Rounds vs Communication Tradeoffs for Maximal Independent Sets

Sepehr Assadi
Department of Computer Science
Rutgers University
New Brunswick, USA
sepehr.assadi@rutgers.edu

Gillat Kol
Department of Computer Science
Princeton University
Princeton, USA
gillat.kol@gmail.com

Zhijun Zhang
Department of Computer Science
Princeton University
Princeton, USA
zhijunz@princeton.edu

Abstract—We consider the problem of finding a maximal independent set (MIS) in the shared blackboard communication model with vertex-partitioned inputs. There are n players corresponding to vertices of an undirected graph, and each player sees the edges incident on its vertex – this way, each edge is known by both its endpoints and is thus *shared* by two players. The players communicate in simultaneous rounds by posting their messages on a shared blackboard visible to all players, with the goal of computing an MIS of the graph. While the MIS problem is well studied in other distributed models, and while shared blackboard is, perhaps, the simplest broadcast model, lower bounds for our problem were only known against one-round protocols.

We present a lower bound on the round-communication tradeoff for computing an MIS in this model. Specifically, we show that when r rounds of interaction are allowed, at least one player needs to communicate $\Omega(n^{1/20^{r+1}})$ bits. In particular, with logarithmic bandwidth, finding an MIS requires $\Omega(\log \log n)$ rounds. This lower bound can be compared with the algorithm of Ghaffari, Gouleakis, Konrad, Mitrović, and Rubinfeld [PODC 2018] that solves MIS in $O(\log \log n)$ rounds but with a logarithmic bandwidth for an *average* player. Additionally, our lower bound further extends to the closely related problem of maximal bipartite matching.

The presence of edge-sharing gives the algorithms in our model a surprising power and numerous algorithmic results exploiting this power are known. For a similar reason, proving lower bounds in this model is much more challenging, as this sharing in the players' inputs prohibits the use of standard number-in-hand communication complexity arguments. Thus, to prove our results, we devise a new round elimination framework, which we call partial-input embedding, that may also be useful in future work for proving *round-sensitive* lower bounds in the presence of shared inputs.

Finally, we discuss several implications of our results to multi-round (adaptive) distributed sketching algorithms, broadcast congested clique, and to the welfare maximization problem in two-sided matching markets.

Index Terms—distributed sketching, shared blackboard, communication complexity, maximal independent set, maximum matching

I. Introduction

Consider the following communication model: there are n players corresponding to vertices of an undirected graph $G = (V, E)$ and each player only sees the edges incident on its vertex – this way, each edge of the graph is *shared* by the two players at its endpoints. The goal of the players is to solve some fixed problem on G, for instance, finding a spanning forest of G. To do so, the players communicate in synchronous rounds wherein all parties simultaneously write a message on a shared blackboard visible to all. The messages communicated by the players are only functions of their own inputs and the content of the blackboard. When the protocol concludes, an additional party, called the *referee*, computes the output of the protocol as a function of the blackboard content. We are interested in the tradeoff between the number of rounds of the protocol and the *per-player* communication, defined as the worst-case length of any message sent by any player in any round.

In the communication complexity terminology, this model is referred to as the multi-party communication model with *shared blackboard* and *vertex-partitioned inputs*. However, it has also been studied by different communities under different names, such as *broadcast congested clique* [1]–[4], or *(adaptive) distributed sketching* [5]–[8]. At this point, there is quite a large body of algorithmic results in this model [5], [5], [6], [8]–[17] (see Section I-B for more details). The source of power behind these results is a crucial aspect of this model: *edge-sharing*, or in other words, the fact that each edge of the graph is seen by both its endpoints[1]. This sharing in the players' inputs makes this model an "intermediate" model lying between the *number-in-hand* model (with no input sharing) and the notorious *number-on-forehead* model (with arbitrary input sharing). As a result, lower bounds are more scarce in this model [1], [2], [7], [18]–[22].

We study the *maximal independent set (MIS)* problem in this model. While MIS is one of the most studied problems in other distributed models (see, e.g., [23]–[27]), and while shared blackboard is, perhaps, the simplest broadcast model, not much is known about MIS in this model. We do note that Luby's celebrated MIS algorithm [23] implies an $O(\log n)$-round $O(1)$-per-player communication algorithm in this model. Ghaffari, Gouleakis, Konrad, Mitrovic, and

[1]The interested reader is referred to [5] to see this in a surprising algorithm that solves graph connectivity using only a single round and $O(\log^3 n)$ communication bits per player.

Rubinfeld [28] gave an algorithm that runs in $O(\log \log n)$ rounds, but only bounds the communication of an *average* player by $O(\log n)$. I.e., the total communication by all players in a round is $O(n \log n)$, but some players may need to communicate $\omega(\log n)$ bits[2]. Moreover, Assadi, Kol, and Oshman proved that any *one*-round protocol requires almost $(n^{1/2})$ per-player communication. This state-of-affairs raises the following question:

> *What is the complexity of MIS in the shared black-board model with vertex-partitioned inputs? In particular, what are the possible round-communication tradeoffs in this model?*

We make progress on this fundamental open question by presenting a new lower bound on the round-communication tradeoff for the MIS problem. The key contribution of our work is a new technique for proving multi-round lower bounds, even in the presence of edge-sharing. This also allows us to prove a similar lower bound for another fundamental problem, namely, the maximal bipartite matching problem.

Our work can be viewed as a direct continuation of two lines of work: the first line of work is on *number-in-hand* multi-round communication complexity, where we follow up on the result of Alon, Nisan, Raz, and Weinstein [29]. They give lower bounds for the bipartite maximal matching problem, where only parties on one side of the partition are allowed to communicate. The second line of work is the aforementioned lower bound of Assadi, Kol, and Oshman [7], which works in our model, but only considers one-round protocols. In the following, we elaborate more on our results, techniques, and their connections to other settings.

We remark that most proofs are omitted in this version. Interested readers are referred to the full version [30].

A. Our Contributions

Our main result is a multi-round lower bound for computing MIS in the shared blackboard model.

> **Result 1.** Any r-round multi-party protocol (deterministic or randomized) in the shared blackboard model for finding a maximal independent set on n-vertex graphs requires $\Omega(n^{1/20^{r+1}})$ bits of communication per player. In particular, $\Omega(\log \log n)$ rounds are needed for protocols with polylog(n) per-player communication.

Previously, the only known lower bound for MIS in our model was the (almost) $\Omega(n^{1/2})$-communication lower bound of [7] for one-round protocols. Indeed, to the best of our knowledge, there has been no prior communication lower bound in this model for any natural problem that is *sensitive* to the number of rounds (the lower bounds were either for

one-round protocols, e.g., [7], [21], [22], or arbitrary number of rounds, e.g., [1], [2][3]).

The tradeoff achieved in Result 1 asymptotically matches the aforementioned $O(\log \log n)$-round algorithm of [28] for finding MIS, except that, as mentioned before, the protocol of [28] only bounds the communication of an *average* player by $O(\log n)$ bits and a few players need to communicate way more than polylog(n) bits. Thus, the two results do not directly match. It remains an interesting open question to either improve the guarantee of the algorithm of [28] to per-player communication bound or improve our lower bound to average-case communication.

Our techniques in establishing Result 1 are quite general and, as a corollary to our proof, also allow us to prove a lower bound for another fundamental problem, namely, maximal matching.

> **Result 2.** Any r-round multi-party protocol (deterministic or randomized) in the shared blackboard model for finding a maximal matching or any constant factor approximation to maximum matching on n-vertex (bipartite) graphs requires $\Omega(n^{1/20^{r+1}})$ bits of communication per player. As such, $\Omega(\log \log n)$ rounds are needed for protocols with polylog(n) per-player communication.

As in the case of MIS, the only known lower bound prior to our work was the one-round lower bound of [7]. However, for the *number-in-hand* variant of our communication model, wherein each edge of the graph is only seen by one of its endpoints, a series of papers [29], [31], [32] proved a nearly-logarithmic round lower bound for the matching problem (we elaborate on this line of work later). Yet, the number-in-hand model is algorithmically much weaker than the edge-sharing model studied in our paper; for instance, the lower bound of [32] also holds for finding a spanning forest of the input in that model, while finding spanning forests in our model can be done with $O(\log^3 n)$ communication in just one round [5]. We refer the reader to [7] for discussions on the inherit difference of number-in-hand model and our model that allows for edge-sharing and thus is "one step closer" to the notorious number-on-forehead model.

a) Our techniques: We shall go over our techniques in detail in the streamlined overview of our approach in Section III. For now, we only mention the high level bits of our techniques.

Our techniques unify and generalize the lower bounds of [7] for one-round protocols in our model, as well as the lower bounds of [29] for multi-round protocols in the number-in-hand model. To this end, we need several substantially new

---

[2]This algorithm is designed for the (unicast) congested clique model, but given its connection to the distributed sketching/dynamic streaming algorithm of [17]—that solves MIS as a subroutine in correlation clustering—it can be directly implemented in our model with the mentioned bounds.

