger than over finite fields.

to Alice is *independent of the index* i, hence he is essentially "forwarding" a small ($o(n)$) portion of Charlie's message to Alice (this effectively eliminates Bob from communicating, making it similar to a 2-party problem). While our restricted NOF model is formally incomparable to [CEEP12] (as in our model, Alice is the first speaker), Theorem 1.2 in fact subsumes it by a simple modification (see full version). The model we study seems fundamentally stronger than 1.5-protocols, as it inherently involves multiparty NOF communication.

To best of our knowledge, all previous lower bounds ultimately reduce the Multiphase problem to a *2-party* communication game, which makes the problem more amenable to **compression-based** arguments. This is the main departure point of our work. We remark that most of our information-theoretic tools in fact apply to general dynamic data structures.

2. Technical Overview

Here we provide a streamlined overview of our main technical result, Theorem 1.2. For the full proof, we refer the reader to the full version of the paper.

As discussed earlier in the introduction, a naïve approach to the Multiphase Game $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$ is a "round elimination" approach: Since Charlie's advice consists of only $|U| = o(m)$ bits and he has no knowledge of the index of the interesting subproblem $i \in_R [m]$, his advice U to Bob should convey $o(1)$ bits of information about the interesting set S_i and hence Alice and Bob might hope to simply "ignore" his advice U and use such efficient NOF protocol Γ to generate a too-good-to-be-true *2-party* protocol for set disjointness (by somehow "guessing" Charlies's message which appears useless, and absorbing the error). The fundamental flaw with this intuition is that Charlie's advice is a function of *both* players' inputs, hence conditioning on $U(\vec{S}, T)$ *correlates* the inputs in an arbitrary way, extinguishing the standard "rectangular" (Markov) property of 2-party protocols in the second phase interaction $\Pi^{A\leftrightarrow B}$ between Alice and Bob (This is the notorious feature preventing "direct sum" arguments in NOF communication models). In particular, Chattopadhyay et. al [CEEP12] show that a small advice ($|U| = \tilde{O}(\sqrt{n})$) can already decrease the communication complexity of the multiphase problem to at most the 2-party complexity of set-disjointness under *product* distributions, yielding a surprising 2-round $\tilde{O}(\sqrt{n})$ upper bound on the Multiphase game. (This justifies why our hard distribution for $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$ will be a product distribution to begin with, i.e., $\vec{S} \perp T$). Alas, even if the inputs are originally independent ($I(\vec{S}; T) = 0$), they may *not* remain so throughout Π, and it is generally possible that $I(S_i; T|\Pi) \gg 0$. This means that, unlike 2-party protocols, $\Pi = \Pi(U, \vec{S}, T)$ *introduces correlation* between the inputs, and as such, is not amenable to the standard analysis of 2-party communication techniques.

Nevertheless, one might still hope that if the advice U is small enough, then this correlation will be small for an average index S_i when the inputs are independently chosen. At a high level, our proof indeed shows that if only the *first* message of Alice can (directly) depend on her entire input $\vec{S} = S_1, \ldots, S_m$ (whereas her subsequent messages Π^τ are only a function of S_i, i and the transcript history $[\Pi(\vec{S}, T, i)]^{<\tau}$), then it is possible to *simultaneously* control the information cost and correlation of Π, so long as the advice U is small enough ($o(m)$). This in turn facilitates a "robust" direct-sum style argument for *approximate protocols*. More formally, our proof consists of the following two main steps:

• **A low correlation and information random process for computing** AND**.** The first part of the proof shows that an efficient Restricted NOF protocol Γ for the Multiphase game $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$ (under the natural *product* distribution on $\vec{S}, T$) can be used to design a certain random process $Z(X, Y)$ computing the 2-bit AND function (on 2 independent bits $\sim \mathcal{B}_{\Theta(1/\sqrt{n})}$), which *simultaneously* has *low information cost* w.r.t X, Y and *small correlation*, meaning that the input bits remain roughly independent at any point during the process, i.e, $I(X; Y|Z) = o(1/n)$. Crucially, Z is *not* a valid 2-party protocol (Markov chain) – if this were the case, then we would have $I(X; Y|Z) = 0$ since a deterministic 2-party protocol cannot introduce any additional correlation between the original inputs (this is also the essence of the the celebrated *"Cut-and-Paste" Lemma* [BYJKS02]). Nevertheless, we show that for restricted NOF protocols Γ (equivalently, unrestricted protocols in our 4-party model, cf. Figure 2), it is possible to design such random variable $Z(X, Y)$ from Γ, which is *close enough* to a Markov chain. The design of Z requires a careful choice conditioning variables (to ensure that the correlation $\sim |U|/m$ doesn't accumulate over rounds) as well as a "coordinate sampling" step for reducing entropy, though the analysis of this part ultimately uses standard tools (the chain rule and subadditivity of mutual information).We first design a Z' with similar guarantees for single-copy *disjointness*, and then use (a variation of) the standard direct-sum information cost argument to "scale down" the information and correlation of Z so as to extract the desired random process for 2-bit AND. An important observation in this last step is that the direct sum property of information cost holds not just for communication protocols, but in fact for more general random variables.

