Electronic Colloquium on Computational Complexity, Report No. 54 (2015)

Welfare Maximization with Limited Interaction

Noga Alon * Noam Nisan † Ran Raz ‡ Omri Weinstein §

April 1, 2015

Abstract

We continue the study of welfare maximization in unit-demand (matching) markets, in a distributed information model where agent's valuations are unknown to the central planner, and therefore communication is required to determine an efficient allocation. Dobzinski, Nisan and Oren (STOC'14) showed that if the market size is n, then r rounds of interaction (with logarithmic bandwidth) suffice to obtain an $n^{1/(r+1)}$-approximation to the optimal social welfare. In particular, this implies that such markets converge to a stable state (constant approximation) in time logarithmic in the market size.

We obtain the first multi-round lower bound for this setup. We show that even if the allowable per-round bandwidth of each agent is $n^{\varepsilon(r)}$, the approximation ratio of any r-round (randomized) protocol is no better than $\Omega(n^{1/5^{r+1}})$, implying an $\Omega(\log\log n)$ lower bound on the rate of convergence of the market to equilibrium.

Our construction and technique may be of interest to round-communication tradeoffs in the more general setting of combinatorial auctions, for which the only known lower bound is for simultaneous ($r = 1$) protocols [DNO14].

*Tel Aviv University and Microsoft Research. Research supported in part by BSF grant 2012/107, by ISF grant 620/13, and by the Israeli I-Core program. nogaa@post.tau.ac.il.

†Department of Computer Science, Hebrew University, and MSR Israel. Supported by ISF grants 230/10 and 1435/14 administered by the Israeli Academy of Sciences. noamn@microsoft.com.

‡Weizmann Institute of Science, Israel, and the Institute for Advanced Study, Princeton, NJ. Research supported by the Israel Science Foundation grant No. 1402/14, by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation, by the Simons Foundation, by the Fund for Math at IAS, and by the National Science Foundation grant No. CCF-1412958. ran.raz@weizmann.ac.il.

§Department of Computer Science, Princeton University. Research Supported by a Simons Fellowship award in TCS and a Siebel scholarship. oweinste@cs.princeton.edu.

ISSN 1433-8092

1 Introduction

This paper studies the tradeoff between the amount of communication and the number of rounds of interaction required to find an (approximately) optimal matching in a bipartite graph. In our model there are n "players" and m "items". Each player initially knows a subset of the items to which it may be matched (i.e. m bits of information). The players communicate in rounds: in each round each player writes a message on a shared blackboard. The message can only depend on what the player knows at that stage: his initial input and all the messages by all other players that were written on the blackboard in previous rounds.

This problem was recently introduced by [DNO14] as a simple market scenario: the players are unit-demand bidders and our goal is to find an (approximately) welfare-maximizing allocation of the items to players. The classic auction of [DGS86] – that may be viewed as a simple Walrasian-like market process for this setting – can be implemented as to find an approximately optimal allocation where each player needs only send $O(\log n)$ bits of communication (on the average). The question considered by [DNO14] was whether such a low communication burden suffices without using multiple rounds of interaction. As a lower bound, they proved that a non-interactive protocol, i.e. one that uses a *single round of communication*, cannot get a $n^{1/2-\epsilon}$-factor approximation (for any fixed $\epsilon > 0$) with $n^{o(1)}$ bits of communication per player. As upper bounds they exhibited (I) an $O(\log n)$-round protocol, where each player sends $O(\log n)$ bits per round, that gets a $\frac{1}{1-\delta}$-factor approximation (for any fixed $\delta > 0$) and (II) for any fixed $r \geq 1$, an r-round protocol, where each player sends $O(\log n)$ bits per round, that gets an $O(n^{1/(r+1)})$-approximation.

The natural question at this point is whether there are r-round protocols with better approximation factors that still use $n^{o(1)}$ bits of communication per player. This question was left open in [DNO14], where it was pointed out that it was even open whether the exactly optimal matching can be found by 2-round protocols that use $O(\log n)$ bits of communication per player. We answer this open problem by proving lower bounds for any fixed number of rounds.

Theorem: For every $r \geq 1$ there exists $\epsilon(r) = \exp(-r)$, such that every (deterministic or randomized) r-round protocol requires $n^{\epsilon(r)}$ bits of communication per player in order to find a matching whose size is at least $n^{-\epsilon(r)}$ fraction of the optimal matching.

Our Techniques. We construct a recursive family of hard distributions for every fixed number of communication rounds, and use information theoretic machinery to analyze it. Our proof uses a type of direct-sum based round-reduction argument for multiparty communication complexity. Unlike standard round-elimination arguments in the two-party model, our instance size (and thus the number of players) scales with the number of allowable rounds, and therefore eliminating a communication round essentially requires embedding a "low dimensional" instance (with fewer players) into a "higher dimensional" protocol (operating over a larger input), from its second round onwards. In order to carry out such an embedding, we need a way of sampling the rest of the inputs to the higher dimensional protocol (including the remaining players) *conditioned on the first message of the protocol*, with no extra communication. The main obstacle is that conditioning on the first message of the "high dimensional" protocol *correlates the private inputs of the players* (i.e., the inputs to the "lower dimensional" protocol) *with the "missing" inputs*, and it is not hard to see that, in general, such sampling cannot be done without communication! Circumventing this major obstacle calls for a subtle construction and analysis, which ensures the aforementioned correlations remain "local" and therefore allows to perform the embedding using a combination of private and public randomness. Our constructed family of distributions is designed to facilitate such embedding (using certain conditional independence properties) on one hand, and yet retain a "marginal indistinguishability" property which is essential to keep the information argument above valid (we discuss this further in Section 4).

1.1 More context and related models

The bipartite matching problem is clearly a very basic one and obviously models a host of situations beyond the economic one that was the direct motivation of [DNO14] and of this paper. Despite having been widely studied, even its algorithmic status is not well understood, and it is not clear whether a nearly-linear time algorithm exists for it. (The best known running time (for the dense case) is the 40-year old $O(n^{2.5})$

algorithm of [HK73], but for special cases like regular or near-regular graphs nearly linear times are known (e.g. [Alo03, Yus13])). In parallel computation, a major open problem is whether bipartite matching can be solved in deterministic parallel poly-logarithmic time with a polynomial number of processors (Randomized parallel algorithms for the problem [MVV87, KUW85] have been known for over 25 years). It was suggested in [DNO14] that studying the problem in the communication complexity model is an approach that might lead to algorithmic insights as well.

The bipartite matching problem has been studied in various other multi-party models that focus on communication as well. In particular, strong and tight bounds for approximate matching are known in the weaker "message passing" or "private channels" models [HRVZ13] that have implications to models of parallel and distributed computation . Related work has also been done in networked distributed computing models, e.g., [LPSP08]. "One-way" communication models are used to analyze streaming or semi-streaming models and some upper bounds (e.g., [Kap12]) as well as weak lower bounds [GKK12] are known for approximate matchings in these models. For "r-way" protocols, a super-linear communication lower bound was recently shown by [GO13] for *exact* matchings, in an incomparable model[1]. A somewhat more detailed survey of these related models can be found in the appendix of [DNO14].

It should be noted that the open problems mentioned above remain so even in the standard two-party setting where each of the two players holds all the information of $n/2$ of our players. We do not know any better upper bounds than what is possible in the multi-player model, and certainly, as the model is stronger, no better lower bounds are known. We also do not know whether our lower bound (or the single round one of [DNO14]) applies also in this stronger two-player model.

1.2 Open problems

There are many open problems related to our work. Let us mention a few of the most natural ones. Our first open problem is closing the gap between our lower bound and the upper bound: We show that $r = \Omega(\log \log n)$ rounds of communication are required to achieve constant approximation ratio using poly-logarithmic bits per player, while the upper bound is $r = O(\log n)$. We believe that the upper bound is in fact tight, and improving the lower bound is left as our first and direct open problem.