[3]Specifically, the latter ones bound the *total* communication needed to solve the problem and use this to get a lower bound on the number of rounds *times* communication per round. Such lower bounds cannot capture more nuanced round-communication tradeoffs (e.g., like the ones exhibited by [23] or [28] for MIS).

ideas[4]. The main novelty of our work is in developing a new *round elimination* argument that is tailored to our edge-sharing model. Similar to standard round elimination arguments, say, the one in [29], our approach is also based on *simulating* an r-round protocol on "large" instances in only $(r-1)$ rounds for smaller "embedded" instances (with fewer players and smaller inputs). Prior work perform such a simulation by generating an *input* for the "missing" players of the large r-round instance with *low correlation* with the actual embedded $(r-1)$-round hard instance. As we argue, such an approach is doomed to fail for our model with its edge-sharing aspects. Instead, we introduce a *partial-input embedding* argument that implements this simulation via generating only the *messages* of the missing players. We then use information-theoretic tools to track the gradual increase in the *correlation* of these messages with the embedded hard instance throughout the *entire* simulation (not only the first round which is sufficient for "input-sampling" protocols of prior work).

B. Further Implications of Our Results to Related Models

We conclude this section by listing further implications of our results to other well-studied settings.

a) Broadcast congested clique: The communication model studied in our paper is equivalent to the broadcast congested clique model studied in various prior work, e.g., in [1]–[4]. Specifically, our Result 1 and Result 2 imply $\Omega(\log \log n)$ round lower bounds for both MIS and maximal matching on any broadcast congested clique algorithm with polylog(n) bandwidth. Incidentally, in the stronger *unicast* congested clique model, $O(\log \log n)$-round algorithms are known for both MIS [28] and maximal matching [33]. We note that, as shown in [1], proving lower bounds in the unicast model implies strong circuit lower bounds and thus is beyond the reach of current techniques.

b) Distributed sketching: Our model is also equivalent to the distributed sketching model that was initiated in the breakthrough work of [5]. Starting from the connectivity sketch of [5], there has been tremendous progress on efficient distributed sketching algorithms for various other problems in *one* round, e.g., cut sparsifiers [6], spectral sparsifiers [9], [10], vertex connectivity [11], densest subgraph [12], $(\Delta+1)$-coloring [13], Δ-coloring [14], and in *multiple* rounds, e.g., minimum spanning trees [5], matchings [5], [15], spanners [8], [16], and MIS and correlation clustering [17]. Given the strength of this model, proving lower bounds in this model has been a highly challenging task (see, e.g. [7], [8]), and only a handful of lower bounds are known including $\Omega(\log^3 n)$ bits for connectivity [21], [22] and $\Omega(n^{1/2})$ bits for MIS and maximal matching [7] for *one*-round sketches. Our results contribute to this line of work by providing the first *round-*

sensitive lower bounds in this model, and our techniques can be of independent interest here as well.

c) Dynamic streaming algorithms: One key motivation of [5] in introducing graph sketching was their application to *dynamic* (semi-)streaming algorithms that can process streams of insertions and deletions of edges with $O(n \cdot \text{poly} \log (n))$ memory (*all* sketches mentioned above also imply dynamic streaming algorithms). Multi-round sketching protocols, similar to the ones in our model, then correspond to multi-pass streaming algorithms. Currently, the best known multi-pass dynamic semi-streaming algorithms for MIS and maximal matching require $O(\log \log n)$ passes [17] and $O(\log n)$ passes [5], [15], respectively. On the lower bound front however, only single-pass lower bounds are known for either problem [13], [34]–[36] (there has been recent progress on multi-pass lower bounds for computing *exact* maximum matchings [37]–[39] in logarithmic passes or even $(1+o(1))$-approximation in two passes [40] but they do *not* apply to maximal matching in any way). While our results do *not* imply streaming lower bounds, they do rule out certain popular techniques of vertex-partitioned graph sketching for obtaining $o(\log \log n)$-pass algorithms for either problem. Thus, they can form a starting point for proving multi-pass lower bounds for all dynamic streaming algorithms as well.

d) Welfare maximization and interaction: A beautiful line of work initiated by [31] and followed up in [29], [32], [41], [42], studies the role played by the *interaction* of participating agents in the efficiency of markets. One formalization, corresponding to unit-demand agents in a matching market, is as follows: we have n agents who are interested in getting any one of their private subset of n items; the goal is to allocate these items in a way that maximizes the *welfare*, defined as the number of agents who receive an item of their liking. The market proceeds in rounds wherein the agents communicate polylog(n)-bit messages about their desired items. How many rounds of interaction are needed to maximize the welfare to within a constant factor?

This problem can be seen as approximating matchings on the bipartite graph consisting of agents on one side that have edges to their preferred items on the other side. The model of communication is also identical to the one in our paper with the crucial difference that only vertices on one side of the bipartition, namely, the agents, are communicating. In this model, [31] gave an $O(\log n)$-round algorithm and ruled out one-round algorithms. [29] improved the lower bound to $\Omega(\log \log n)$ rounds and subsequently [32] obtained a nearly tight $\Omega(\frac{\log n}{\log \log n})$ lower bound (similar lower bounds are also obtained for the more general setting of combinatorial auctions in [41]).

All these results are restricted to one-sided markets. Our Result 2 generalizes (some of) these results to *two-sided* matching markets [43], wherein *both* sides of the market consist of communicative agents that know in advance if they make a good match. A canonical example of two-sided matching markets is college admissions and the celebrated Gale-Shapley algorithm for stable marriage [44]. Another example, perhaps

---

[4]Braverman and Oshman [32] gave stronger lower bounds than [29], that work for nearly logarithmic number of rounds. However, their techniques seem "too tailored" to the number-in-hand model and approximate matchings, and thus are *not* suitable for us (given the algorithm of [28] for MIS, which, even though not exactly in our model, seem quite close, it is not clear if one can get a logarithmic lower bound in our model).

more closely related to the setting of our paper, is assigning users to servers in a large distributed Internet service [45]. Our Result 2 suggests that even when both sides of the market are able to communicate with a limited bandwidth, at least a modest amount of interaction is necessary for maximizing welfare (approximately).

II. PRELIMINARIES

a) Notation: For an integer $t \in \mathbb{N}$, we write $[t]$ as a shorthand for the set $\{1, \ldots, t\}$. Let $h : A \to B$ be an arbitrary function for two sets A, B. For any subset $Z \subseteq A$, we use $h(Z) = \{h(z) \mid z \in Z\}$. For a tuple $X = (X_1, \ldots, X_t)$ and integer $i \in [t]$, we define $X_{<i} = (X_1, \ldots, X_{i-1})$ (we also define X_{-i} and $X_{\leqslant i}$ analogously). For a graph $G = (V, E)$ and a permutation σ over V, we denote by $\sigma(G)$ the graph on the same vertex set in which $\sigma(u)$ and $\sigma(v)$ are connected if and only if $(u, v) \in E$.

When there is room for confusion, we use sans-serif letters for random variables (e.g. A) and the same normal letters for their realizations (e.g. A). For random variables A, B, we use $\mathrm{supp}(\mathsf{A})$ as the support of A, $\mathbb{H}(\mathsf{A})$ as the *Shannon entropy*, $\mathbb{I}(\mathsf{A} \,;\mathsf{B})$ as the *mutual information*, $\mathbb{D}(\mathsf{A} \,\|\, \mathsf{B})$ as the *KL-divergence*, and $\|\mathsf{A} - \mathsf{B}\|_{\mathrm{tvd}}$ as the *total variation distance*. We refer the interested readers to the textbook by Cover and Thomas [46] for necessary background on information theory, including the definitions and basic tools used in our paper.

A. Multi-Party Shared Blackboard Model with Vertex-Partitioned Inputs

We work in the multi-party shared blackboard model with vertex-partitioned inputs, also known as the broadcast congested clique model in the literature. The communication model is defined formally as follows. Consider a simple graph $G = (V, E)$ with one player assigned to each of the $n = |V|$ vertices. For convenience, we identify a vertex with its associated player in the rest of this paper and use the two terms interchangeably. There is a *shared blackboard*, initially empty, that is readable and writable by all players. The player associated to a vertex $v \in V$ is presented as input with n, a unique ID of v in the range $[n]$, and IDs of all of v's neighbors $N_G(v) = \{u \in V \mid (v, u) \in E\}$. Thus, each edge $(u, v) \in E$ is *shared* by both players u and v.

Communication proceeds in $r \in \mathbb{N}$ *synchronous* rounds. For each round $t \in [r]$, the players compute their messages based on their initial input as well as the current content of the blackboard, and post them to the blackboard *simultaneously*. In a randomized protocol, the players may also use both public and private randomness. After the last round, the final content of the blackboard constitutes the *transcript*, denoted by Π, of the protocol. Then, a *referee* computes the output of the protocol depending on Π (and possibly public randomness of all players and its own private randomness). The *bandwidth* of a protocol is defined to be the *maximum* number of bits ever communicated by any player in any round.

We are interested in round-communication tradeoff of the following problems:

a) Maximal Independent Set: We say a protocol computes a maximal independent set (MIS) with error probability $\delta \in [0, 1]$ if the output of the referee is a valid MIS of G with probability at least $1 - \delta$ over the randomness of the protocol. The protocol may err by outputting a subset of vertices which is not independent or not maximal.

b) Approximate Matching: We say a protocol computes an α-*approximate matching* ($\alpha \geqslant 1$) if the output $\Gamma(\Pi)$ of the referee: (1) is always a set of *disjoint* pairs of vertices; and (2) satisfies $\mathbb{E} |\Gamma(\Pi) \cap E| \geqslant \mu(G)/\alpha$, where $\mu(G)$ is the size of the maximum matching of G and expectation is taken over the randomness of the protocol. This definition allows the referee to output *non-existing* edges as long as they are disjoint but only the correct ones in E are counted. This is a less restrictive error-model than requiring the algorithm to output a valid matching with certain probability and our lower bound holds even in this less restrictive setting; see also [29].