• **A "robust" Cut-and-Paste Lemma.** The second part is proving that such random variable $Z(X, Y)$ cannot exist, i.e., ruling out a random process $Z(X, Y)$ computing AND (with 1-sided error under $X, Y \sim^{iid} \mathcal{B}_{\Theta(1/\sqrt{n})}$) which simultaneously has *low information* *cost* and *small correlation* ($o(1/n)$). The high-level intuition is that, if $Z(X, Y)$ introduces little correlation, then the distribution over X and Y conditioned on $Z(X, Y)$ should remain *approximately* a product distribution, i.e., close to a rectangle. By the correctness guarantee of Z, the distribution on $\{0,1\}^2$ conditioned on $Z(X, Y) = \mathsf{AND}(X, Y) = 0$ must have 0 mass on the $(1,1)$ entry. But if this conditional distribution does not contain $(1,1)$ in its support and *close to a rectangle*, a KL-divergence calculation shows that $Z(X, Y)$ must reveal a lot of information about either X or Y (this calculation crucially exploits the fact that Pinsker's inequality is loose "near the ends", i.e., $\mathsf{D}_{KL}(p||q) \approx \|p - q\|_1$ for $p, q = o(1)$, and there is no quadratic loss). Our argument can be viewed as a generalization of the Cut-and-Paste Lemma to more robust settings of random variables ("approximate protocols"). We remark that while the proof of the original Cut-and-Paste Lemma [BYJKS02] heavily relies on properties of the Hellinger distance, this technique does not seem to easily extend to small-correlation random variables. This forces us to find a more direct argument, which may be of independent interest.

Sketch of nonlinear network lower bound (Theorem 1.6) We prove Theorem 1.6 via reduction from our communication lower bound (Theorem 1.2). Let $A \in \{0,1\}^{m \times n}$ be the random matrix where every entry is $\mathcal{B}_{\Theta(1/\sqrt{n})}$ (i.e., Alice's input in Theorem 1.2), and respectively, let $x = T \in \{0,1\}^n$ be Bob's input in the Multiphase game. We claim that a cheap circuit C_A for computing Ax (over the boolean semiring), with only W wires (where gates compute width-k DNFs), implies an efficient *Restricted* NOF protocol for computing $(Ax)_i = \mathsf{DISJ}(A_i, x)$, which would violate Theorem 1.2. The key point in this reduction is using Charlie's advice to *kill the high fanin gates in* C_A: By a standard averaging argument, there can be $o(m)$ gates with fanin $\omega(W/m)$. Since in Theorem 1.2 Charlie is allowed to send $o(m)$ bits of advice and he sees *both* A and x (but not i), his advice to Bob will consist of the *outputs* of C_A on these $o(m)$ high fanin gates. Now, all the remaining gates of C_A have fanin $O(W/m)$, and since C_A has depth d, there can be at most $O((W/m)^d)$ such gates in the induced sub circuit whose root is the ith output gate (as this is a tree of hight d with branching-factor $O(W/m)$). Since each gate computes a degree-k polynomial but has *low fan-in*, Bob can afford to send the explicit function computed at this gate using $O((W/m)^k)$ bits. This induces a (1-round) Restricted protocol for the Multiphase game, with $|\Pi| = O((W/m)^{k+d})$ communication, after which Bob can compute the output $(Ax)_i$. By Theorem 1.2, $|\Pi| > \Omega(\sqrt{n})$, which gives the desired bound on the number of wires W.