Another interesting direction is trying to extend our lower bound technique to obtain similar-in-spirit round-communication tradeoffs for the more general setup of combinatorial auctions, also studied by [DNO14]. From a communication complexity perspective, lower bounds in this setup are more compelling, since player valuations require *exponentially* many bits to encode, hence interaction has the potential to reduce the overall communication (required to obtain efficient allocations) from exponential to polynomial. Indeed, it is shown in [DNO14] that, in the case of *sub-additive bidders*, there is an r-round randomized protocol that obtains an $\tilde{O}(r \cdot m^{1/(r+1)})$-approximation to the optimal social welfare, where in each round each player sends $poly(m, n)$ bits. Once again, an (exponential in m) lower bound on the communication was given only for the case of simultaneous protocols ($r = 1$) and the natural question is to extend it to multiple rounds as well.

A more general open problem advocated by [DNO14] is to analyze the communication complexity of finding an *e*xact optimal matching. One may naturally conjecture that $n^{\Omega(1)}$ rounds of interaction are required for this if each player only sends $n^{o(1)}$ bits in each round, but no super-logarithmic bound is known. The communication complexity of the problem without any limitation on the number of rounds is also open: no significantly super linear, $\omega(n \log n)$, bound is known, while the best upper bound known is $\tilde{O}(n^{3/2})$.

[1] Besides of the fact that the lower bound in [GO13] applies only for testing *exact* matchings and not approximate matchings, their model consists of a small number of parties (*constant or logarithmic* in n) who are communicating in some fixed number of *sequential* rounds (not simultaneous). The input itself of each player is therefore *super-linear* in the number of nodes of the input graph (n), and indeed they prove a super-linear communication lower bound, which is clearly impossible in our model. The [GO13] model was motivated by streaming lower bounds and does not seem to capture the economic scenario we attempt to model in this paper (i.e., that of private-valuations) and therefore this result is incomparable to ours, as also evidenced by the distinct proof-techniques.

2 preliminaries

We reserve capital letters for random variables, and calligraphic letters for sets. The ℓ_1 (statistical) distance between two distributions in the same probability space is denoted $|\mu - \nu| := \frac{1}{2} \cdot \sum_a |\mu(a) - \nu(a)|$. We write $X \perp Y \mid Z$ to denote that X and Y are statistically independent conditioned on the random variable Z. For a vector random variable $X = X_1 X_2 \ldots X_s$, we sometimes use the shorthands $X_{\leq i}$ and X_{-i} to denote $X_1 X_2 \ldots X_i$ and $X_1 X_2 \ldots X_{i-1}, X_{i+1}, \ldots \ldots X_s$ respectively (similarly, $X^{-i} := X^1 X^2 \ldots X^{i-1} X^{i+1} \ldots X^s$). We write $A \in_R \mathcal{U}$ to denote a uniformly distributed random variable over the set $\mathcal{U}$. We use the terms "bidders" and "players" interchangeably throughout the paper.

2.1 Communication Model

Our framework is the *number-in-hand* (NIH) multiparty communication complexity model with shared blackboard. In this model, n players receive inputs $(x_1, x_2, \ldots, x_n) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \mathcal{X}_n$ respectively. In our context, each of the n players (bidders) is associated with a node $u \in U = [n]$ of some bipartite graph $G = (U, V, E)$, and her input is the set of incident edges on her node (her demand set of items in $V = [m]$). The players' goal is to compute a maximum set of disjoint connected pairs $(u, v) \in E(G)$, i.e., a maximum matching in G (we define this formally below).

The players communicate in some fixed number of rounds r, where in each communication round, players *simultaneously* write (at most) ℓ bits each on a *shared blackboard* which is viewable to all parties. We sometimes refer to the parameter ℓ as the *bandwidth* of the protocol. In a deterministic protocol, each player's message should be completely determined by the content of the blackboard and her own private input x_i. In a randomized protocol, the message of each player may further depend on both public and private random coins. When player's inputs are distributional ($(x_1, x_2, \ldots, x_n) \sim \mu$) which is the setting in this paper, we may assume without loss of generality that the protocol is *deterministic*, since the averaging principle asserts that there is always some fixing of the randomness that will achieve the same performance with respect to μ. We remark that by the averaging principle, our main result applies to the randomized setting as well[2].

The transcript of a protocol π (namely, the content of the blackboard) when executed on an input graph G is denoted by $\Pi(G)$, or simply Π when clear from context. At the end of the r'th communication round, a *referee* (the "central planner" in our context) computes a matching $\hat{\mathcal{M}}(\Pi)$, which is completely determined by Π. We call this the *output* of the protocol.

We will be interested in protocols that compute *approximate matchings*. To make this more formal, let $\mathcal{G}(n, m)$ denote the family of bipartite graphs on (n, m)-vertex sets respectively, and denote by $\mathcal{F}(n, m)$ the family of all matchings in $\mathcal{G}(n, m)$ (not necessarily maximum matchings). Denote by $|\mathcal{M}(G)|$ the size of a maximum matching in the input graph G. We require that the output of any protocol satisfies $\hat{\mathcal{M}}(\Pi) \in \mathcal{F}(n, m)$. The following definition is central to this work.

Definition 2.1 (Approximate Matchings). *We say that a protocol π computes an α-*approximate matching *($\alpha \geq 1$) if $|\hat{\mathcal{M}}(\Pi) \cap E(G)|$ is at least $\frac{1}{\alpha} \cdot |\mathcal{M}(G)|$, i.e., if the number of matched pairs $(u, v) \in E(G)$ is at least a $(1/\alpha)$-fraction of the maximum matching in G. Similarly, when the input graph G is distributed according to some distribution μ (i.e., $(x_1, x_2, \ldots, x_n) \sim \mu$), we say that the* approximation ratio *of π is α if*

$$\mathop{\mathbb{E}}_{G \sim \mu}[|\hat{\mathcal{M}}(\Pi) \cap E(G)|] \geq \frac{1}{\alpha} \cdot \mathop{\mathbb{E}}_{G \sim \mu}[|\mathcal{M}(G)|].$$

The *expected matching size* of π is $\mathbb{E}_\mu[|\hat{\mathcal{M}}(\Pi) \cap E(G)|]$ (we remark that the "hard" distribution we construct in the next section will satisfy $|\mathcal{M}(G)| \equiv n$ for all G in the support of μ, so the quantity $\mathbb{E}_{G \sim \mu}[|\mathcal{M}(G)|]$ will always be n). Note that these definitions in particular allow the protocol to be *erroneous*, i.e., the referee is allowed to output "illegal" pairs $(u, v) \notin E(G)$, but we only count the correctly matched pairs. Our lower bound holds even with respect to this more permissive model.

[2] More formally, if there is a distribution μ on players inputs such that the approximation ratio of any r-round *deterministic* protocol with respect to μ is at most α in expectation, then fixing the randomness of the protocol would yield a deterministic protocol with the same performance, thus the former lower bound applies to randomized r-round protocols as well.

2.2 Information theory

Our proof relies on basic concepts from information theory. For a broader introduction to the field, and proofs of the claims below, we refer the reader to the excellent monograph of [CT91].

For two distributions μ and ν in the same probability space, the *Kullback-Leiber* divergence between μ and ν is defined as

$$\mathbb{D}\left(\mu(a)\|\nu(a)\right) := \mathbb{E}_{a\sim\mu}\left[\log\frac{\mu(a)}{\nu(a)}\right]. \tag{1}$$

The following well known inequality upper bounds the statistical distance between two distributions in terms of their KL Divergence:

Lemma 2.2 (Pinsker's inequality)**.** *For any two distributions μ and ν,*

$$|\mu(a)-\nu(a)|^2 \leq \frac{1}{2}\cdot\mathbb{D}\left(\mu(a)\|\nu(a)\right).$$

A related measure which is central to this paper is that of *mutual information*, which captures correlation between random variables.