III. TECHNICAL OVERVIEW

As our proof is quite dense and technical and involves various information theoretic maneuvers that are daunting to parse, we use this section to unpack our main ideas and give a streamlined overview of our approach. We emphasize that this section oversimplifies many details and the discussions will be informal for the sake of intuition.

The starting point of our approach is a lower bound of [29] for approximate matchings in the *number-in-hand* multi-party communication model. We first give a detailed discussion of this result as our techniques need to inevitably subsume this work (since our result implies theirs as well). We then discuss the challenges of extending this result to our model that allows for edge-sharing and present a technical overview of our work. We stick with approximate matchings in this overview as it is easier to work with and to compare with [29].

A. A Detailed Overview of [29]

[29] considers the same communication setting as ours on bipartite graphs $G = (L \sqcup R, E)$ with the key difference that the players are only associated with vertices in L, and thus each edge is seen by only a single player. They prove that any protocol that uses polylog(n) communication per player and computes an $O(1)$-approximate matching requires $\Omega(\log \log n)$ rounds in this model.

The proof in [29] is via **round elimination**: to lower bound polylog (n)-communication r-round protocols π_r, they start with $p_r \approx n^{4/5}$ *independent* $(r-1)$-round "hard" instances $I_1, \ldots, I_{p_r}$, called *principal* instances. These instances are supported on disjoint sets of $\approx n^{1/5}$ vertices each, and are then "embedded" in a single graph G to form an r-round instance I. This instance is such that the *first* message of π_r cannot reveal much information about principal instances and thus π_r cannot solve them in its remaining $r-1$ rounds given their (inductive) hardness.

To limit the information revealed by π_r about principal instances, [29] further "packs" the graph, for every principal instance $i \in [p_r]$, with $f_r \approx n^{2/5}$ *fooling* instances $J_{i,*} :=$
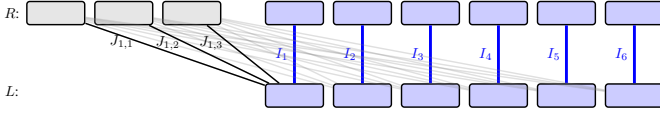
Fig. 1. An illustration of the lower bound instances of [29] with parameters $f_r = 3$ and $p_r = 6$. The top right vertices (blue) are used in principal instances, while top left vertices (gray) are fooling instances. The heavy (blue) edges are from principal instances and the light (gray) edges are from fooling instances – to avoid clutter, only the edges in fooling instances of the first principal instance are drawn (solid black edges). To find a large matching in this graph, one needs to find sufficiently large matchings in many of the principal instances.

$J_{i,1}, \ldots, J_{i,f_r}$. This packing ensures that: (1) these fooling instances are supported on a small set of vertices on the R-side of the bipartition and so π_r still has to solve *most of* the underlying principal instances in order to solve I; and (2) each player in I "plays" in $f_r + 1$ instances, consisting of only one principal instance, while being *oblivious* to which instance is the principal one. An ingenious idea in [29] is that these fooling instances need *not* actually be hard $(r-1)$-round instances! Instead, they form a product distribution where for each vertex $v \in L$, only the marginal distribution of v is the same under fooling and principal instances. This ensures that in the first round (and only in this round), v cannot distinguish between principal and fooling instances.

    a) Round elimination embedding: We can now discuss how [29] *eliminates* the first round of π_r and obtains an $(r-1)$-round protocol σ for solving an $(r-1)$-round hard instance $I^\star$.

Embedding argument of [29]:

(i) The players in σ sample the first message $M^{(1)}$ of π_r using *public* randomness.

(ii) Then, they will sample an index $i \in [p_r]$ uniformly and let $I_i = I^\star$ in the instance I.

(iii) Next, they sample $J_{i,1}, \ldots, J_{i,f_r}$ conditioned on $M^{(1)}$ and $I_i = I^\star$ using *private* randomness. This is a non-trivial sampling process which, on a high level, is doable only because fooling instances are product distributions (with only the marginals matching principal ones).

    More specifically, each player v *independently* sample its own input $J_{i,*}(v)$ in all the fooling instances, conditioned on only its actual input $I_i(v)$ in its principal instance I_i, and $M^{(1)}$.

(iv) Finally, the players of σ sample the remaining $p_r - 1$ principal instances I_{-i} and $(p_r - 1) \cdot f_r$ fooling instances $J_{-i,*}$ conditioned on $M^{(1)}$ to have a complete instance I.

    At this point, the players in σ already have the first message $M^{(1)}$ of π_r as well as inputs of all underlying instances without any communication. So, they can continue running π_r from

its second round, by each player of σ on $I^\star$ communicating the messages of corresponding player of π_r in I_i, and simulating messages of π_r for players outside I_i with no communication. As π_r will also need to solve I_i for a random $i \in [p_r]$, this gives a $(r-1)$-round protocol σ for $I^\star = I_i$.

    At a high level, the correctness of this approach can be argued as follows:

- The *right* distribution of all underlying variables for π_r can be expressed as (by chain rule):

$$M^{(1)} \times (I_i \mid M^{(1)}) \times (J_{i,*} \mid I_i, M^{(1)})$$
$$\times (I_{-i}, J_{-i,*} \mid J_{i,*}, I_i, M^{(1)}). \quad (1)$$

- The distribution sampled from in the protocol σ on the other hand is:

$$\underbrace{M^{(1)}}_{\text{publicly}} \times \underbrace{I_i}_{\text{input}} \times (\underbrace{\bigtimes_v J_{i,*}(v) \mid I_i(v), M^{(1)})}_{\text{privately}}$$
$$\times \underbrace{(I_{-i}, J_{-i,*} \mid M^{(1)})}_{\text{publicly}}. \quad (2)$$

    Let us show that these distributions are $o(1)$-close in total variation distance, which implies that π_r also works (almost) as good on sampled instances, giving us the desired $(r-1)$-round protocol σ for $I^\star$. Here, the first terms are the same. For the second terms,

$$\|I_i - (I_i \mid M^{(1)})\|_{\text{tvd}}^2 \leqslant \mathbb{I}(I_i ; M^{(1)})$$
$$\leqslant \frac{1}{f_r + 1} \cdot \mathbb{I}(J_{i,*}, I_i ; M_i^{(1)}) \leqslant o(1). \quad (3)$$

In (3), the first inequality is standard. The second inequality uses the fact that the players in I_i in π_r are oblivious to origins of their edges in I_i vs. $J_{i,*} = J_{i,1}, \ldots, J_{i,f_r}$ (by the marginal indistinguishability of these instances); thus, the information revealed by their messages $M_i^{(1)}$ is "spread" over these instances; also, other players of π_r cannot reveal any information about these instances as they do not see them. The final inequality holds because the messages communicated by $\approx n^{1/5}$ players in I_i have collective size much smaller than $f_r \approx n^{2/5}$.

    Finally, the third and fourth terms in (1) and (2) also have the same distributions in both cases, which at a high level, follows from the rectangle property of communication protocols: for instance, since $J_{i,*}(u)$ and $J_{i,*}(v)$ were independent originally, they remain independent even after conditioning on $M^{(1)}$ – this is sufficient to show the equivalence of corresponding distributions. This concludes the closeness of these distributions and our overview of the work of [29].

B. Our Approach and New Ideas

    The very first obvious challenge in using construction of [29] in our model is that it can be easily solved in just a single round once *both* sides of the bipartite graph can speak (the maximum matching of instances created is incident on vertices with degree one in R who can just communicate their edge directly on the blackboard). This brings us to the first and most obvious of our ideas.
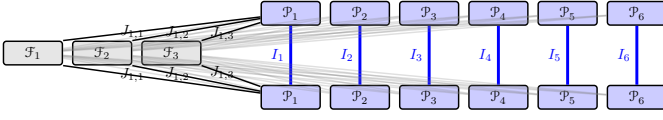
Fig. 2. An illustration of our lower bound instances with parameters $f_r = 3$ and $p_r = 6$. The top and bottom vertices (blue) are principal blocks, while middle left vertices (gray) are fooling blocks. The heavy (blue) edges are from principal instances and the light (gray) edges are from fooling instances – to avoid clutter, only the edges in fooling instances of the first principal instance are drawn (solid black edges). Note that fooling blocks participate only in fooling instances while principal blocks participate both in principal and fooling instances.

1) Idea One: Symmetrizing the Input Distribution: The first step is to symmetrize the input distribution in [29]. Basically, to create a hard r-round instance, we again start with $(r-1)$-round hard principal instances $I_1, \ldots, I_{p_r}$. We then also add f_r sets of vertices $\mathcal{F}_1, \ldots, \mathcal{F}_{f_r}$ called the *fooling blocks* and use vertices on *both* sides of each principal instance I_i, called *principal block* $\mathcal{P}_i$, and the fooling blocks to form fooling instances $J_{i,1}, \ldots, J_{i,f_r}$ – as before, these fooling instances are *not* hard $(r-1)$-round distributions, but only that the input of principal blocks match the "right" distribution *marginally*.

This step of symmetrizing the input distribution is a straightforward extension of [29], and we claim no novelty in this part. The interesting part is how to analyze this distribution in our model in light of the following key differences from [29]: (1) in addition to principal blocks, vertices in $\mathcal{F}_1, \ldots, \mathcal{F}_{f_r}$ can now also communicate; and (2) there is an edge-sharing aspect in our model; in particular, sharing of edges between fooling blocks and principal blocks allows fooling blocks to communicate even about edges directly inside principal instances (!), and yet fooling blocks themselves are not even fooled anymore in the distribution. We discuss our approach for handling these parts in the following three subsections.