3. Lower Bounds on Nonlinear Networks for computing Linear Operators

Circuits with arbitrary gates As mentioned in Section 1, a long-standing open problem in circuit complexity is whether *non-linear* gates can significantly (polynomially) reduce the number of wires of circuits computing linear operators [JS10]. We consider Valiant's depth-2 circuit model [Val77] with arbitrary gates, and its generalizations to arbitrary depths. More formally, consider a circuit computing a linear operator $x \mapsto Ax$ where A is an $m \times n$ matrix with $m = \mathrm{poly}(n)$, using unbounded fan-in, and where gates are allowed to be *arbitrary* functions. Clearly, such circuits can trivially compute any f with m gates. As such, the interesting complexity measure in this model is the minimum number of **wires** (W) to computing the function f. This measure captures how much "information" needs to be transferred between different components of the circuit, in order to compute the function. For a more thorough exposition and motivation on circuits with arbitrary gates, we refer the reader to [Juk12].

Previous Works In contrast to arithmetic circuit models (e.g., [Val77] where allowed functions are simple functions such as AND, OR or PARITY), it is a long-standing open problem [JS10], [Juk12], [Dru12] whether non-linear circuits can compute *any* linear operator A with near-linear ($\tilde{O}(m)$) number of wires. Indeed, for linear circuits, this is a simple counting argument. Counting argument shows that $\Omega(mn/\log(mn))$ wires are necessary for linear circuits, and this is tight [Lup56]. But once again, for arbitrary circuits, counting argument completely fails. The number of possible functions over n bits is already doubly exponential in n. In fact, counting argument fails even when we consider only width-3 DNFs. While there are only 2^{mn} different linear operators, there are at least 2^{n^3} different width-3 DNF gates.

[Juk10] initiated works on analyzing the complexity of *representing* a random matrix, that is computing Ax when x is restricted to having only one 1. In other words, compute Ae_i for $i \in [n]$. [Dru12] showed that when restricted to *representing* a matrix, $\Omega(m \log m)$ is necessary for depth 2 circuit, complementing the previous upper bound of $O(m \log m)$ by [Juk10].

The lower bound of [Dru12] immediately implies $\Omega(m \log m)$ lower bound for computing a matrix using depth 2 circuit, since any circuit that computes a matrix must represent a matrix as well. But no better bounds were known in case of **computing** the linear operator A. We refer the reader to the full version for the full summary of known results on wire bounds for circuits computing linear operator.

Our Result First, we show that our communication lower bound (i.e. Theorem 1.2) implies a trade-off between degree of gates, depth and number of wires required to compute a random linear operator (over the

Boolean semi-ring) using a variant of a reduction in [Vio18]. To the best of our knowledge, this is the first polynomial lower bound on the number of wires for any non-linear circuit model.

We then show that Conjecture 3.4 implies a polynomial lower bound on nonadaptive *static* data structures against a "semi-explicit" static problem—computing set-disjointness queries w.r.t n^2 *random* sets. Using the reduction of [Vio18] once again, we show that this static data structure lower bound implies a polynomial wire lower bound on depth-d circuits with *arbitrary* gates ($k = n$) for computing random linear operators. This is the content of the next two subsections.

3.1. Width-k DNF Lower Bound

First we show that if there exists a circuit with width-k DNF gates and small number of wires, then there exists a good 4-party communication protocol.

Lemma 3.1. *If there exists a depth-d circuit with width-k DNF gates with W-wires for computing Ax where $A \in \{0,1\}^{m\times n}$ with $m = \mathrm{poly}(n)$. Then there exists a 4-party communication protocol for computing $A_i x$ (for any given $i \in [m]$) with $|U| \le o(m)$, $|\Pi_i^{\mathcal{M}}| \le O\left((W/m)^{k+d}\right)$ and $|\Pi_i| \le O\left((W/m)^d \log n\right)$*

Proof. Suppose we have a depth-d circuit with width-k DNF gates with W wires. Then we argue that this induces an efficient 4-party communication protocol.

First set Megan's input as the linear operator A and index in question i, Bob's input as x and i, Charlie's input as A and x, Alice's input as A_i and i.

Charlie and Megan's message. We set Charlie and Megan's message in the following manner. Consider the set of gates G with fan-in $\omega(W/m)$. Since there are total of W wires, we know that $|G| < o(m)$. Also note that G has no dependence on i.