Definition 2.3 (Conditional Mutual Information)**.** *Let A,B,C be jointly distributed random variables. The* Mutual Information *between A and B conditioned on C is*

$$I(A;B|C) := \mathop{\mathbb{E}}_{\mu(cb)}\mathbb{D}\left(\mu(a|bc)\|\mu(a|c)\right) = \mathop{\mathbb{E}}_{\mu(ca)}\mathbb{D}\left(\mu(b|ac)\|\mu(b|c)\right) = \sum_{a,b,c}\mu(abc)\log\frac{\mu(a|bc)}{\mu(a|c)}.$$

The above definition can be interpreted as follows: $I(A;B|C)$ is large if the distribution $(A|B=b,C=c)$ is "far" from $(A|C=c)$ for typical values of b,c, which means that B provides a lot of information about A conditioned on C. We note that an equivalent, more intuitive definition of (conditional) mutual information is $I(A;B|C) = H(A|C) - H(A|BC)$, where $H(A|C)$ is the (expected) *Shannon Entropy* of the random variable A conditioned on C. Thus A and B have large mutual information conditioned on C, if further conditioning on B significantly reduces the entropy of A. We prefer Definition 2.3 as it is more appropriate for our proof, but we note that the latter one immediately implies

Fact 2.4. $I(A;C|D) \leq H(A|D) \leq H(A) \leq |A|$*,*

where the last term denotes the cardinality of $\log|\mathsf{Supp}(A)|$ of the random variable A, and the second transition follows since conditioning never increases entropy.

The most important property of mutual information is that it satisfies the following chain rule:

Fact 2.5 (Chain rule for mutual information)**.** *Let A,B,C,D be jointly distributed random variables. Then $I(AB;C|D) = I(A;C|D) + I(B;C|AD)$.*

Lemma 2.6 (Conditioning on independent variables increases information)**.** *Let A,B,C,D be jointly distributed random variables. If $I(A;D|C)=0$, then it holds that $I(A;B|C) \leq I(A;B|CD)$.*

Proof. We apply the chain rule twice. On one hand, we have $I(A;BD|C) = I(A;B|C) + I(A;D|CB) \geq I(A;B|C)$, since mutual information is nonnegative. On the other hand, $I(A;BD|C) = I(A;D|C) + I(A;B|CD) = I(A;B|CD)$, since $I(A;D|C)=0$ by assumption. Combining both equations completes the proof. □

On the other hand, the following lemma asserts a condition under which conditioning *decreases* information:

Lemma 2.7. *Let A,B,C,D be jointly distributed random variables such that $I(B;D|AC)=0$. Then it holds that $I(A;B|C) \geq I(A;B|CD)$.*

Proof. Once again, we apply the chain rule twice. We have $I(A;B|CD) = I(AD;B|C) - I(D;B|C) = I(A;B|C) + I(D;B|AC) - I(D;B|C) = I(A;B|C) - I(D;B|C) \leq I(A;B|C)$. □

Fact 2.8 (Data processing inequality, general case)**.** *Let $X \to Y \to Z$ be a Markov chain ($I(X;Z|Y) = 0$). Then $I(X;Z) \le I(X;Y)$.*

Fact 2.9 (Data processing inequality, special case)**.** *Let A, B, C be three jointly distributed random variables, where the domain of B is Ω, and let $f : \Omega \longrightarrow \mathcal{U}$ be any deterministic function. Then $I(A;B|C) \ge I(A;f(B)|C)$.*

Fact 2.10. *Let μ and ν be two probability distributions over a non-negative random variable X, whose value is bounded by X_{max}. Then $\mathbb{E}_\nu[X] \le \mathbb{E}_\mu[X] + |\mu - \nu| \cdot X_{max}$.*

3 A hard distribution for r-round protocols

We begin by defining a family of hard distributions for protocols with r rounds. Recall that $\mathcal{G}(n,m)$ is the family of bipartite graphs on (n,m) vertex-sets. For any given number of rounds r, we define a hard distribution μ_r on bipartite graphs in $\mathcal{G}(n_r, m_r)$. μ_r is recursively defined in Figure 1.

A recursive definition of the hard distribution μ_r

In what follows, ℓ is a parameter (denoting the bandwidth of the communication channel).

1. For $r = 0$, $G^0 = (U^0, V^0, E^0)$ consists of a set of n_0 bidders $U^0 = \{b_1, \dots, b_{n_0}\}$ and a set of m_0 items $V^0 = \{j_1 \dots, j_{m_0}\}$, such that $n_0 = m_0 = \ell^5$. E^0 is then obtained by selecting a random permutation $\sigma \in_R S_{\ell^5}$ and connecting $(b_i, j_{\sigma(i)})$ by an edge. This specifies μ_0.

2. For any $r \ge 0$, the distribution μ_{r+1} over $G^{r+1} = (U^{r+1}, V^{r+1}, E^{r+1})$ is defined as follows:
   Vertices:
   - The set of bidders is $U^{r+1} := \bigcup_{i=1}^{n_r^4} B_i$ where $|B_i| = n_r$. Thus, $n_{r+1} = n_r^5$.
   - The set of items is $V^{r+1} := \bigcup_{j=1}^{n_r^4 + \ell\cdot n_r^2} T_j$ where $|T_j| = m_r$. Thus, $m_{r+1} = (n_r^4 + \ell \cdot n_r^2) \cdot m_r$.

   Edges: Let d_r be the degree of each vertex (bidder) in the graph G^r (this is well defined as, by induction, it holds that the degree of any vertex is fixed for every graph in the support of μ_r). The distribution on edges is obtained by first choosing $\ell \cdot n_r^2$ random indices $\{a_1, a_2, \dots a_{\ell\cdot n_r^2}\}$ from $[n_r^4 + \ell \cdot n_r^2]$, and a random invertible map $\sigma : [n_r^4] \longrightarrow [n_r^4 + \ell \cdot n_r^2] \setminus \{a_1, a_2, \dots a_{\ell\cdot n_r^2}\}$. Each bidder $u \in B_i$ is connected to d_r random items *in each one* of the blocks $T_{a_1}, T_{a_2}, \dots, T_{a_{\ell\cdot n_r^2}}$, using independent randomness for each of the blocks and for each bidder. The entire block B_i is further connected to the entire block $T_{\sigma(i)}$ *using an independent copy of the distribution* μ_r. Note that this is well defined, as $|B_i| = n_r$, $|T_{\sigma(i)}| = m_r$ and μ_r is indeed a distribution on bipartite graphs from $\mathcal{G}(n_r, m_r)$.

Figure 1: A hard distribution for r-round protocols.

Remark 3.1. *A few remarks are in order:*

(i) As standard, the input of each bidder $u \in U^{r+1}$ is the set of incident edges on the vertex u (defined by μ_{r+1}). Note that every graph in the support of μ_{r+1} has a perfect matching ($|\mathcal{M}(G^{r+1})| = n_{r+1}$).

(ii) It is easy to see by induction that : (a) $n_r = \ell^{5^{r+1}}$; and (b) $m_r \le n_r^2$.
(Proof of (b): By induction on r, $m_{r+1} := (n_r^4 + \ell \cdot n_r^2) \cdot m_r \le (n_r^4 + \ell \cdot n_r^2) \cdot n_r^2 \le 2n_r^6 < n_{r+1}^2$).

(iii) Note that in μ_{r+1}, each block of bidders B_i is connected to its "hidden item block" $T_{\sigma(i)}$ using a copy of the joint distribution μ_r, and to each of the "fooling item blocks" T_{a_j}, using the product of the marginals *of μ_r, i.e., according to $\times_{u \in n_r}(\mu_r|u)$. This property will be crucial.*

(iv) Throughout the paper, we assume the bandwidth parameter ℓ is larger than some large enough absolute constant (note that by (ii) above, in fact $\ell = \omega_r(1)$).