2) Idea Two: Bounding Revealed Information on Average: Our goal as before is to do a round elimination argument and embed an $(r-1)$-round instance inside an r-round one. Our embedding argument in the *first* round is going to be the same as that of [29], except that we also sample the first message $M_F^{(1)}$ of fooling blocks using public randomness (there are *no* such players in [29]). We will then have all the messages of round one, namely, $M^{(1)} = (M_P^{(1)}, M_F^{(1)})$, as well as edges incident on the principal block $\mathcal{P}_i$, namely, $I_i, J_{i,*}$, inside I without having done any communication.

Specifically, we design a protocol σ that given an $(r-1)$-round instance $I^\star$, creates an r-round instance I and uses a $\mathrm{polylog}(n)$-communication r-round protocol π_r on I to solve $I^\star$ as follows. Let us argue that the joint distribution of obtained random variables at this point is close to that of the actual distribution induced by π_r (similar to (1) and (2) for [29]):

Our embedding argument – first round:

(i) Players in σ sample the first message $M_P^{(1)}, M_F^{(1)}$ of principal and fooling blocks *publicly*.

(ii) Then, they will sample an index $i \in [p_r]$ uniformly and let $I_i = I^\star$ in the instance I; thus, players in σ will play the role of principal block $\mathcal{P}_i$ in π_r from now on.

(iii) Next, they sample $J_{i,*}$ conditioned only on $M_P^{(1)}$ and $I_i = I^\star$ using *private* randomness by each vertex v of σ independently sampling $J_{i,*}(v)$ only conditioned on $I_i(v), M_P^{(1)}$.

- The *right* distribution of the underlying variables for π_r can be expressed as:

$$(M_P^{(1)}, M_F^{(1)}) \times (I_i \mid M_P^{(1)}, M_F^{(1)})$$
$$\times (J_{i,*} \mid I_i, M_P^{(1)}, M_F^{(1)}). \quad (4)$$

- The distribution sampled from in the protocol σ is:

$$\underbrace{(M_P^{(1)}, M_F^{(1)})}_{\text{publicly}} \times \underbrace{I_i}_{\text{input}} \times \underbrace{(\times_v J_{i,*}(v) \mid I_i(v), M_P^{(1)})}_{\text{privately}}. \quad (5)$$

The first terms are the same. For the second terms, similar to (3), we have,

$$\|I_i - (I_i \mid M_P^{(1)}, M_F^{(1)})\|_{\mathrm{tvd}}^2$$
$$\leqslant \mathbb{I}(I_i ; M_P^{(1)}, M_F^{(1)})$$
$$= \mathbb{I}(I_i ; M_P^{(1)}) + \mathbb{I}(I_i ; M_F^{(1)} \mid M_P^{(1)}), \quad (6)$$

using the chain rule of mutual information in the equality. The first term in RHS above can still be bounded by $o(1)$ by the same logic that principal blocks are oblivious to identity of principal instance edges in their input. But such a statement is not true about fooling blocks in the second term, as those vertices themselves are not fooled. Consider the following 1-bit protocol.

Example. Suppose we direct each edge of the graph randomly to one of its endpoints using public randomness. Principal blocks[a] send the XOR of their *outgoing* edges and fooling blocks send the XOR of their *incoming* edges incident on $J_{i,*}$ for some $i \in [p_r]$. Taking the XOR of messages sent by $\mathcal{P}_i$, $M_{P,i}^{(1)}$, and fooling blocks, $M_F^{(1)}$, reveals XOR of all edges inside I_i as each such edge will be outgoing for exactly one endpoint and edges in $J_{-i,*}$ cancel out in this XOR. This reveals one bit of information about I_i, making $\mathbb{I}(I_i ; M_F^{(1)} \mid M_P^{(1)}) \geqslant 1$. (Ideas like this are used in actual distributed sketching protocols, e.g., in [5], [10].)

---
[a]A player can know whether it is principal or fooling simply based on its degree.

Instead, we show that fooling blocks cannot reveal much about I_i for an *average* $i \in [p_r]$:

$$\mathbb{E}_i[\mathbb{I}(\mathsf{I}_i \,; \mathsf{M}_F^{(1)} \mid \mathsf{M}_P^{(1)})] \leqslant \frac{1}{p_r} \cdot \mathbb{I}(\mathsf{I}_1, \ldots, \mathsf{I}_{p_r} \,; \mathsf{M}_F^{(1)} \mid \mathsf{M}_P^{(1)})$$
$$\leqslant o(1), \tag{7}$$

where in the second inequality we used the fact that the polylog(n)-bit messages of *all* $f_r \approx n^{2/5}$ fooling blocks of size $\approx n^{1/5}$ cannot reveal more than $o(p_r)$ information as $p_r \approx n^{4/5}$ (this idea is similar to the "public-vs-private" vertices of [7] for one-round lower bounds in the distributed sketching model). This allows us to bound the LHS of (6) on average for $i \in [p_r]$. A similar type of argument can be applied to the third terms also to "drop" the conditioning on $\mathsf{M}_F^{(1)}$, while changing the distribution only by $o(1)$ in total variation distance. This implies that

$$\mathbb{E}_i \|(\mathsf{J}_{i,*} \mid \mathsf{I}_i, \mathsf{M}_P^{(1)}, \mathsf{M}_F^{(1)}) - (\mathsf{J}_{i,*} \mid \mathsf{I}_i, \mathsf{M}_P^{(1)})\|_{\mathrm{tvd}} \leqslant o(1).$$

By [29], the second distribution here matches the product distribution sampled privately by the players (the third term of (5)). This is now sufficient for simulating the *first round* of π_r (almost) faithfully with no communication as $i \in [p_r]$ is also chosen randomly in the embedding[5].

It is tempting to consider our job done as we successfully simulated the first round of π_r with no communication, and thus we *eliminated* a round. But in fact, this is just the start of **the unique challenges of our model**. Unlike [29], it is not clear how we can continue running π_r in the subsequent rounds: in σ, we have only decided on the input of principal block $\mathcal{P}_i$ in I – the input to other principal blocks and all fooling blocks are still undecided, and so π_r is not well defined for the subsequent rounds. We now need to deviate entirely from [29] to handle this.

3) Idea Three: Partial-Input Embedding and Non-Simultaneous Simulation: To continue running π_r from its second round onwards, we should be able to simulate *all* players in I, not only the principal block $\mathcal{P}_i$ responsible for $I_i = I^\star$. Let us consider a standard approach.

a) Standard approach for handling remaining instances: The standard approach is to sample input of remaining players in π_r using public randomness and let the "actual" players of σ simulate them "in their head" with no communication (this corresponds to step (iv) of embedding of [29]). This approach fails completely for us. Consider the fooling blocks first: at this point in the protocol σ, the players have sampled $J_{i,*}$ *privately* which was necessary in the first round (given the correlation of $J_{i,*}(v)$ with $I_i(v)$ via $\mathsf{M}_P^{(1)}$ and that $I_i(v)$ was only known to v). But given that the other endpoints of these edges are in fooling blocks, this means that *no* single player of σ can even know the edges incident on a single vertex in fooling blocks, leaving no player to simulate players of π_r in fooling blocks (or sampling rest of their inputs).

A more subtle issue happens when it comes to the rest of principal blocks, which on the surface, should be fine given they share no edges with principal block $\mathcal{P}_i$. To be able to sample instances $I_{-i}, J_{-i,*}$ publicly in the last step of embedding, we need the following two distributions to be close:

$$\underbrace{(\mathsf{I}_{-i}, \mathsf{J}_{-i,*} \mid \mathsf{J}_{i,*}, \mathsf{I}_i, \mathsf{M}_P^{(1)}, \mathsf{M}_F^{(1)})}_{\text{right distribution}}$$
$$\text{vs.} \quad \underbrace{(\mathsf{I}_{-i}, \mathsf{J}_{-i,*} \mid \mathsf{M}_P^{(1)}, \mathsf{M}_F^{(1)})}_{\text{"input-sampling"-protocol distribution}}$$

Yet, even a 1-bit communication protocol can turn these two distributions far from each other:

> **Example.** Suppose principal blocks remain silent and each fooling block sends the XOR of their incident edges. Then conditioned on the messages $\mathsf{M}_F^{(1)}$, once we additionally know $J_{i,*}$, we learn the parity of edges in $J_{-i,*}$ which changes the distribution of $J_{-i,*}$ by $\Omega(1)$.

All in all, when it comes to our edge-sharing model, the standard approach of sampling the remaining instances inherently fails: (i) fooling blocks are directly incident on edges in $J_{i,*}$ which are part of the input to players in $\mathcal{P}_i$ in π_r; (ii) worse yet, the messages of fooling blocks even correlate inputs of the rest of principal vertices with those of $\mathcal{P}_i$, meaning that *all* principal players can reveal information about I_i not only the ones in $\mathcal{P}_i$ that are directly incident on it.

b) Our approach for handling remaining instances: A key idea we use in the rest of our protocol is what we call **partial-input embedding**: we only generate the rest of the input for players $\mathcal{P}_i$ and for all the remaining players, we will simulate them solely by *sampling their messages* without ever committing to their input. Thus, our embedding keeps going even beyond the first round as we will need to generate the messages of remaining players throughout the entire execution of π_r.