Set Megan's message $\Pi_i^{\mathcal{M}}$ as the description of circuit computing $A_i x$, without gates in G. Note that $\Pi_i^{\mathcal{M}}$ contains at most $O((W/m)^d)$ gates since each gate not in G has fan-in at most $O(W/m)$. Furthermore, description of each gate requires at most $O\left((W/m)^k + (W/m)\log(m + (W/m)^d)\right)$ bits, $O\left((W/m)^k\right)$ bits for describing the function and $O\left((W/m)\log(m + (W/m)^d)\right)$ bits for describing the inputs.

Therefore with $m = \mathrm{poly}(n)$, we get

$$
\begin{aligned}
|\Pi_i^{\mathcal{M}}| &\le O((W/m)^{k+d} \\
&\quad + (W/m)^{d+1}\log(m + (W/m)^d)) \\
&= O((W/m)^{k+d}).
\end{aligned}
$$

Furthermore we set Charlie's message U as the value of gates in G. Therefore we have $|U| \le o(m)$.

Alice and Bob's message. Alice queries Bob the gate values of x_i's and G required for computing $A_i x$ as given by $\Pi_i^{\mathcal{M}}$. Bob can answer Alice's query since Bob knows x and gate value of G from U. Since fan-in is bounded as $O(W/m)$ and it is a depth-d circuit, we know there are at most $O((W/m)^d)$ gate values required for computing $A_i x$. Therefore we have

$$
\begin{aligned}
|\Pi_i^{A\leftrightarrow B}| &\le O((W/m)^d \log(n + |G|)) \\
&= O((W/m)^d \log n)
\end{aligned}
$$

With Bob's message, Alice can compute $A_i x$ by using the circuit from $\Pi_i^{\mathcal{M}}$ with Bob's response as the input. □

Now our 4-party communication lower bound (Theorem 1.2) yields the following circuit lower bound via the reduction given in Lemma 3.1.

Theorem 3.2. *There exists a linear operator $A \in \{0,1\}^{m\times n}$ such that any depth-d circuit with width-k DNF gates computing Ax must have wire $W \ge \Omega\left(m \cdot n^{\frac{1}{2(d+k)}}\right)$.*

Proof. Suppose for all A, there exists a circuit for A with wire $W \le o(m \cdot n^{\frac{1}{2(d+k)}})$. Then this yields a 4-party communication for $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$ with $|\Pi_i^{\mathcal{M}}| < o(\sqrt{n})$, $|U| < o(m)$ and $|\Pi_i^{A\leftrightarrow B}| < o(\sqrt{n})$ from Lemma 3.1. But this is a contradiction to Theorem 1.2. □

Remark 3.3. *We remark that it is possible to obtain a weaker $\Omega(m + n^{1+1/(k+d)})$ lower bound more directly* without *using our communication lower bound Theorem 1.2. The proof was discovered in personal communication with Swastik Kopparty and Sepehr Assadi, and is based on a certain random-restriction argument for eliminating high fan-in gates, combined with the error-correcting properties of random linear operators. This argument, however, can only eliminate $o(n)$ gates (since every elimination shrinks the remaining input space $\{0,1\}^n$ by half), and thus quickly becomes trivial when the number of outputs is $m > n^{1.1}$ (say). By contrast, our lower bound can eliminate $o(m)$ gates, which allows us to prove a polynomial lower bound in the* number of outputs *so long as $m = \mathrm{poly}(n)$.*

3.2. NOF conjecture implications

In this section, we show that Pătraşcu's NOF Conjecture on the original Multiphase Game, even against 3-round protocols, would imply a breakthrough in circuit complexity. This complements our restricted NOF model, as it shows that allowing even *two* of Alice's messages to depend arbitrarily on her entire input $\vec{S}, i$, would resolve a decades-old open problem in circuit lower bounds. It also suggests that attacking the Multiphase conjecture for (general) dynamic data structures via the the NOF Game, should exploit the fact that data structures induce highly restricted NOF protocols.