Notation. To facilitate our analysis, the following notation will be useful. Notice that each block B_i of players is connected to exactly $\ell \cdot n_r^2 + 1$ blocks of items whose indices we denote by

$$\mathcal{I}_i := \{\sigma(i), a_1, a_2, \ldots a_{\ell \cdot n_r^2}\}.$$

For each B_i, let $\tau_i : \mathcal{I}_i \longrightarrow [\ell \cdot n_r^2 + 1]$ be the bijection that maps any index in $\mathcal{I}_i$ to its location in the sorted list of $\mathcal{I}_i$ (i.e., $\tau_i^{-1}(1)$ is the smallest index in $\mathcal{I}_i$, $\tau_i^{-1}(2)$ is the second smallest index in $\mathcal{I}_i$ and so forth). We henceforth denote by G_j^i the (induced) subgraph of $G = G^{r+1}$ on the sets $(B_i, T_{\tau_i^{-1}(j)})$, for each $j \in [\ell \cdot n_r^2 + 1]$. By a slight abuse of notation, we will sometimes write $G_j^i = (B_i, T_{\tau_i^{-1}(j)})$ to denote the specific set of edges of G_j^i. Similarly, for a bidder $u \in B_i$, let $G_j^u = (u, T_{\tau_i^{-1}(j)})$ denote the (induced) subgraph of G on the sets $(u, T_{\tau_i^{-1}(j)})$. In this notation, the entire input of a player $u \in B_i$ is $\Gamma_u := \{G_1^u, G_2^u, \ldots, G_{\ell \cdot n_r^2+1}^u\}$. Let

$$J_i := \tau_i(\sigma(i))$$

denote the index of the "hidden graph" $G_{J_i}^i = (B_i, T_{\sigma(i)})$. To avoid confusion (with the other indices j), we henceforth write

$$G(J_i) := G_{J_i}^i.$$

Note that by symmetry of our construction, the index J_i is uniformly distributed in $[\ell \cdot n_r^2 + 1]$. The following fact will be crucial to our analysis:

Fact 3.2 (Marginal Indistinguishability)**.** *For any bidder $u \in B_i$, it holds that $I(\Gamma_u; J_i \mid \mathcal{I}_i) = 0$.*

Proof. Recall that $\Gamma_u = \{G_1^u, G_2^u, \ldots, G_{\ell \cdot n_r^2+1}^u\}$ is the input of bidder u. The claim follows directly from property (iii) in Remark 3.1, since by definition of our construction, the distribution of edges of $G_j^u = (u, T_{\tau_i^{-1}(j)})$ is $(\mu_r|u)$ for all $j \in [\ell \cdot n_r^2 + 1]$. We remark that the above fact implies that, up to a permutation on the names of the items in V^{r+1}, $G_j^u \sim G_k^u$ for any bidder $u \in B_i$ and any $j \neq k \in [\ell \cdot n_r^2 + 1]$. ☐

Finally, Let $\mathcal{B}$ denote the partition of bidders in $U := U^{r+1}$ into the blocks B_i, and $\mathcal{T}$ denote the partition of items in $V := V^{r+1}$ into the blocks T_j. Throughout the proof, we think of $\mathcal{T}$ and $\mathcal{B}$ as fixed, while we think of the names of the bidders in each block of $\mathcal{B}$ and items in each block of $\mathcal{T}$ as random. Since $\mathcal{T}$ and $\mathcal{B}$ are fixed (publicly known) in the distribution μ_{r+1}, our entire analysis is performed under the implicit conditioning on $\mathcal{T}, \mathcal{B}$. Note that $\mathcal{T}$ does not reveal the identity of the "fooling blocks" T_{a_j}, but only the items belonging to each block.

4 The lower bound

In this section we prove our main result. Recall that the expected matching size of π (with respect to μ) is $\mathbb{E}_\mu[|\hat{\mathcal{M}}(\Pi) \cap E(G)|]$. We shall prove the following theorem.

Theorem 4.1 (Main Result)**.** *The expected matching size of any r-round protocol under μ_r is at most $5n_r^{1-1/5^{r+1}}$. This holds as long as the number of bits sent by each player at any round is at most $\ell = n_r^{1/5^{r+1}}$. In particular, since μ_r has a perfect matching, the approximation ratio of any r-round protocol is no better than $\Omega\left(n^{1/5^{r+1}}\right)$.*

The intuition behind the proof is as follows. Consider some $(r+1)$-round protocol π (with bandwidth ℓ), and let $M_{B_i} = M^1_{B_i} M^2_{B_i}, \ldots, M^{n_r}_{B_i}$ denote the (concatenated) messages sent by all of the bidders in a block B_i in the first round of π. From this point on, we will assume that π is a deterministic protocol (since by the averaging principle we may fix its randomness without harming the performance). Informally speaking, the distribution μ_{r+1} is designed so that messages of bidders in B_i ($M^u_{B_i}$) convey little information about the "hidden" graph $G(J_i)$. Intuitively, this will be true since the *marginal* distribution of the hidden graph $G^u_{J_i}$ for any bidder $u \in B_i$ is indistinguishable from the rest of the "fooling graphs" (Fact 3.2) and therefore a bidder in B_i will not be able to distinguish between vertices (items) in $\bigcup_{j=1}^{\ell \cdot n_r^2} T_{a_j}$ and in $T_{\sigma(i)}$. Using the conditional independence properties of the distribution μ_{r+1} and the simultaneity of the protocol, we will show that the latter condition also implies that the *total* information conveyed by M_{B_i} on $G(J_i)$ is small. In order to make this information $\ll 1$ bit, the parameters are chosen so that n_r grows doubly-exponentially in r ($n_r = \ell^{5^{r+1}}$), and this choice is the cause for the approximation ratio we eventually obtain. Intuitively, the fact that little information is conveyed by each block on the "hidden graph" implies that the distribution of edges in the graph $G(J_i)$ is still close to μ_r even *conditioned* on the first message of the i'th block M_{B_i}. Now suppose an $(r+1)$-round protocol finds a large matching with respect to the original distribution μ_{r+1} (in expectation). Then the expected induced matching size on $G(J_i)$ must be large on average as well. Hence, "ignoring" the first round of the protocol, we would like to argue that the original protocol essentially induces an r-round protocol for finding a large matching with respect to the distribution μ_r, up to some error term (indeed, *some* information about $G(J_i)$ may have already been discovered in the first round of the protocol, but the argument above ensures that this information is small). Doing so essentially reduced the problem to finding a large matching under μ_r using only r rounds, so we may use an inductive approach to upper bound the latter expected matching size.

Making the latter intuition precise is complicated by the fact that, unlike standard "round-elimination" arguments in the two-party setting, in our setup one cannot simply "project" an r-round n_{r+1}-party protocol (with inputs $\sim \mu_{r+1}$) directly to the distribution μ_r, since a protocol for the latter distribution has only n_r players (inputs). To remedy this, we crucially rely on the conditional independence properties of our construction (Lemma 4.6 below) together with an embedding argument to obtain the desired lower bound.

The embedding part of the proof (Claim 4.7) is subtle, since in general, conditioning on the first message M_1 correlates the (private) inputs of the players with the "missing" inputs to the "higher-dimensional" protocol (the "fooling item blocks" of μ_{r+1}), so it is not clear how the players can sample these "missing" inputs without communicating. Luckily and crucially, the edges to the "fooling blocks" T_{a_j} in μ_{r+1} were chosen independently for each bidder $u \in U$ (unlike the hidden graphs $G(J_i)$ in which players have correlated edges). This independence is what allows to embed a lower-dimensional graph $H \sim \mu_r$ and "complete" the rest of the graph (using a combination of public and private randomness) according to the conditional distribution $(G|M_1, H)$ without any communication, thus "saving" one round of communication.