In particular, after running the embedding part of the first round, for any round $t > 1$, the players in the protocol σ will simulate the t-th round of π_r as follows:

Our embedding argument – after first round:

(i) The players in σ communicate messages of $\mathcal{P}_i$ using the current content of the blackboard $M^{(<t)}$, and their inputs $J_{i,*}, I_i$ sampled for the first round, and send the messages $M_{P,i}^{(t)}$.

(ii) *After* this message is revealed, the players use *public* randomness to sample the t-th message of remaining players $M_{-i}^{(t)} := (M_{P,-i}^{(t)}, M_F^{(t)})$ conditioned on public knowledge $M^{(<t)}, M_{P,i}^{(t)}$.

---

[5] [29] also works with a random $i \in [p_r]$ but only to ensure that the underlying instance I_i needs to be solved by π_r as most but not all principal instances are solved in π_r – *all* information-theoretic guarantees for π_r mentioned for the embedding of [29] hold for *arbitrary* $i \in [p_r]$ unlike ours.

It is worth pointing out a rather strange aspect of this embedding. In π_r itself, the messages $M_{P,i}^{(t)}$ and $M_{-i}^{(t)}$ are communicated *simultaneously* with each other. Yet, in our simulation of π_r, we are crucially using messages principal block $\mathcal{P}_i$ to help us generate the remaining messages! We will discuss the necessity of this **non-simultaneous simulation** of a round in the next subsection.

As before, let us examine the underlying distributions in the first t rounds for $t > 1$:

- The right distribution of the underlying variables up until this point in π_r is:

$$\underbrace{(\mathsf{M}^{(<t)}, \mathsf{J}_{i,*}, \mathsf{I}_i)}_{\text{prior rounds}} \times (\mathsf{M}_{P,i}^{(t)} \mid \mathsf{M}^{(<t)}, \mathsf{J}_{i,*}, \mathsf{I}_i)$$
$$\times (\mathsf{M}_{-i}^{(t)} \mid \mathsf{M}_{P,i}^{(t)}, \mathsf{M}^{(<t)}, \mathsf{J}_{i,*}, \mathsf{I}_i). \quad (8)$$

- The distribution sampled from in the protocol σ is:

$$\underbrace{(\mathsf{M}^{(<t)}, \mathsf{J}_{i,*}, \mathsf{I}_i)}_{\text{prior rounds}} \times \underbrace{(\bigtimes_v \mathsf{M}_{P,i}^{(t)}(v) \mid \mathsf{M}^{(<t)}, \mathsf{J}_{i,*}(v), \mathsf{I}_i(v))}_{\text{communication}}$$
$$\times \underbrace{(\mathsf{M}_{-i}^{(t)} \mid \mathsf{M}_{P,i}^{(t)}, \mathsf{M}^{(<t)})}_{\text{publicly}}. \quad (9)$$

The first terms can be shown to be $o(1)$-close inductively (with base case being success of our simulation in the first round). The second terms are identical since the messages $M_{P,i}^{(t)}$ in π_r are simply generated simultaneously by each vertex $v \in \mathcal{P}_i$ looking at its own neighborhood $J_{i,*}(v), I_i(v)$ and the blackboard $M^{(<t)}$. For the last terms to be close, similar to (3) and (6), we need to bound the mutual information between $M_{-i}^{(t)}$ and $J_{i,*}, I_i$ at this point of the protocol, namely:

$$\|(\mathsf{M}_{-i}^{(t)} \mid \mathsf{M}_{P,i}^{(t)}, \mathsf{M}^{(<t)}) - (\mathsf{M}_{-i}^{(t)} \mid \mathsf{M}_{P,i}^{(t)}, \mathsf{M}^{(<t)}, \mathsf{J}_{i,*}, \mathsf{I}_i)\|_{\text{tvd}}^2$$
$$\leqslant \mathbb{I}(\mathsf{M}_{-i}^{(t)} ; \mathsf{J}_{i,*}, \mathsf{I}_i \mid \mathsf{M}_{P,i}^{(t)}, \mathsf{M}^{(<t)}). \quad (10)$$

Yet, while the RHS of this equation may seem similar to that of (7), this is a much more challenging term to bound as we shall discuss in the next subsection. For now, we only mention that our proof eventually bounds this information term *on average* for $i \in [p_r]$ with $o(1)$ which allows us to continue the simulation.

Having shown the $o(1)$-closeness of the distribution of π_r and the one used in our embedding, the proof ends as follows. The players of σ can continue running π_r by playing the role of principal block $\mathcal{P}_i$ in π_r explicitly with proper communication and keep sampling messages of remaining players as done in the embedding. At the end of the last round, they will obtain an almost faithful simulation of the entire protocol π_r which allows them to solve $I^\star = I_i$ as π_r likely needs to solve I_i for a random $i \in [p_r]$. This will then give us an $(r-1)$-round protocol for $I^\star$ which in turn allows us to use the inductive hardness of these instances to infer the lower bound for r-round protocols.

4) Idea Four: Bounding Gradual Correlation of Players' Inputs: The main technical part of our proof is to bound the information term in the RHS of (10), namely, the information other players can reveal about the input of principal block $\mathcal{P}_i$ in a single round. By the definition of $M_{-i}^{(t)} = (M_{P,-i}^{(t)}, M_F^{(t)})$ and chain rule, we have,

$$\text{RHS of (10)} = \mathbb{I}(\mathsf{M}_{P,-i}^{(t)} ; \mathsf{J}_{i,*}, \mathsf{I}_i \mid \mathsf{M}_{P,i}^{(t)}, \mathsf{M}^{(<t)})$$
$$+ \mathbb{I}(\mathsf{M}_F^{(t)} ; \mathsf{J}_{i,*}, \mathsf{I}_i \mid \mathsf{M}_P^{(t)}, \mathsf{M}^{(<t)}). \quad (11)$$

Recall that by the construction of the instance I, we have $\mathsf{J}_{i,*}, \mathsf{I}_i \perp \mathsf{J}_{-i,*}, \mathsf{I}_{-i}$. By the rectangle property of communication protocols, if the input of players are independent of each other, then even after communication, their corresponding input remains independent. *Assuming* we have this conditional independence here, one can easily prove both of the following properties:

$$\mathbb{I}(\mathsf{M}_{P,-i}^{(t)} ; \mathsf{J}_{i,*}, \mathsf{I}_i \mid \mathsf{M}_{P,i}^{(t)}, \mathsf{M}^{(<t)}) = 0,$$

and

$$\mathbb{E}_i[\mathbb{I}(\mathsf{M}_F^{(t)} ; \mathsf{J}_{i,*}, \mathsf{I}_i \mid \mathsf{M}_P^{(t)}, \mathsf{M}^{(<t)})]$$
$$\leqslant \frac{1}{p_r} \cdot \mathbb{I}(\mathsf{M}_F^{(t)} ; \mathsf{J}, \mathsf{I} \mid \mathsf{M}_P^{(t)}, \mathsf{M}^{(<t)})$$
$$\leqslant o(1). \quad \text{(similar to (7))}$$

So then what is the problem here? **Short answer: edge-sharing between the players!**

While $\mathsf{J}_{i,*}, \mathsf{I}_i \perp \mathsf{J}_{-i,*}, \mathsf{I}_{-i}$ is true initially, having fooling blocks that are able to see (subsets of) *both* these sets from the other endpoints, means that their messages can *correlate* these inputs as well. In other words, it can be that $\mathsf{J}_{i,*}, \mathsf{I}_i \not\perp \mathsf{J}_{-i,*}, \mathsf{I}_{-i} \mid \mathsf{M}_F^{(<t)}$ already from the second round. What is even more problematic is that even principal blocks in $\mathcal{P}_i$ and $\mathcal{P}_{-i}$ will see messages of these fooling blocks, so after the second round, even messages of other principal blocks correlate their originally independent inputs – more formally, this means that $\mathsf{J}_{i,*}, \mathsf{I}_i \not\perp \mathsf{J}_{-i,*}, \mathsf{I}_{-i} \mid \mathsf{M}_P^{(t)}$ (with no direct conditioning on fooling blocks' messages) can also happen after the second round!

The following example helps to motivate our approach.

Example. Consider the following two protocols:
- Protocol 1: in the second round, every principal block *except for* $\mathcal{P}_i$ sends XOR of their edges to fooling blocks[a] $J_{-i,*}$, while fooling blocks send XOR of all their edges in J.
- Protocol 2: in the second round, every principal block sends XOR of their edges in J while fooling blocks send XOR of all their edges in J.

In the first protocol, conditioned on $M_F^{(2)}$, the messages $M_{P,-i}^{(2)}$ reveal the XOR of edges in $J_{i,*}$, and thus the first mutual information term in (11) is 1 bit (note that here $M_{P,i}^{(2)} = \varnothing$).

In the second protocol, while $M_{P,-i}^{(2)}, M_F^{(2)}$ still reveal the

This example shows that one can have protocols that for *some* values of $i \in [p_r]$, principal blocks in $\mathcal{P}_{-i}$ can reveal non-trivial information about inputs of a principal block $\mathcal{P}_i$ also. But the given protocol (Protocol 1) is quite sensitive to the choice of index i, and for other indices $j \neq i$, this revealing of information no longer happens in this specific protocol. On the other hand, making the protocol less sensitive to the choice of i by "symmetrizing" the actions of players breaks its information-revealing property as players in $\mathcal{P}_i$ themselves will reveal the information offered by others. We exploit this by bounding the first term of (11) *on average* for $i \in [p_r]$. Note that this is precisely the step that our non-simultaneous simulation of a round, alluded to in Section III-B3, kicks in: the messages of $M_{P,-i}^{(2)}$ are still correlated heavily with $J_{i,*}, I_i$ even in Protocol 2; but conditioning on $M_{P,i}^{(2)}$ allows us to "break" this correlation and thus generate these messages even in the absence of public knowledge of $J_{i,*}, I_i$. We argue this is true for all protocols in the following.