First, note that Conjecture 1.1 in particular implies the following special case:

Conjecture 3.4 (3-round NOF Conjecture)**.** *Any 3-round NOF protocol for the 3-party Multiphase Game with $|U| = o(m)$ bits of advice must have $|\Pi| > n^{\varepsilon}$ communication for some constant $\varepsilon > 0$.*

Static Data Structure Lower Bound First, we consider the following class of static data structure problems $\mathcal{P}_A^f(x)$ defined by a query matrix $A \in \{0,1\}^{m \times n}$ and a function $f : \{0,1\}^{2n} \to \{0,1\}$:

1) Given a *fixed* matrix A with rows $A_1, \ldots, A_m$, preprocess an input database $x \in \{0,1\}^n$.
2) Given $i \in [m]$ as a query, the data structure needs to output $f(A_i, x)$.

Note that A is *hard-wired* to the problem, i.e., the data structure can access A for free during both preprocessing and query stage[5]. In particular with word size w, with $s = m/w$ cells, one can store the (boolean) answers for all possible queries and the problem becomes trivial ($t = 1$), whereas without any preprocessing, the query algorithm needs to read x (but not A) to compute the answer $f(A_i, x)$, giving (worst case) query time $t \sim n/w$. Accordingly, the query algorithm is *non-adaptive* if the cell addresses that are probed only a function of $i \in [m]$ and A.

We show that Conjecture 3.4 implies the following lower bound on $\mathcal{P}_A^f(x)$ where $f := \mathsf{DISJ}_n$.

Lemma 3.5 (Polynomial Static Lower Bound for Random Set Disjointness Queries)**.** *Suppose Conjecture 3.4 holds. Let $m = \omega(n)$. Then there exists a collection of m sets $A := A_1, \ldots, A_m \subseteq [n]^m$ such that any non-adaptive static data structure solving $\mathcal{P}_A^{\mathsf{DISJ}_n}$ must either use $s \geq \Omega(\frac{m}{w})$ space, or have $t \geq \Omega(n^{\varepsilon}/w)$ query time, in the cell-probe model with word size w.*

Proof. Suppose for any $A \in \{0,1\}^{m \times n}$ there is a non-adaptive static data structure D_A computing $\mathcal{P}_A^{\mathsf{DISJ}_n}$ with $s \leq o(\frac{m}{w})$ space and $t \leq o(n^{\varepsilon}/w)$ cell probes. We show that this induces a too-good-to-be-true (3-round) NOF protocol for the Multiphase game $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$, violating Conjecture 3.4. Indeed, consider the following simple 3-party protocol for simulating D_A:

Charlie's "advice" U in Phase 1 of the Multiphase game will be the contents of the s memory cells of D_A, Alice's message during Phase 2 (i.e., $\Pi_i^{A \to B}$) will be the memory addresses probed by the D_A for answering $\mathsf{DISJ}(A_i, x)$, and Bob's messages $\Pi_i^{B \to A}$ are the contents of cells probed by Alice. Note this protocol is well defined: Indeed, by the definition of $\mathcal{P}_A^{\mathsf{DISJ}_n}$, U only depends on A and x at preprocessing time, and if D_A is non-adaptive, then $\Pi_i^{A \to B}$ is only a function of A and i; Finally, $\Pi_i^{B \to A}$ depends on the previous transcript and $U = U(A, x)$ which Bob possesses. We therefore have a valid 3-round NOF protocol for $\mathsf{SEL}^m_{\mathsf{DISJ}_n}$ with

$$|U| + |x| \leq sw + n \leq o(m) + n = o(m), \quad \text{and}$$
$$|\Pi_i| = |\Pi_i^{A \to B}| + |\Pi_i^{B \to A}| \leq 2tw \leq o(n^{\varepsilon})$$

bits, which contradicts Conjecture 3.4. □

We consider the following parameter for the circuit lower bound.

Corollary 3.6. *If $m = \omega(n)$, and $s = o(m/w)$ then $t \geq \Omega(n^{\varepsilon}/w)$.*

Circuit Lower Bound. Now we show that Conjecture 3.4 implies circuit lower bounds using reduction by [Vio18] from Lemma 3.5. We use the following translation theorem for lower bounds in arbitrary depths.

Theorem 3.7 ([Vio18])**.** *Suppose function $f : \{0,1\}^n \to \{0,1\}^m$ has a circuit of depth d with W wires, consisting of unbounded fan-in, arbitrary gates. Then for any r there exists a static data structure (with non-adaptive query) with space $s = n + r$, query time $(W/r)^d$, and word size $\max\{\log n, \log r\} + 1$ which solves the following problem*

1) Preprocess input x depending on x and f
2) Given $i \in [m]$, output $f_i(x)$.