We now turn to formalize the above intuition. From this point on, let us use the shorthands

$$\mathbf{J} := J_1, \ldots, J_{n_r^4} \ , \ \mathcal{I} := \mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_{n_r^4}.$$

Also, for the remainder of the proof, let us define for simplicity

$$\Delta_r := \frac{1}{n_r} \ .$$

Let π be an $(r+1)$-round deterministic protocol. For a given message $m_{B_i} := m^1_{B_i} m^2_{B_i}, \ldots, m^{n_r}_{B_i}$ sent in π by the bidders in block B_i in the first round of π, a fixing of the index $J_i = j_i$ of the "hidden" block of items, and of the partition $\mathcal{I}_i$, let

$$\psi^i_r := (G(J_i) \mid M_{B_i} = m_{B_i}, J_i = j_i, \mathcal{I}_i)$$

denote the distribution of the "hidden graph" $G(J_i)$ conditioned on $M_{B_i}, \mathcal{I}_i$ and J_i. The following lemma asserts that, in expectation over the first communication round of π, the marginal distribution of $G(J_i)$ is very close to its original distribution μ_r.

Lemma 4.2. *For every $i \in [n_r^4]$,*

$$\mathop{\mathbb{E}}_{m_{B_i}, \mathcal{I}_i, j_i} \left[|\psi^i_r - \mu_r| \right] \le \Delta_r^{1/2}.$$

Proof. We begin by showing that the local message $M^u_{B_i}$ of any bidder $u \in B_i$ conveys little information on $G(J_i)$. Note that Fact 3.2 (and the Data Processing inequality (Fact 2.9)) together imply that, for any block B_i of bidders and any $u \in B_i$,

$$I(M^u_{B_i}; J_i \mid \mathcal{I}_i) = 0. \tag{2}$$

(Note that in contrast, $I(M_{B_i}; J_i \mid \mathcal{I}_i) \neq 0$. In fact, J_i may be almost determined by the entire message of the i'th block (let alone by the entire message M_1 of π), as the induced distribution of $G(J_i)$ is different than that of G^i_j, $j \neq J_i$. This is where we crucially use the *simultaneaty* of bidder's messages). We will also need the following proposition:

Proposition 4.3. *For any bidder $u \in B_i$ and any $j \in [\ell \cdot n_r^2 + 1]$, it holds that*

$$I(M^u_{B_i}; G^i_j \,|\mathcal{I}_i, J_i = j) \leq I(M^u_{B_i}; G^u_j \,|\mathcal{I}_i).$$

Proof. Recall that $\Gamma_u = \{G^u_1, G^u_2, \ldots, G^u_{\ell \cdot n_r^2+1}\}$ is the input of bidder u, and notice that for any $j \in [\ell \cdot n_r^2 + 1]$,

$$(M^u_{B_i}|\mathcal{I}_i, J_i = j) \to (\Gamma_u|\mathcal{I}_i, J_i = j) \to (G^u_j|\mathcal{I}_i, J_i = j) \to (G^i_j|\mathcal{I}_i, J_i = j)$$

is a Markov chain where the left chain holds since, conditioned on $(\Gamma_u, \mathcal{I}_i, J_i = j)$, $M^u_{B_i}$ is completely determined and therefore independent of G^i_j and G^u_j, and the right chain holds since conditioned on $\mathcal{I}_i, J_i = j$ and the graph G^u_j, the rest of the edges of the graph G^i_j are independent of Γ_u by construction. Therefore, by the (general) Data Processing inequality (Fact 2.8), we have

$$I(M^u_{B_i}; G^i_j \,|\mathcal{I}_i, J_i = j) \leq I(M^u_{B_i}; G^u_j \,|\mathcal{I}_i, J_i = j). \tag{3}$$

Now, by Fact 3.2, we know that the distribution of $(\Gamma_u|\mathcal{I}_i)$ is independent of the event "$J_i = j$". Since $M^u_{B_i}$ and G^u_j are deterministic functions of Γ_u (conditioned on $\mathcal{I}_i$), this also implies that the joint distribution of $(M^u_{B_i}, G^u_j|\mathcal{I}_i)$ is independent of the event "$J_i = j$". Therefore, we conclude by (3) that

$$I(M^u_{B_i}; G^i_j \,|\mathcal{I}_i, J_i = j) \leq I(M^u_{B_i}; G^u_j \,|\mathcal{I}_i, J_i = j) = I(M^u_{B_i}; G^u_j \,|\mathcal{I}_i).$$

□

We proceed to prove the Lemma. We may now write for any $u \in B_i$

$$I(M^u_{B_i}; G(J_i) \,|J_i, \mathcal{I}_i) = \frac{1}{\ell \cdot n_r^2 + 1} \cdot \sum_{j=1}^{\ell \cdot n_r^2+1} I(M^u_{B_i}; G^i_j \,|\mathcal{I}_i, J_i = j)$$

(By definition of conditional mutual information and since $J_i \in_R [\ell \cdot n_r^2 + 1]$ and by (2))

$$\leq \frac{1}{\ell \cdot n_r^2 + 1} \cdot \sum_{j=1}^{\ell \cdot n_r^2+1} I(M^u_{B_i}; G^u_j \,|\mathcal{I}_i) \qquad \text{(By Proposition 4.3)}$$

$$\leq \frac{1}{\ell \cdot n_r^2 + 1} \cdot \sum_{j=1}^{\ell \cdot n_r^2+1} I(M^u_{B_i}; G^u_j \,|G^u_1, G^u_2, \ldots, G^u_{j-1}, \mathcal{I}_i) \tag{4}$$

$$= \frac{1}{\ell \cdot n_r^2 + 1} \cdot I(M^u_{B_i}; G^u_1, G^u_2, \ldots, G^u_{\ell \cdot n_r^2+1}|\mathcal{I}_i) \qquad \text{(by the chain rule)}$$

$$\leq \frac{H(M^u_{B_i})}{\ell \cdot n_r^2 + 1} \qquad \text{(by Fact 2.4)}$$

$$\leq \frac{|M^u_{B_i}|}{\ell \cdot n_r^2 + 1} \leq \frac{\ell}{\ell \cdot n_r^2 + 1} < \frac{1}{n_r^2} = \Delta_r^2 \tag{5}$$

where the inequality in (4) follows from Lemma 2.6 taken with $A = G^u_j, B = M^u_{B_i}, C = \mathcal{I}_i, D = G^u_{<j}$, since G^u_j is independent of $G^u_{<j}$ for all j, conditioned on $\mathcal{I}_i$.

Now, we claim that, for each bidder $u \in B_i$, conditioning on the previous messages of the bidders ($M^{<u}_{B_i} := M^1_{B_i} M^2_{B_i} \ldots M^{u-1}_{B_i}$) can only *decrease* the information $M^u_{B_i}$ reveals on the hidden graph $G(J_i)$:

Claim 4.4. $I(M^u_{B_i}; G(J_i) \mid M^{<u}_{B_i}, J_i, \mathcal{I}_i) \le I(M^u_{B_i}; G(J_i) \mid J_i, \mathcal{I}_i)$.

Proof. By construction of μ_{r+1}, conditioned on $G(J_i), \mathcal{I}_i$ and J_i, the inputs of bidders u and $B_i \setminus \{u\}$ are *independent*. In particular, this fact and the data processing inequality (Fact 2.9) together imply that

$$I(M^u_{B_i}; M^{<u}_{B_i} \mid G(J_i), J_i, \mathcal{I}_i) = 0,$$

since π was assumed to be a deterministic protocol. By non-negativity of information and the chain rule,

$$\begin{aligned}
&I(M^u_{B_i}; G(J_i) \mid M^{<u}_{B_i}, J_i, \mathcal{I}_i) \le I(M^u_{B_i}; G(J_i), M^{<u}_{B_i} \mid J_i, \mathcal{I}_i)\\
&= I(M^u_{B_i}; G(J_i) \mid J_i, \mathcal{I}_i) + I(M^u_{B_i}; M^{<u}_{B_i} \mid G(J_i), J_i, \mathcal{I}_i)\\
&= I(M^u_{B_i}; G(J_i) \mid J_i, \mathcal{I}_i).
\end{aligned}$$

□

We conclude that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{m_{B_i}, j_i, \mathcal{I}_i}\left[\mathbb{D}\left(\psi^i_r \| \mu_r\right)\right] = I(M_{B_i}; G(J_i) \mid J_i, \mathcal{I}_i) \quad \text{(by Definition 2.3 of conditional mutual information)}\\
&= \sum_{u \in B_i} I(M^u_{B_i}; G(J_i) \mid M^{<u}_{B_i}, J_i, \mathcal{I}_i) \quad \text{(by the chain rule)}\\
&\le \sum_{u \in B_i} I(M^u_{B_i}; G(J_i) \mid J_i, \mathcal{I}_i) \quad \text{(by Claim 4.4)}\\
&\le |B_i| \cdot \Delta^2_r \quad \text{(by (5))}\\
&= n_r \cdot \Delta^2_r = \Delta_r. \qquad (6)
\end{aligned}$$

Combining (6), Pinsker's inequality (Lemma 2.2) and convexity of $\sqrt{\cdot}$ completes the entire proof of the lemma.

