Economic Efficiency Requires Interaction

[Extended Abstract] *

Shahar Dobzinski†
Weizmann Institute of Science
shahar.dobzinski@weizmann.ac.il

Noam Nisan
Hebrew University and Microsoft Research
noam@cs.huji.ac.il

Sigal Oren‡
Hebrew University and Microsoft Research
sigalo@cs.huji.ac.il

ABSTRACT
We study the necessity of interaction between individuals for obtaining approximately efficient economic allocations. We view this as a formalization of Hayek's classic point of view that focuses on the information transfer advantages that markets have relative to centralized planning. We study two settings: combinatorial auctions with unit-demand bidders (bipartite matching) and combinatorial auctions with subadditive bidders. In both settings we prove that non-interactive protocols require exponentially larger communication costs than interactive ones, even those that only use a modest amount of interaction.

Categories and Subject Descriptors
J.4 [Social and behavioral sciences]: Economics

General Terms
Economics, Theory

Keywords
Auctions, Matching, Communication Complexity

1. INTRODUCTION
The most basic economic question in a social system is arguably how to determine an efficient allocation of the economy’s resources. This challenge was named the economic calculation problem by von Mises, who argued that markets do a better job than centralized systems can. In his classic paper [16], Hayek claimed that the heart of the matter is the distributed nature of “input”, i.e., that the central planner does not have the information regarding the costs and utilities of the different parties:

knowledge of the circumstances of which we must make use never exists in concentrated or integrated form but solely as the dispersed bits of ... knowledge which all the separate individuals possess.

Hayek therefore proposes that the question of which economic system is better (market-based or centrally-planned) can essentially be reduced to determining which of them is able to better transfer the information needed for economic efficiency:

which of these systems is likely to be more efficient depends ... on whether we are more likely to succeed in putting at the disposal of a single central authority all the knowledge which ... is initially dispersed among many different individuals, or in conveying to the individuals such additional knowledge as they need in order to enable them to fit their plans with those of others.

This paper formalizes Hayek’s question in specific technical terms that quantify the amount of information exchange needed for obtaining economic efficiency. We consider the main distinction – in terms of information transfer – between a centralized system and a distributed market-based one to be that of interaction: in a centralized system all individuals send information to the central planner who must then determine an efficient allocation, while market-based systems are by nature interactive.1

Our main results support Hayek’s point of view. We exhibit situations where interaction allows exponential savings in information transfer, making the economic calculation problem tractable for interactive markets even when it is intractable for a centralized planner. We have two conceptually similar, but technically disjoint, sets of results along this line. The first set of results considers the classic simple setting of unit-demand bidders, essentially a model of matching in bipartite graphs. The second and more complicated setting concerns combinatorial auctions with subadditive bidders. In both settings we show that non-interactive protocols have exponentially larger communication costs relative to interactive ones.

1One may of course conceive of interactive versions of centralized planning such as the “market socialism” models proposed by Lange and Lerner [22], but these become indistinguishable from market mechanisms in terms of their informational abilities. As Hayek [16] puts it: “... when Professor Abba P. Lerner rediscovers Adam Smith ...”.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright and reproduction permission requests for materials in the ACS Publications database should be submitted online at http://acsl-photocopy-permission.org. This includes requests to photocopy articles that are distributed only by permission of the ACS. Permission to republish or repurpose articles or portions of articles for commercial, marketing, or similar purposes also requires obtaining ACS permission and must be submitted via our RightsLink service at http://acsl-photocopy-permission.org. Requests can be made online at http://acsdm.org/publications/rights. Requests to repurpose ACS content online, including but not limited to posting to a digital repository or website, must be submitted via our RightsLink service at http://acsl-photocopy-permission.org. This includes requests to repurpose articles or portions of articles as Creative Commons licenses or other open access licenses. Copyright © 2014 by the Association for Computing Machinery, Inc.
In a complementary set of results we also show that exponential savings in communication costs can be realized with even limited interaction.

In technical terms, we formalize this problem in the realm of communication complexity. Non-interactive systems are modeled as simultaneous communication protocols, where all agents simultaneously send messages to a central planner who must decide on the allocation based on these messages alone. Interactive systems may use multiple rounds of communication and we measure the amount of interactiveness of a system by the number of communication rounds. In both of our settings we prove lower bounds on the simultaneous communication complexity of finding an approximately efficient allocation as well as exponentially smaller upper bounds for protocols that use a modest number of rounds of interaction.

We now elaborate in more details on the two settings that we consider and describe our results. We begin with the technically simpler setting of bipartite matching.

### 1.1 Bipartite Matching

In this simple matching scenario there are $n$ players and $m$ goods. Each player $i$ is interested in acquiring a single item from some privately known subset $S_i$ of the goods and our goal is to allocate the items to the players in a way that maximizes the number of players who get an item from their desired set. This is of course a classic problem in economics (matching among unit demand bidders) as well as in computer science (bipartite matching).

We first consider simultaneous protocols. Each of the players is allowed to send a small amount of information, $l$ bits with $l << n$, to the centralized planner who must then output a matching.

**Theorem:**

- Every deterministic simultaneous protocol where each player $i$ sends at most $n^l$ bits of communication cannot approximate the size of the maximum matching to within a factor better than $O(n^{1-\epsilon})$.
- Any randomized simultaneous protocol where each player $i$ sends at most $n^l$ bits of communication cannot approximate the size of the maximum matching to within a factor better than $O(n^{1/2-2\epsilon})$.