To continue, by using chain rule on the first term of (11), we get that,

$$\mathbb{I}(M_{P,-i}^{(t)} ; J_{i,*}, I_i \mid M_{P,i}^{(t)}, M^{(<t)})$$
$$= \mathbb{I}(M^{(<t)}, M_P^{(t)} ; J_{i,*}, I_i) - \mathbb{I}(M^{(<t)}, M_{P,i}^{(t)} ; J_{i,*}, I_i) \quad (12)$$

where RHS is all the information revealed by the protocol about $J_{i,*}, I_i$ *minus* the information revealed already by players $\mathcal{P}_i$ and content of the blackboard. Now, in the absence of any conditioning, one can use the fact that $J_{i,*}, I_i \perp J_{-i,*}, I_i$ to bound:

$$\mathbb{E}_i[\mathbb{I}(M^{(<t)}, M_P^{(t)} ; J_{i,*}, I_i)] \leqslant o(1) + \frac{1}{p_r} \cdot \mathbb{I}(M_P^{(\leqslant t)} ; J, I \mid M_F^{(<t)}),$$

i.e., argue that fooling blocks can only reveal $o(1)$ bits about the input of an average principal block and the rest is the *average* information revealed by principal blocks themselves about the entire input. The second term of (12) is lower bounded by (via a simple application of chain rule and non-negativity of mutual information),

$$\mathbb{E}_i[\mathbb{I}(M^{(<t)}, M_{P,i}^{(t)} ; J_{i,*}, I_i)] \geqslant \mathbb{E}_i[\mathbb{I}(M_P^{(\leqslant t)} ; J_{i,*}, I_i) \mid M_F^{(<t)}].$$

Last step of the proof is to bound the second terms of the two equations above by showing that

$$\mathbb{I}(M_P^{(\leqslant t)} ; J, I \mid M_F^{(<t)}) \leqslant \sum_{i=1}^{p_r} \mathbb{I}(M_P^{(\leqslant t)} ; J_{i,*}, I_i \mid M_F^{(<t)}).$$

In words, this means that the total information revealed by principal blocks about the entire instance is bounded by the sum of the information revealed by them about each individual principal block's input $J_{i,*}, I_i$ for $i \in [p_r]$ *after* we condition on the messages of fooling blocks. This step requires a detailed calculation that at its core boils down to the fact that once we

condition on $M_F^{(<t)}$, we can "isolate" the information revealed by each message $M_{P,i}^{(t)}$ solely to $J_{i,*}, I_i$ – in other words, the principal blocks cannot generate correlation with other principal blocks' inputs on their own beyond what is already forced by fooling blocks.

Plugging in these bounds all together in (12) bounds the RHS by $o(1)$. A similar exercise, allows us to bound the second term in (11) by $o(1)$ also, which bounds the total information revealed about $J_{i,*}, I_i$ by players other than the ones in $\mathcal{P}_i$ by $o(1)$. This concludes the $o(1)$ bound on the mutual information term in (10), and implies the correctness of our simulation.

To conclude, we managed to simulate *all* rounds of π_r almost faithfully by continuing the embedding throughout the protocol and as a result solve the underlying instance $I^\star$ in $(r-1)$ rounds using a protocol with polylog(n)-size messages. We can now repeat this argument for $(r-1)$-round protocols and since in each recursion, the size of underlying instances drops by a factor of $\approx n^{1/5}$, we will end up with a non-trivial instance for any $r = o(\log \log n)$ that needs to be solved by a 0-round protocol – a contradiction that implies our desired lower bound.

IV. A Hard Distribution for Maximal Independent Set

The following is a formal restatement of Result 1.

Theorem 1 (Result 1, formal). *For $r \geqslant 0$ and any r-round multi-party protocol (deterministic or randomized) in the shared blackboard model for computing a maximal independent set on n-vertex graphs with constant error probability, there must exist some vertex communicating at least $\Omega(n^{1/20^{r+1}})$ bits in some round.*

In this section, we give a recursive definition of the hard distribution for maximal independent set that we are going to use for our proofs in Section V. The base case is the following hard distribution $\mathcal{D}_{\mathsf{MIS}}^{(0)}$ for protocols without any communication.

Distribution 1. The hard distribution $\mathcal{D}_{\mathsf{MIS}}^{(0)}$ for protocols computing a maximal independent set without any communication.
Parameters: bandwidth k, number of vertices $n_0 = 2k$.
1) Let E be an arbitrary, *fixed* perfect matching over n_0 vertices.
2) For $e \in E$, drop e with probability $1/2$ independently.
3) Return the graph G sampled above.

An immediate observation about $\mathcal{D}_{\mathsf{MIS}}^{(0)}$ is that any valid maximal independent set uniquely determines the set of matching edges that is dropped from E: for $e = (u, v) \in E$, e is dropped from E if and only if both of u, v are present in the maximal independent set. So for any deterministic referee, it can output a valid maximal independent set with probability at most 2^{-k}
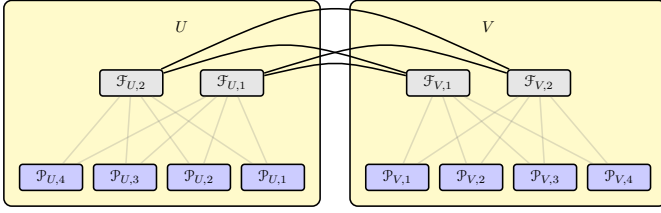
Fig. 3. An illustration of our lower bound instances for maximal independent set with parameters $\hat{f}_r = 2$ and $\hat{p}_r = 4$. The bottom vertices (blue) are principal blocks, while top vertices (gray) are fooling blocks. The heavy (solid black) edges fully connect fooling vertices from two "half instances" (yellow boxes). Note that these are the only edges across two "half instances". To find a maximal independent set in this graph, one needs to find maximal independent sets in all principal instances of at least one of "half instances".

over $\mathcal{D}_{\mathsf{MIS}}^{(0)}$ if it gets no information from the vertices. Note that this distributional bound naturally generalizes to randomized referees by an averaging argument, which is summarized in the following lemma.

Lemma IV.1 (Base Case). *Any 0-round protocol for computing a maximal independent set can only succeed with probability 2^{-k} over $\mathcal{D}_{\mathsf{MIS}}^{(0)}$.*

Building upon $\mathcal{D}_{\mathsf{MIS}}^{(0)}$, we construct the r-round hard distribution $\mathcal{D}_{\mathsf{MIS}}^{(r)}$ recursively. Assume we are given the $(r-1)$-round hard distribution $\mathcal{D}_{\mathsf{MIS}}^{(r-1)}$ over n_{r-1} vertices. The construction consists of two steps: first defining an auxiliary "half distribution" $\mathcal{H}_{\mathsf{MIS}}^{(r)}$ and then using $\mathcal{H}_{\mathsf{MIS}}^{(r)}$ to get the desired $\mathcal{D}_{\mathsf{MIS}}^{(r)}$, as shown below. The "half instances" roughly correspond to the hard instances we talk about in Section III. See Figure 3 for an illustration.

Distribution 2. The "half distribution" $\mathcal{H}_{\mathsf{MIS}}^{(r)}$ over graphs with vertex set V ($r \geqslant 1$).
Parameters: bandwidth k, number of **fooling blocks** $\hat{f}_r = k^6 \cdot n_{r-1}^3$, number of **principal blocks** $\hat{p}_r = k^6 \cdot n_{r-1}^3 \cdot \hat{f}_r$, number of vertices $\hat{n}_r = (n_{r-1} - 1) \cdot \hat{f}_r + n_{r-1} \cdot \hat{p}_r$, and vertex set V with $|V| = \hat{n}_r$.
  1) Partition V into disjoint sets of vertices $\mathcal{P}_1, \ldots, \mathcal{P}_{\hat{p}_r}, \mathcal{F}_1, \ldots, \mathcal{F}_{\hat{f}_r}$ such that $\forall i \in [\hat{p}_r]$: $|\mathcal{P}_i| = n_{r-1}$ and $\forall j \in [\hat{f}_r]$: $|\mathcal{F}_j| = n_{r-1} - 1$. Define $\mathcal{P}(V) := \bigcup_{i \in [\hat{p}_r]} \mathcal{P}_i$ and $\mathcal{F}(V) := \bigcup_{j \in [\hat{f}_r]} \mathcal{F}_j$.
  2) For $i \in [\hat{p}_r]$, sample an independent instance of $\mathcal{D}_{\mathsf{MIS}}^{(r-1)}$ on $\mathcal{P}_i$.
  3) For $u \in \mathcal{P}(V)$ and $j \in [\hat{f}_r]$, sample an independent instance of $\mathcal{D}_{\mathsf{MIS}}^{(r-1)}$ on $\mathcal{F}_j \cup \{u\}$ and only keep the edges adjacent to u (dropping all the edges between vertices in $\mathcal{F}_j$).
  4) Return the graph G sampled above.