□

We are now ready to prove Theorem 4.1. To this end, for any r-round protocol π, input graph G, and induced subgraph $H \subseteq G$, let

$$N_\pi(G, H) := |\hat{\mathcal{M}}(\Pi(G)) \cap E(H)|$$

denote the size of the matching computed from π's transcript with respect to the subgraph H (note that $N_\pi(G, H)$ is a random variable depending on G). For notational convenience, we use the shorthand $N_\pi(G) := N_\pi(G, G)$. Theorem 4.1 will follow directly from the following theorem:

Theorem 4.5. *Let π be an r-round (deterministic) communication protocol with bandwidth ℓ. Then*

$$\mathop{\mathbb{E}}_{G \sim \mu_r}[N_\pi(G)] \le 5n_r \cdot \left(\sum_{k=0}^{r-1} \Delta_k^{1/2}\right) + 1.$$

Proof. We prove the theorem by induction on r. Let us denote

$$t(r) := 5n_r \cdot \left(\sum_{k=0}^{r-1} \Delta_k^{1/2}\right) + 1.$$

For $r = 0$ (namely, with no communication at all), the expected number of edges the referee guesses correctly under μ_0 is at most $n_0 \cdot \frac{1}{n_0} = 1 = t(0)$ (as $G^0 \sim \mu_0$ is a random permutation on $[n_0]$).

Suppose the theorem statement holds for all integers up to r. Thus, the expected matching produced by any r-round protocol θ (with bandwidth $\le \ell$) under μ_r satisfies

$$\mathop{\mathbb{E}}_{G \sim \mu_r}[N_\theta(G)] \le t(r). \qquad (7)$$

We need to show that the expected matching produced by any $(r+1)$-round protocol π (with bandwidth $\leq \ell$) under μ_{r+1} satisfies

$$\mathop{\mathbb{E}}_{G\sim\mu_{r+1}}[N_\pi(G)] \leq t(r+1). \tag{8}$$

Let π be an $(r+1)$-round protocol. Recall that $G \sim \mu_{r+1}$ consists of n_r^4 "blocks" B_i of bidders, each of which is connected to exactly $|\mathcal{I}_i| = \ell \cdot n_r^2 + 1$ item blocks. Let $M_1 := M_{B_1} M_{B_2} \dots M_{B_{n_r^4}}$ denote the messages sent by each *block* of bidders in the first round of π (where $M_{B_i} = M_{B_i}^1, M_{B_i}^2, \dots, M_{B_i}^{n_r}$ is the concatenated message of all bidders $u \in B_i$). Recall that for every $i \in [n_r^4]$, $G(J_i)$ denotes the induced subgraph of G on $(B_i, T_{\tau_i^{-1}(J_i)})$, and that for every bidder $u \in B_i$, $G_{J_i}^u = (u, T_{\tau_i^{-1}(J_i)})$ denotes the induced subgraph between bidder u and the "hidden graph" of the i'th block to which u belongs. In the same spirit, for every block B_i and every bidder $u \in B_i$, let

$$G(T_i) := \left(B_i, \bigcup_{j=1}^{\ell\cdot n_r^2} T_{a_j}\right) \quad , \quad G_T^u := \left(u, \bigcup_{j=1}^{\ell\cdot n_r^2} T_{a_j}\right)$$

denote the induced subgraph on the block B_i (on the bidder $u \in B_i$) and all "fooling blocks" respectively. As usual, for any subset $S \subseteq [n_r^4]$, we write $G(T_S) := \left(\bigcup_{i\in S} B_i, \bigcup_{j=1}^{\ell\cdot n_r^2} T_{a_j}\right)$ and use the convention $\mathbf{T} := T_{[n_r^4]}$. In what follows, $G(\mathbf{J}) := G(J_1)G(J_2)\dots G(J_{n_r^4})$ denotes the (concatenation of the) "hidden" graphs. The following proposition will be essential for the rest of our argument:

Lemma 4.6 (Conditional Subgraph Decomposition)**.** *The following conditions hold:*

1. $((G_T^1, G_T^2, \dots, G_T^{n_r}) \mid M_1, G(J_1), \mathbf{J}, \mathcal{I}) \sim \bigtimes_{u\in B_1}(G_T^u | M_1, G_{J_1}^u, \mathbf{J}, \mathcal{I})$, *where* $\{1, 2, \dots, n_r\}$ *are the bidders of the first block* B_1.
2. $(G(\mathbf{J}), G(\mathbf{T}) \mid M_1, \mathbf{J}, \mathcal{I}) \sim \bigtimes_{i\in[n_r^4]}(G(J_i)G(T_i) \mid M_{B_i}, \mathbf{J}, \mathcal{I})$.

That is, the joint distribution of the "fooling subgraphs" G_T^u *of each bidder* $u \in B_1$ *conditioned on the entire message* M_1 *and the "hidden graph"* $G(J_1)$ *of the first block, is a product of the marginal distributions* G_T^u *conditioned only on the "local hidden part"* $G_{J_1}^u$ *of each bidder and* M_1*.*

Furthermore, the joint distribution of the subgraphs induced on each block $(G(J_i)G(T_i))$ *conditioned on the entire message* M_1 *is a product distribution of the marginal distributions of the* i*'th block, conditioned only on the "local" message of the respective block* M_{B_i} *(In particular, these graphs remain independent even conditioned on* M_1*).*

The intuition behind the second proposition is clear: Since in the original distribution μ_{r+1}, the graphs of each block are independent by construction, this remains true even when conditioned on the first (deterministic) message of each block. The first proposition is more subtle, since within the same block (say B_1), the inputs of the bidders $u \in B_1$ are correlated (via the hidden graph $G(J_1)$). However, conditioned on knowing the hidden block $(\mathbf{J}, \mathcal{I})$, the marginal distribution of the fooling graph G_T^u is *independent* for each u by construction, and therefore the only correlation between $G(J_1)$ and G_T^u created by conditioning on the message M_1, is correlation between the "local hidden graph" of bidder u ($G_{J_1}^u$) and G_T^u. We remark that this fact will be used crucially in the embedding argument below (Claim 4.7). We proceed to the formal proof.

Proof of Lemma 4.6. We repeatedly use Lemma 2.7.