For completeness, we attach the proof here. Though [Vio18] did not remark on queries being non-adaptive, we remark that data structures derived from circuits are intrinsically non-adaptive, i.e. the memory cells probed only depends on query itself and f (but not on the content of cells probed along the way).

Proof of Theorem 3.7. Consider circuit C_f which computes f. Now set G as the set of gates with fan-in $> W/r$. Since the number of wires are bounded by W, $|G| < r$. Now given x, store the values of these gates G in space r.

Now we argue inductively on the level of the gate. We show that if the gate is at level ℓ, that the number of non-adaptive queries made to compute the value of the gate is at most $(W/r)^{\ell}$. As base case, suppose if it were a level 1 gate. If it lies in G, then the non-adaptive query required is 1, by calling to its address in G, which requires $\log |G|$-bits. Otherwise, the query required is at most W/r, each of which requires $\log n$-bits for the address, and they are non-adaptive. Therefore, the word size required is $\max\{\log n, \log r\} + 1$.

Now as induction step, suppose for any $j < \ell$, level j gates require at most $(W/r)^j$ non-adaptive queries to compute its value with word size $\max\{\log n, \log r\} + 1$. Consider a level ℓ gate. If the gate lies in G, again it only requires 1 non-adaptive query, by calling to its address in G. Otherwise, it can be answered by computing at most W/r level $\ell - 1$ gates, and these queries are non-adaptive. Each of these level $\ell - 1$ gates require $(W/r)^{\ell-1}$ non-adaptive queries by induction hypothesis, and the word size required is $\max\{\log n, \log r\} + 1$.

[5]This is a generalization of Valiant's model [Val77] in that the circuit itself is allowed to depend arbitrarily on A.

Therefore, the total non-adaptive query required is at most $(W/r) \cdot (W/r)^{\ell-1} = (W/r)^{\ell}$.

Since the final output gate that we are interested in is at level d, we get $(W/r)^d$ as the upper bound on the number of queries. $\square$

Here we crucially used the fact that wirings are fixed if we fix C_f. Note that G is determined by C_f. Then a contrapositive of Theorem 3.7 then states that data structure lower bound with non-adaptive query implies circuit lower bounds. When restricted to $o(m)$ additional space usage (i.e. $r = o(m)$, $f(n)$ query time lower bound translates to $\Omega\left(mf(n)^{1/d}\right)$ lower bound for W. Using the contrapositive along with Corollary 3.6, we show the following circuit lower bound.

Corollary 3.8. *Assuming Conjecture 3.4 and $m = \omega(n)$, there exists a matrix $A \in \{0,1\}^{m\times n}$ such that any depth-d circuit that computes*[6] *Ax requires $m \cdot n^{\Omega(\varepsilon/d)}$ wirings. In particular, if $d = 2$, there exists A that requires $m \cdot n^{\frac{\varepsilon}{2}-o(1)}$ wirings.*

Proof. We use contrapositive of Theorem 3.7, that is data structure lower bound implies circuit lower bounds. Consider Corollary 3.6. Setting $r = O(m)$, and $m = \omega(n)$. Then Corollary 3.6 with Theorem 3.7 implies that for some A, $(W/r)^d \geq \Omega(n^{\varepsilon}/\log n)$. Since $r = o(m)$, rewriting in terms of W, we get

$$W \geq \Omega\left(m \cdot n^{\frac{\varepsilon}{d} - \frac{\log\log n}{d \log n}}\right) = m \cdot n^{\Omega\left(\frac{\varepsilon}{d}\right)}.$$

Now setting $d = 2$, we get $W \geq m \cdot n^{\frac{\varepsilon}{2}-o(1)}$. $\square$

Acknowledgement

We are grateful to Huacheng Yu for his valuable comments and insightful discussions on this project, and to Swastik Kopparty and Sepehr Assadi for the discussion on Theorem 3.2.

[6]Over the Boolean semi-ring, with addition as OR and multiplication as AND

References

[BFS86] László Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory (preliminary version). In *27th Annual Symposium on Foundations of Computer Science*, pages 337–347, 1986.