Distribution 3. The hard distribution $\mathcal{D}_{\mathsf{MIS}}^{(r)}$ for r-round protocols computing a maximal independent set ($r \geqslant 1$).
Parameters: bandwidth k, number of **fooling blocks** $f_r = 2\hat{f}_r$, number of **principal blocks** $p_r = 2\hat{p}_r$, number of vertices $n_r = 2\hat{n}_r$.
  1) Let U and V be two disjoint sets of vertices, each of size $\hat{n}_r$. Sample two independent instances of $\mathcal{H}_{\mathsf{MIS}}^{(r)}$ on U and V.
  2) For $u \in \mathcal{F}(U)$ and $v \in \mathcal{F}(V)$, add an edge (u, v).
  3) Let G' be the graph sampled above. Sample a uniformly random permutation σ over $U \cup V$ and return $G = \sigma(G')$.

Remark IV.2. *A few remarks are in order.*

  1) *In the construction of the "half distribution" $\mathcal{H}_{\mathsf{MIS}}^{(r)}$, we call the sets of vertices $\mathcal{P}_1, \ldots, \mathcal{P}_{\hat{p}_r}$ the principal blocks, and the sets of vertices $\mathcal{F}_1, \ldots, \mathcal{F}_{\hat{f}_r}$ the fooling blocks. All vertices in $\mathcal{P}(V)$ and $\mathcal{F}(V)$ are the principal vertices and the fooling vertices, respectively.*

  2) *With a slight abuse of notation, we write $\sigma(\mathcal{P}_1), \ldots, \sigma(\mathcal{P}_{p_r})$ to denote all p_r principal blocks of $\sigma(U \cup V)$, and similarly $\sigma(\mathcal{F}_1), \ldots, \sigma(\mathcal{F}_{f_r})$ for all fooling blocks, in the construction of the hard distribution $\mathcal{D}_{\mathsf{MIS}}^{(r)}$.*

  3) *It is not hard to see that $n_r \leqslant k^{20^{r+1}}$ for $r \geqslant 0$. Indeed, $n_0 = 2k \leqslant k^{20}$ and by induction, the number of fooling blocks is $\hat{f}_r \leqslant k^6 \cdot k^{3 \cdot 20^r} \leqslant k^{9 \cdot 20^r}$, the number of principal blocks is $\hat{p}_r \leqslant k^6 \cdot k^{3 \cdot 20^r} \cdot \hat{f}_r \leqslant k^{18 \cdot 20^r}$, and thus $n_r \leqslant 2 \cdot 2 \cdot k^{20^r} \cdot \hat{p}_r \leqslant k^{20^{r+1}}$ for $r \geqslant 1$. Throughout the paper we assume the bandwidth parameter k is at least some sufficiently large constant.*

One important property about $\mathcal{D}_{\mathsf{MIS}}^{(r)}$, which justifies our use of two "half instances", is that any valid maximal independent set for G must also be maximal for the induced subgraph on either $\sigma(\mathcal{P}(U))$ or $\sigma(\mathcal{P}(V))$. The implication is that solving a hard instance drawn from $\mathcal{D}_{\mathsf{MIS}}^{(r)}$ requires to solve at least one of the "half instances" drawn from $\mathcal{H}_{\mathsf{MIS}}^{(r)}$. Formally, we have the following claim.

Claim IV.3. *Let Γ be any valid maximal independent set for a graph G drawn from $\mathcal{D}_{\mathsf{MIS}}^{(r)}$. Then at least one of the following must hold:*

  1) *$\Gamma \cap \sigma(\mathcal{P}(U))$ is a valid maximal independent set for the induced subgraph on $\sigma(\mathcal{P}(U))$.*

  2) *$\Gamma \cap \sigma(\mathcal{P}(V))$ is a valid maximal independent set for the induced subgraph on $\sigma(\mathcal{P}(V))$.*

Note that our construction in Distribution 3 has no edge between principal blocks, so Claim IV.3 further implies that solving an r-round instance requires to solve at least half of the principal $(r-1)$-round instances.

V. The Lower Bound for Maximal Independent Set

We prove the following theorem in this section. Theorem 1 is a straightforward corollary by an averaging argument, namely the easy direction of Yao's minimax principle [47]. Note that by the third statement of Remark IV.2, $n_r \leqslant k^{20^{r+1}}$ so we know $k \geqslant n_r^{1/20^{r+1}}$.

Theorem 2. *For $r = o(\log k)$, any r-round protocol for computing a maximal independent set that communicates at most k bits per vertex in every round can only succeed with probability less than 0.1 over $\mathcal{D}_{\mathsf{MIS}}^{(r)}$.*

Our proof to Theorem 2 for r-round protocols in general is by repeatedly applying the following round elimination lemma.

Lemma V.1 (Round Elimination). *For $r = o(\log k)$ and $\delta \in [0, 1]$, if there exists an r-round protocol for computing a maximal independent set that communicates at most k bits per vertex in every round and succeeds with probability δ over $\mathcal{D}_{\mathsf{MIS}}^{(r)}$, then there also exists an $(r-1)$-round protocol for computing a maximal independent set that communicates at most k bits per vertex in every round and succeeds with probability $\delta/2 - 1/n_{r-1}$ over $\mathcal{D}_{\mathsf{MIS}}^{(r-1)}$.*

Before proving Lemma V.1, which is the main part of this section, we first show it easily implies Theorem 2.

Proof of Theorem 2: Suppose for the purpose of contradiction that there exists an r-round protocol that communicates at most k bits per vertex in every round and that has success probability 0.1 over $\mathcal{D}_{\mathsf{MIS}}^{(r)}$. Applying Lemma V.1 for r times, we obtain a 0-round protocol having success probability

$$\frac{0.1}{2^r} - \sum_{t \in [r]} \frac{1}{2^{t-1} \cdot n_{t-1}} \geqslant \frac{0.1}{2^r} - \frac{1}{n_0} \cdot \sum_{t \in [r]} \frac{1}{2^{t-1}}$$
$$(\text{as } n_{t-1} \text{ is increasing})$$
$$\geqslant \frac{0.1}{2^r} - \frac{2}{n_0}$$
$$= \frac{1}{k^{o(1)}},$$

over $\mathcal{D}_{\mathsf{MIS}}^{(0)}$, where the last step follows from the assumption $r = o(\log k)$. Recall that $n_0 = 2k$ so the second term above is $\Theta(1/k)$ and can be ignored. However, the existence of such a 0-round protocol contradicts the lower bound of Lemma IV.1. This concludes the proof of the theorem. ∎

The proof of Lemma V.1 is deferred to the full version [30].

VI. The Lower Bound for Approximate Bipartite Matching

In this section we adapt the techniques for maximal independent set to prove the following formal version of Result 2.

Theorem 3 (Result 2, formal). *For $r \geqslant 0$ and any r-round multi-party protocol (deterministic or randomized) in the shared blackboard model for computing a maximal matching or any constant factor approximation to maximum matching on n-vertex (bipartite) graphs, there must exist some vertex communicating at least $\Omega(n^{1/20^{r+1}})$ bits in some round.*

Intuitively, Distribution 3 makes little use of any property specific to independent sets so most of our previous argument holds for matchings as well. Indeed, with minor adjustment to the hard distributions to better fit the need of approximation, almost the same ideas as in Sections IV and V yield the lower bound for approximate matching for *general* graphs. Meanwhile, there is a simple reduction to the bipartite case (by only considering edges across a random bipartition), concluding the proof of Theorem 3. Interested readers are referred to the full version [30] for formal proofs of both parts.

References

[1] A. Drucker, F. Kuhn, and R. Oshman, "On the power of the congested clique model," in *ACM Symposium on Principles of Distributed Computing, PODC '14, Paris, France, July 15-18, 2014*, M. M. Halldórsson and S. Dolev, Eds. ACM, 2014, pp. 367–376.

[2] F. Becker, P. Montealegre, I. Rapaport, and I. Todinca, "The impact of locality on the detection of cycles in the broadcast congested clique model," in *LATIN 2018: Theoretical Informatics - 13th Latin American Symposium, Buenos Aires, Argentina, April 16-19, 2018, Proceedings*, ser. Lecture Notes in Computer Science, M. A. Bender, M. Farach-Colton, and M. A. Mosteiro, Eds., vol. 10807. Springer, 2018, pp. 134–145.

[3] T. Jurdzinski, K. Lorys, and K. Nowicki, "Communication complexity in vertex partition whiteboard model," in *Structural Information and Communication Complexity - 25th International Colloquium, SIROCCO 2018, Ma'ale HaHamisha, Israel, June 18-21, 2018, Revised Selected Papers*, ser. Lecture Notes in Computer Science, Z. Lotker and B. Patt-Shamir, Eds., vol. 11085. Springer, 2018, pp. 264–279.

[4] T. Jurdzinski and K. Nowicki, "Connectivity and minimum cut approximation in the broadcast congested clique," in *Structural Information and Communication Complexity - 25th International Colloquium, SIROCCO 2018, Ma'ale HaHamisha, Israel, June 18-21, 2018, Revised Selected Papers*, ser. Lecture Notes in Computer Science, Z. Lotker and B. Patt-Shamir, Eds., vol. 11085. Springer, 2018, pp. 331–344.

[5] K. J. Ahn, S. Guha, and A. McGregor, "Analyzing graph structure via linear measurements," in *Proceedings of the Twenty-third Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '12. SIAM, 2012, pp. 459–467. [Online]. Available: http://dl.acm.org/citation.cfm?id=2095116.2095156

[6] ——, "Graph sketches: sparsification, spanners, and subgraphs," in *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2012, Scottsdale, AZ, USA, May 20-24, 2012*, 2012, pp. 5–14. [Online]. Available: http://doi.acm.org/10.1145/2213556.2213560

[7] S. Assadi, G. Kol, and R. Oshman, "Lower bounds for distributed sketching of maximal matchings and maximal independent sets," in *PODC '20: ACM Symposium on Principles of Distributed Computing, Virtual Event, Italy, August 3-7, 2020*, Y. Emek and C. Cachin, Eds. ACM, 2020, pp. 79–88.