Proof of (1) It suffices to show that for every $u \in B_1$, $I(G_T^u; G_T^{-u} G_{J_1}^{-u} | M_1, \mathbf{J}, \mathcal{I}, G_{J_1}^u) = 0$. To this end, observe that

$$I(G_T^u; M_1^{-u} | M_1^u, G(J_1), G_T^{-u}, \mathbf{J}, \mathcal{I}) \leq H(M_1^{-u} | G(J_1), G_T^{-u}, \mathbf{J}, \mathcal{I}) = 0, \tag{9}$$

since the message M_1^{-u} of all bidders in B_1 except bidder u is fully determined by the inputs $(G(J_1), G_T^{-u})$. For the same reason,

$$I(G_T^{-u} G_{J_1}^{-u}; M_1^u | G_{J_1}^u, G_T^u, \mathbf{J}, \mathcal{I}) \leq H(M_1^u | G_{J_1}^u, G_T^u, \mathbf{J}, \mathcal{I}) = 0. \tag{10}$$

Therefore,

$$\begin{aligned}
&I(G_T^u; G_T^{-u} G_{J_1}^{-u} | M_1, \mathbf{J}, \mathcal{I}, G_{J_1}^u) = I(G_T^u; G_T^{-u} G_{J_1}^{-u} | M_1^u, M_1^{-u}, \mathbf{J}, \mathcal{I}, G_{J_1}^u) \\
&\leq I(G_T^u; G_T^{-u} G_{J_1}^{-u} | M_1^u, \mathbf{J}, \mathcal{I}, G_{J_1}^u) \quad \text{(By Lemma 2.7 with } D = M_1^{-u}\text{, and (9))} \\
&\leq I(G_T^u; G_T^{-u} G_{J_1}^{-u} | \mathbf{J}, \mathcal{I}, G_{J_1}^u) \qquad \text{(By Lemma 2.7 with } D = M_1^u\text{, and (10))} \\
&= 0, \quad \text{as desired.}
\end{aligned}$$

Proof of (2) It suffices to show $I(G(J_i)G(T_i); G(J_{-i})G(T_{-i})M_{B_{-i}} | M_{B_i}, \mathbf{J}, \mathcal{I}) = 0$. Once again, applying Lemma 2.7 with $D = M_{B_i}$, we have

$$I(G(J_i)G(T_i); G(J_{-i})G(T_{-i}) | M_{B_i}, \mathbf{J}, \mathcal{I}) \leq I(G(J_i)G(T_i); G(J_{-i})G(T_{-i}) | \mathbf{J}, \mathcal{I}) \tag{11}$$

since $I(M_{B_i}; G(J_{-i})G(T_{-i}) | G(J_i), G(T_i), \mathbf{J}, \mathcal{I}) \leq H(M_{B_i} | G(J_i), G(T_i), \mathbf{J}, \mathcal{I}) = 0$ where the last equality is because M_{B_i} is determined by the input of block B_i. The same argument implies

$$I(G(J_i)G(T_i); M_{B_{-i}} | M_{B_i}, \mathbf{J}, \mathcal{I}, G(J_{-i})G(T_{-i})) \leq I(G(J_i)G(T_i); M_{B_{-i}} | \mathbf{J}, \mathcal{I}, G(J_{-i})G(T_{-i})) \tag{12}$$

since once again, $I(M_{B_i}; M_{B_{-i}} | G(\mathbf{J}), G(\mathbf{T}), \mathbf{J}, \mathcal{I}) \leq H(M_{B_i} | G(\mathbf{J}), G(\mathbf{T}), \mathbf{J}, \mathcal{I}) = 0$. Combining equations (11) and (12), we conclude by the chain rule that

$$\begin{aligned}
&I(G(J_i)G(T_i); G(J_{-i})G(T_{-i})M_{B_{-i}} | M_{B_i}, \mathbf{J}, \mathcal{I}) \\
&= I(G(J_i)G(T_i); G(J_{-i})G(T_{-i}) | M_{B_i}, \mathbf{J}, \mathcal{I}) + I(G(J_i)G(T_i); M_{B_{-i}} | M_{B_i}, \mathbf{J}, \mathcal{I}, G(J_{-i})G(T_{-i})) \\
&\leq I(G(J_i)G(T_i); G(J_{-i})G(T_{-i}) | \mathbf{J}, \mathcal{I}) + I(G(J_i)G(T_i); M_{B_{-i}} | \mathbf{J}, \mathcal{I}, G(J_{-i})G(T_{-i})) \\
&= 0,
\end{aligned} \tag{13}$$

where the last transition follows from the definition of μ_{r+1}, and since $M_{B_{-i}}$ is a deterministic function of $G(J_{-i})G(T_{-i})$ conditioned on $\mathbf{J}, \mathcal{I}$.

□

We now proceed to prove (8), the inductive step of the proof. Recall that π is assumed to be deterministic, but M_1 is still a random variable (under the input distribution μ_{r+1}). Hence, we may equivalently draw $G \sim \mu_{r+1}$ by first sampling the first message $m_1 \sim M_1$, and then sampling $G \sim \mu_{r+1} | m_1$. Let us denote by $\pi | m_1$ the protocol which is the subtree of π conditioned on the first message being m_1. Note that $\pi | m_1$ has only r rounds of communication. We therefore have

$$\mathop{\mathbb{E}}_{G \sim \mu_{r+1}} [N_\pi(G)] \leq \mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J}, \mathcal{I}}} \mathop{\mathbb{E}}_{G | m_1, \mathbf{J}, \mathcal{I}} \left[\ell \cdot n_r^2 \cdot m_r + \sum_{i=1}^{n_r^4} N_{\pi | m_1}(G, G(J_i)) \right] \tag{14}$$

since any matching in G can at most match all of the items in $\bigcup_{j=1}^{\ell \cdot n_r^2} T_{a_j}$ and each block T_{a_j} contains m_r edges by definition of μ_{r+1}, and the rest of the matched edges are contained in $G(J_1), G(J_2), \ldots, G(J_{n_r^4})$. Recall that by definition of μ_{r+1}, $G(J_i) \sim \mu_r$. Hence by linearity of expectation and the second proposition of Lemma 4.6, we may equivalently write the above as

$$\begin{aligned}
&= \ell \cdot n_r^2 \cdot m_r + \sum_{i=1}^{n_r^4} \mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J}, \mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_i) | (m_{B_i}, \mathbf{J}, \mathcal{I}) \\ G | \ (G(J_i), m_1, \mathbf{J}, \mathcal{I})}} \left[N_{\pi | m_1}(G, G(J_i)) \right] \\
&= \ell \cdot n_r^2 \cdot m_r + \sum_{i=1}^{n_r^4} \mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J}, \mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_i) \sim \psi_r^i \\ G | \ (G(J_i), m_1, \mathbf{J}, \mathcal{I})}} \left[N_{\pi | m_1}(G, G(J_i)) \right],
\end{aligned} \tag{15}$$

by definition of ψ_r^i (actually, ψ_r^i is defined conditioned only on $J_i, \mathcal{I}_i$ but conditioning on all indices $\mathbf{J}, \mathcal{I}$ clearly doesn't change the distribution). By symmetry of the distribution μ_{r+1}, it suffices to upper bound the first term in the above summation

$$\mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_1)\sim\psi_r^1 \\ G|\ (G(J_1),m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right].$$

To this end, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_1)\sim\psi_r^1 \\ G|\ (G(J_1),m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right] \\
&\leq \mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \left[|\psi_r^1 - \mu_r|\right] \cdot n_r + \mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_1)\sim\mu_r \\ G|\ (G(J_1),m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right] && (16) \\
&\leq \Delta_r^{1/2} \cdot n_r + \mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_1)\sim\mu_r \\ G|\ (G(J_1),m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right] && (17)
\end{aligned}$$

where (16) follows from fact 2.10, since trivially $N_{\pi|m_1}(G, G(J_1) \leq |U^r| = n_r$ for any $G(J_1)$, and the last transition (17) follows from Lemma 4.2. We now wish to use the inductive hypothesis to argue that the rightmost term of (17) cannot exceed the expected matching size of an r-round protocol over μ_r. We do so using an embedding argument, which is the heart of the proof.

Claim 4.7 (r-round Embedding)**.**

$$\mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_1)\sim\mu_r \\ G|\ (G(J_1),m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right] \leq t(r).$$

Proof. Notice that the protocol $\pi|m_1$ is defined over inputs from μ_{r+1} and not μ_r, so we cannot apply the inductive hypothesis directly to obtain our desired upper bound. Instead, we will "embed" $H \sim \mu_r$ into $\pi|m_1$ by simulating the rest of the players using public randomness (and then fix the public coins to obtain a deterministic protocol). To this end, for every bidder $u \in U^r$, denote by H_u the induced subgraph of H on the vertex u (i.e., the input of bidder u in μ_r).