[BK02] Jérémy Barbay and Claire Kenyon. Adaptive intersection and t-threshold problems. In *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '02, pages 390–399, Philadelphia, PA, USA, 2002. Society for Industrial and Applied Mathematics.

[BL15] Joshua Brody and Kasper Green Larsen. Adapt or die: Polynomial lower bounds for non-adaptive dynamic data structures. *Theory Comput.*, 11:471–489, 2015.

[BPP07] Philip Bille, Anna Pagh, and Rasmus Pagh. Fast evaluation of union-intersection expressions. In Takeshi Tokuyama, editor, *Algorithms and Computation*, pages 739–750, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[BY04] Ricardo Baeza-Yates. A fast set intersection algorithm for sorted sequences. In Suleyman Cenk Sahinalp, S. Muthukrishnan, and Ugur Dogrusoz, editors, *Combinatorial Pattern Matching*, pages 400–408, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[BYJKS02] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, FOCS '02, pages 209–218, Washington, DC, USA, 2002. IEEE Computer Society.

[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, 2012.

[CGL15] Raphaël Clifford, Allan Grønlund, and Kasper Green Larsen. New unconditional hardness results for dynamic and online problems. In *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, FOCS '15, pages 1089–1107, Washington, DC, USA, 2015. IEEE Computer Society.

[CP10] Hagai Cohen and Ely Porat. Fast set intersection and two-patterns matching. *Theor. Comput. Sci.*, 411(40-42):3795–3800, 2010.

[DLOM00] Erik D. Demaine, Alejandro López-Ortiz, and J. Ian Munro. Adaptive set intersections, unions, and differences. In *In Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA*, pages 743–752, 2000.

[Dru12] Andrew Drucker. Limitations of lower-bound methods for the wire complexity of boolean operators. In *Proceedings of the 2012 IEEE Conference on Computational Complexity (CCC)*, CCC '12, pages 170–180, Washington, DC, USA, 2012. IEEE Computer Society.

[GHK+13] Anna Gál, Kristoffer Arnsfelt Hansen, Michal Koucký, Pavel Pudlák, and Emanuele Viola. Tight bounds on computing error-correcting codes by bounded-depth circuits with arbitrary gates. *IEEE Trans. Information Theory*, 59(10):6611–6627, 2013.

[HKNS15] Monika Henzinger, Sebastian Krinninger, Danupon Nanongkai, and Thatchaphol Saranurak. Unifying and strengthening hardness for dynamic problems via the online matrix-vector multiplication conjecture. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing - STOC 15*. ACM Press, 2015.

[HW07] Johan Håstad and Avi Wigderson. The randomized communication complexity of set disjointness. *Theory of Computing*, 3(1):211–219, 2007.

[JS10] S. Jukna and G. Schnitger. Circuits with arbitrary gates for random operators. 2010.

[Juk10] Stasys Jukna. Representing (0, 1)-matrices by boolean circuits. *Discrete Mathematics*, 310(1):184 – 187, 2010.

[Juk12] Stasys Jukna. Boolean function complexity. *Algorithms and Combinatorics*, 2012.

[KPP15] Tsvi Kopelowitz, Seth Pettie, and Ely Porat. Dynamic set intersection. In *Lecture Notes in Computer Science*, pages 470–481. Springer International Publishing, 2015.

[Lar12] Kasper Green Larsen. Higher cell probe lower bounds for evaluating polynomials. In *FOCS 2012*, pages 293–301, 2012.

[Lup56] Oleg Borisovich Lupanov. On rectifier and switching-and-rectifier schemes. *Dokl. Akad. Nauk SSSR*, pages 111:1171–1174, 1956.

[LWY18] Kasper Green Larsen, Omri Weinstein, and Huacheng Yu. Crossing the logarithmic barrier for dynamic boolean data structure lower bounds. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 978–989, 2018.

[MPP05] Christian Worm Mortensen, Rasmus Pagh, and Mihai Patrascu. On dynamic range reporting in one dimension. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 104–111, 2005.

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

[Val77] Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In Jozef Gruska, editor, *Mathematical Foundations of Computer Science 1977*, pages 162–176, Berlin, Heidelberg, 1977. Springer Berlin Heidelberg.

[Vio18] Emanuele Viola. Lower bounds for data structures with space close to maximum imply circuit lower bounds. In *ECCC*, volume 25, 2018.

[Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing. In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.