[8] A. Filtser, M. Kapralov, and N. Nouri, "Graph spanners by sketching in dynamic streams and the simultaneous communication model," in *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, D. Marx, Ed. SIAM, 2021, pp. 1894–1913.

[9] K. J. Ahn, S. Guha, and A. McGregor, "Spectral sparsification in dynamic graph streams," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 16th International Workshop, APPROX 2013, and 17th International Workshop, RANDOM 2013, Berkeley, CA, USA, August 21-23, 2013. Proceedings*, 2013, pp. 1–10.

[10] M. Kapralov, Y. T. Lee, C. Musco, C. Musco, and A. Sidford, "Single pass spectral sparsification in dynamic streams," in *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, 2014, pp. 561–570. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2014.66

[11] S. Guha, A. McGregor, and D. Tench, "Vertex and hyperedge connectivity in dynamic graph streams," in *Proceedings of the 34th ACM Symposium on Principles of Database Systems, PODS 2015, Melbourne, Victoria, Australia, May 31 - June 4, 2015*, 2015, pp. 241–247.

[12] A. McGregor, D. Tench, S. Vorotnikova, and H. T. Vu, "Densest subgraph in dynamic graph streams," in *Mathematical Foundations of Computer Science 2015 - 40th International Symposium, MFCS 2015, Milan, Italy, August 24-28, 2015, Proceedings, Part II*, 2015, pp. 472–482.

[13] S. Assadi, Y. Chen, and S. Khanna, "Polynomial pass lower bounds for graph streaming algorithms," in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019.*, 2019, pp. 265–276.

[14] S. Assadi, P. Kumar, and P. Mittal, "Brooks' theorem in graph streams: A single-pass semi-streaming algorithm for Δ-coloring," *CoRR*, vol. abs/2203.10984, 2022.

[15] S. Lattanzi, B. Moseley, S. Suri, and S. Vassilvitskii, "Filtering: a method for solving graph problems in mapreduce," in *SPAA 2011: Proceedings of the 23rd Annual ACM Symposium on Parallelism in Algorithms and Architectures, San Jose, CA, USA, June 4-6, 2011 (Co-located with FCRC 2011)*, 2011, pp. 85–94.

[16] M. Kapralov and D. P. Woodruff, "Spanners and sparsifiers in dynamic streams," in *ACM Symposium on Principles of Distributed Computing, PODC '14, Paris, France, July 15-18, 2014*, 2014, pp. 272–281.

[17] K. J. Ahn, G. Cormode, S. Guha, A. McGregor, and A. Wirth, "Correlation clustering in data streams," in *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015*, ser. JMLR Workshop and Conference Proceedings, F. R. Bach and D. M. Blei, Eds., vol. 37. JMLR.org, 2015, pp. 2237–2246.

[18] F. Becker, M. Matamala, N. Nisse, I. Rapaport, K. Suchan, and I. Todinca, "Adding a referee to an interconnection network: What can(not) be computed in one round," in *25th IEEE International Symposium on Parallel and Distributed Processing, IPDPS 2011, Anchorage, Alaska, USA, 16-20 May, 2011 - Conference Proceedings*. IEEE, 2011, pp. 508–514.

[19] F. Becker, P. Montealegre, I. Rapaport, and I. Todinca, "The simultaneous number-in-hand communication model for networks: Private coins, public coins and determinism," in *Structural Information and Communication Complexity - 21st International Colloquium, SIROCCO 2014, Takayama, Japan, July 23-25, 2014. Proceedings*, ser. Lecture Notes in Computer Science, M. M. Halldórsson, Ed., vol. 8576. Springer, 2014, pp. 83–95.

[20] T. Jurdzinski, K. Lorys, and K. Nowicki, "Communication complexity in vertex partition whiteboard model," in *Structural Information and Communication Complexity - 25th International Colloquium, SIROCCO 2018, Ma'ale HaHamisha, Israel, June 18-21, 2018, Revised Selected Papers*, ser. Lecture Notes in Computer Science, Z. Lotker and B. Patt-Shamir, Eds., vol. 11085. Springer, 2018, pp. 264–279.

[21] J. Nelson and H. Yu, "Optimal lower bounds for distributed and streaming spanning forest computation," in *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019*, T. M. Chan, Ed. SIAM, 2019, pp. 1844–1860.

[22] H. Yu, "Tight distributed sketching lower bound for connectivity," in *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, D. Marx, Ed. SIAM, 2021, pp. 1856–1873.

[23] M. Luby, "A simple parallel algorithm for the maximal independent set problem," in *Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA*, R. Sedgewick, Ed. ACM, 1985, pp. 1–10.

[24] N. Linial, "Distributive graph algorithms-global solutions from local data," in *28th Annual Symposium on Foundations of Computer Science, Los Angeles, California, USA, 27-29 October 1987*. IEEE Computer Society, 1987, pp. 331–335.

[25] F. Kuhn, T. Moscibroda, and R. Wattenhofer, "Local computation: Lower and upper bounds," *J. ACM*, vol. 63, no. 2, pp. 17:1–17:44, 2016.

[26] M. Ghaffari, "An improved distributed algorithm for maximal independent set," in *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, R. Krauthgamer, Ed. SIAM, 2016, pp. 270–277.

[27] A. Balliu, S. Brandt, J. Hirvonen, D. Olivetti, M. Rabie, and J. Suomela, "Lower bounds for maximal matchings and maximal independent sets," in *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, D. Zuckerman, Ed. IEEE Computer Society, 2019, pp. 481–497.

[28] M. Ghaffari, T. Gouleakis, C. Konrad, S. Mitrovic, and R. Rubinfeld, "Improved massively parallel computation algorithms for mis, matching, and vertex cover," in *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing, PODC 2018, July 23-27, 2018*, 2018, pp. 129–138.

[29] N. Alon, N. Nisan, R. Raz, and O. Weinstein, "Welfare maximization with limited interaction," in *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, V. Guruswami, Ed. IEEE Computer Society, 2015, pp. 1499–1512.

[30] S. Assadi, G. Kol, and Z. Zhang, "Rounds vs communication tradeoffs for maximal independent sets," *arXiv preprint arXiv:2209.09049*, 2022.

[31] S. Dobzinski, N. Nisan, and S. Oren, "Economic efficiency requires interaction," in *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, 2014, pp. 233–242.

[32] M. Braverman and R. Oshman, "A rounds vs. communication tradeoff for multi-party set disjointness," in *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, 2017, pp. 144–155.

[33] S. Behnezhad, M. Hajiaghayi, and D. G. Harris, "Exponentially faster massively parallel maximal matching," in *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, D. Zuckerman, Ed. IEEE Computer Society, 2019, pp. 1637–1649.

[34] S. Assadi, S. Khanna, Y. Li, and G. Yaroslavtsev, "Maximum matchings in dynamic graph streams and the simultaneous communication model," in *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, 2016, pp. 1345–1364.

[35] G. Cormode, J. Dark, and C. Konrad, "Independent sets in vertex-arrival streams," in *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, 2019, pp. 45:1–45:14.

[36] J. Dark and C. Konrad, "Optimal lower bounds for matching and vertex cover in dynamic graph streams," in *35th Computational Complexity Conference, CCC 2020, July 28-31, 2020, Saarbrücken, Germany (Virtual Conference)*, ser. LIPIcs, S. Saraf, Ed., vol. 169. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020, pp. 30:1–30:14.

[37] V. Guruswami and K. Onak, "Superlinear lower bounds for multipass graph processing," in *Proceedings of the 28th Conference on Computational Complexity, CCC 2013, K.lo Alto, California, USA, 5-7 June, 2013*, 2013, pp. 287–298.

[38] S. Assadi and R. Raz, "Near-quadratic lower bounds for two-pass graph streaming algorithms," in *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, S. Irani, Ed. IEEE, 2020, pp. 342–353.

[39] L. Chen, G. Kol, D. Paramonov, R. R. Saxena, Z. Song, and H. Yu, "Almost optimal super-constant-pass streaming lower bounds for reachability," in *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, S. Khuller and V. V. Williams, Eds. ACM, 2021, pp. 570–583.

[40] S. Assadi, "A two-pass (conditional) lower bound for semi-streaming maximum matching," in *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, J. S. Naor and N. Buchbinder, Eds. SIAM, 2022, pp. 708–742.

[41] ——, "Combinatorial auctions do need modest interaction," in *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*, 2017, pp. 145–162.

[42] N. Nisan, "The demand query model for bipartite matching," in *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, D. Marx, Ed. SIAM, 2021, pp. 592–599.

[43] A. E. Roth and M. Sotomayor, "Two-sided matching," *Handbook of game theory with economic applications*, vol. 1, pp. 485–541, 1992.

[44] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.

[45] B. M. Maggs and R. K. Sitaraman, "Algorithmic nuggets in content delivery," *Comput. Commun. Rev.*, vol. 45, no. 3, pp. 52–66, 2015.

[46] T. M. Cover and J. A. Thomas, *Elements of information theory (2. ed.)*. Wiley, 2006.

[47] A. C. Yao, "Probabilistic computations: Toward a unified measure of complexity (extended abstract)," in *18th Annual Symposium on Foundations of Computer Science, Providence, Rhode Island, USA, 31 October - 1 November 1977*, 1977, pp. 222–227.