Consider the following r-round randomized protocol τ for $H \sim \mu_r$: The n_r players use the shared random tape to sample $(M_1, \mathbf{J}, \mathcal{I})$ (according to the probability space of π). Then, they "embed" their inputs (the graph H) to the first block B_1 and each bidder $u \in B_1$ "completes" his missing edges to the "fooling graph" G_T^u according to $(G_T^u | M_1, H_u, \mathbf{J}, \mathcal{I})$ *using private randomness.* Note that this is possible due to the first proposition of Lemma 4.6, since it asserts that $(G_T^u | M_1, H, \mathbf{J}, \mathcal{I}) \sim (G_T^u | M_1, H_u, \mathbf{J}, \mathcal{I})$. The players now use the second proposition of Lemma 4.6 to sample the graphs of the rest of the blocks according to $\times_{i=2}^{n_r^4} (G(J_i) G(T_i) \mid M_{B_i}, \mathbf{J}, \mathcal{I})$, as the proposition asserts that these subgraphs remain independent after the conditioning.

This process specifies a graph G such that $H \sim \mu_r$ and $G \sim \mu_{r+1}$ conditioned on $G(J_1) = H, M_1, \mathbf{J}, \mathcal{I}$. Notice that so far the players have not communicated at all. The players now run the r-round protocol $\pi|m_1$ and outputs its induced matching on H. Notice that this protocol is well defined, as the messages of bidders outside block B_1 in $\pi|m_1$ in every round ($r \in \{2, 3, \ldots, r+1\}$) are completely determined by their respective inputs on the random tape and the content of the blackboard, since π was assumed to be a deterministic protocol. Call the resulting protocol τ.

By construction, the expected matching size of τ (over the private and public randomness $M_1, \mathbf{J}, \mathcal{I}, G$) with respect to H is

$$\mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{H\sim\mu_r \\ G|\ (G(J_1)=H,m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, H)\right] = \mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_1)\sim\mu_r \\ G|\ (G(J_1),m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right].$$

By the averaging principle, there is some fixing of the randomness of τ that obtains (at least) the same expectation as above with respect to $H \sim \mu_r$. Call this deterministic protocol τ'. But τ' is a deterministic

r-round protocol over μ_r, hence the inductive hypothesis asserts that

$$\mathop{\mathbb{E}}_{\substack{m_1\\ \mathbf{J},\mathcal{I}}} \mathop{\mathbb{E}}_{\substack{G(J_1)\sim\mu_r\\ G|\ (G(J_1),m_1,\mathbf{J},\mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right] \ \leq \mathop{\mathbb{E}}_{H\sim\mu_r}\left[N_{\tau'}(H)\right] \ \leq t(r),$$

as claimed.

$\square$

We are now in shape to complete the entire proof of Theorem 4.5. Plugging in the bounds of (17) and Claim 4.7 into equation (15), we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{G\sim\mu_{r+1}}[N_\pi(G)] \leq \ell\cdot n_r^2\cdot m_r + n_r^4\cdot\left[\Delta_r^{1/2}\cdot n_r + t(r)\right]\\
&= \ell\cdot n_r^2\cdot m_r + n_r^4\cdot\left[\Delta_r^{1/2}\cdot n_r + 5n_r\cdot\left(\sum_{k=0}^{r-1}\Delta_k^{1/2}\right)+1\right]\\
&\leq n_r^4\cdot\left[5n_r\cdot\Delta_r^{1/2} + 5n_r\cdot\left(\sum_{k=0}^{r-1}\Delta_k^{1/2}\right)\right] \qquad (\text{since } \ell\cdot n_r^2\cdot m_r + n_r^4 < 4n_r^5\Delta_r^{1/2})\\
&= 5n_r^5\cdot\left(\sum_{k=0}^{r}\Delta_k^{1/2}\right) = 5n_{r+1}\cdot\left(\sum_{k=0}^{r}\Delta_k^{1/2}\right) \qquad (\text{since by definition, } n_{r+1}=n_r^5)\\
&= t(r+1)-1\\
&< t(r+1).
\end{aligned}\tag{18}$$

This proves the induction step (8), and therefore concludes the entire proof of Theorem 4.5. $\square$

Theorem 4.5 immediately implies our desired lower bound:

Proof of Theorem 4.1. By Theorem 4.5, the expected matching size of any r-round protocol π under μ_r is at most

$$\begin{aligned}
\mathop{\mathbb{E}}_{G\sim\mu_r}[N_\pi(G)] &\leq t(r) = 5n_r\cdot\left(\sum_{k=0}^{r-1}\Delta_k^{1/2}\right)+1\\
&= 1+5n_r\cdot\sum_{k=0}^{r-1}\left(\frac{1}{n_k}\right)^{1/2} = 1+5n_r\cdot\sum_{k=0}^{r-1}\left(\frac{1}{\ell^{5^{k+1}}}\right)^{1/2},
\end{aligned}$$

since by definition of μ_r, $n_r = n_{r-1}^5$, and $n_0 := \ell^5$, hence $n_r = \ell^{5^{r+1}}$. Since this is a doubly-exponential decaying series (and as long as ℓ is large enough than some absolute constant), we can upper bound the sum by, say,

$$\begin{aligned}
&\leq 5n_r\cdot\Delta_0^{1/5} = 5n_r\cdot\left(\frac{1}{\ell^5}\right)^{1/5}\\
&= \frac{5}{\ell}\cdot n_r = 5\cdot\frac{n_r}{n_r^{1/5^{r+1}}} = 5\cdot n_r^{1-1/5^{r+1}},
\end{aligned}$$

since, by property (4) in Remark 3.1, $n_r = \ell^{5^{r+1}} \iff \ell = n_r^{1/5^{r+1}}$.

$\square$

References

[Alo03] Noga Alon. A simple algorithm for edge-coloring bipartite multigraphs. *Inf. Process. Lett.*, 85(6):301–302, March 2003.

[CT91] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley series in telecommunications. J. Wiley and Sons, New York, 1991.

[DGS86] Gabrielle Demange, David Gale, and Marilda Sotomayor. Multi-item auctions. *The Journal of Political Economy*, pages 863–872, 1986.

[DNO14] Shahar Dobzinski, Noam Nisan, and Sigal Oren. Economic efficiency requires interaction. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 233–242, 2014.

[GKK12] Ashish Goel, Michael Kapralov, and Sanjeev Khanna. On the communication and streaming complexity of maximum bipartite matching. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 468–485. SIAM, 2012.

[GO13] Venkatesan Guruswami and Krzysztof 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*, pages 287–298, 2013.

[HK73] John E Hopcroft and Richard M Karp. An nˆ5/2 algorithm for maximum matchings in bipartite graphs. *SIAM Journal on computing*, 2(4):225–231, 1973.

[HRVZ13] Zengfeng Huang, Bozidar Radunovic, Milan Vojnovic, and Qin Zhang. Communication complexity of approximate maximum matching in distributed graph data. *Microsoft Technical Report, MSR-TR-2013-35*, 2013.

[Kap12] Michael Kapralov. Better bounds for matchings in the streaming model. *arXiv preprint arXiv:1206.2269*, 2012.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity.* Cambridge University Press, Cambridge, 1997.

[KUW85] Richard M Karp, Eli Upfal, and Avi Wigderson. Constructing a perfect matching is in random nc. In *Proceedings of the seventeenth annual ACM symposium on Theory of computing*, pages 22–32. ACM, 1985.

[LPSP08] Zvi Lotker, Boaz Patt-Shamir, and Seth Pettie. Improved distributed approximate matching. In *Proceedings of the twentieth annual symposium on Parallelism in algorithms and architectures*, pages 129–136. ACM, 2008.

[MVV87] Ketan Mulmuley, Umesh V Vazirani, and Vijay V Vazirani. Matching is as easy as matrix inversion. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 345–354. ACM, 1987.

[Yus13] Raphael Yuster. Maximum matching in regular and almost regular graphs. *Algorithmica*, 66(1):87–92, 2013.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de