Both our bounds are essentially tight. For deterministic protocols, one can trivially obtain an approximation ratio of $n$ with message length $O(n \log n)$: each player sends the index of one arbitrary item that he is interested in. If randomization is allowed, it is not hard to see that when each player sends the index of a random item he is interested in, we get an approximation ratio of $O(\sqrt{n})$. We have therefore established a gap between randomized and deterministic protocols. We also note that the randomized lower bound can in fact be obtained from more general recent results of [18] in a stronger model. For completeness we present a significantly simpler direct proof for our setting.

On the positive side, we show that a few communication rounds suffice to get an almost efficient allocation\(^2\). Our algorithm is a specific instantiation of the well known class of “auction algorithms”. This class of algorithms has its roots in [8] and has been extensively studied from an economic point of view (e.g. [29]) as well as from a computational point of view (starting with [5]).

The standard ascending auction algorithm for this setting begins by setting the price of each item to be 0. Initially, all bidders are “unallocated”. Then, in an arbitrary order, each unallocated player $i$ reports an index of an item that maximizes his profit in the current prices (his “demand”). The price of that item increases by $\epsilon$ and the item is reallocated to player $i$, and the process continues with another unallocated player. It is well known that if the maximum that a player is willing to pay for an item is 1, then this process terminates after at most $O(\frac{1}{\epsilon^2})$ steps and it is not hard to construct examples where this is tight. We show that if in each round every unallocated player reports, simultaneously with the others, an index of a random item that maximizes his profit in the current prices ($O(\log n)$ bits of information) and each reported item is reallocated to an arbitrary player that reported it, then the process terminates in logarithmically many rounds. We are not aware of any other scenario where natural market dynamics provably converge (approximately) to an equilibrium in time that is sub-linear in the market size.

**Theorem:** Fix $\epsilon > 0$. After $O(\log n)$ rounds the randomized algorithm provides in expectation an $(1 + \epsilon)$-approximation to the bipartite matching problem.

We then quantify the tradeoff between the amount of interaction and economic efficiency. We show that for every $k \geq 1$ there is a randomized protocol that obtains an $O(n^{1/(k+1)})$-approximation in $k$ rounds, where at each round each player sends $O(\log n)$ bits of information.

In passing we note that the communication complexity of the exact problem, i.e. of finding an exact perfect matching, when it exists, remains a very interesting open problem. Moreover, we believe that it may shed some light on basic algorithmic challenges of finding a perfect matching in near-linear time as well as deterministically in parallel. We shortly present this direction in appendix A.

### 1.2 Combinatorial Auctions with Subadditive Bidders

Our second set of results concerns a setting where we are selling $m$ items to $n$ bidders in a combinatorial auction. Here each player $i$ has a valuation $v_i$ that specifies his value for every possible subset $S$ of items. The goal is to maximize the “social welfare” $\sum v_i(A_i)$, where $A_i$ is the set of goods that is allocated to player $i$. The communication requirements in such settings have received much attention and it is known that, for arbitrary valuations, exponential amount of communication is required to achieve even $m^{1/2-\epsilon}$-approximation of the optimal welfare [28]. However, it is also known that if the valuations are subadditive, $v_i(S \cup T) \leq v_i(S) + v_i(T)$, then constant factor approximations can be achieved using only polynomial communication [11, 12, 30, 9, 23]. Can this level of approximate welfare be achieved by a direct mechanism, without interaction?

Two recent lines of research touch on this issue. First, several recent papers show that valuations cannot be “compressed”, e.g. approximately, and that any polynomial-length description of subadditive valuations (or even the more restricted XOS valuations) must lose a factor of $O(\sqrt{m})$ in precision [3, 2]. Similar, but somewhat weaker, non-approximation results are also known for the far more restricted subclass of “gross-substitutes” valuations [4] for which exact welfare maximization is possible with polynomial iterative communication. Thus the natural approach for a direct mechanism where each player sends a succinctly specified approximate version of his valuation (a “sketch”) to the central planner cannot lead to a better than $O(\sqrt{m})$ approximation. This does not, however, rule out other approaches for non-interactive allocation, that do not require approximating the whole valuation, as we show next:

\(^2\)Formally, the algorithm works in the so called “blackboard model” – see Appendix A for a definition.
Theorem: There exists a deterministic communication protocol such that each player holds a subadditive valuation and sends (simultaneously with the others) polynomially many bits of communication to the central planner that guarantees an $O(m^{\frac{1}{k+1}})$-approximation to the optimal allocation.

Another line of relevant research considers bidders with such valuations being put in a game where they can only bid on each item separately [7, 6, 15, 13]. In such games the message of each bidder is by definition only $O(m)$ real numbers, each can be specified in sufficient precision with logarithmically-many bits. Surprisingly, it turns out that sometimes this suffices to get constant factor approximation of the social welfare. Specifically one such result [13] considers a situation where the valuation $v_i$ of each player $i$ is drawn independently from a commonly known distribution $D_i$ on subadditive valuations. In such a case, every player $i$ can calculate bids on the items – requiring $O(m \log m)$ bits of communication – based only on his valuation $v_i$ and the distributions $D_j$ of the others (but not their valuations $v_j$). By allocating each item to the highest bidder we get a 2-approximation to the social welfare, in expectation over the distribution of valuations. This is a non-interactive protocol that comes tantalizingly close to what we desire: all that remains is for the 2-approximation to hold for every input rather than in expectation. Using Yao’s principle, this would fare, in expectation over the distribution of valuations. This is a nice result to the highest bidder we get a 2-approximation to the social welfare. We do know that exactly solving the problem for correlated distributions [24, 1] would be possible here too.

Our main technical construction proves a negative answer and shows that interaction is essential for obtaining approximately optimal allocation among subadditive valuations (even for the more restricted XOS valuations):

**Theorem:** No (deterministic or randomized) protocol such that each player holds an XOS valuation and simultaneously with the others sends sub-exponentially many bits of communication to a central planner can guarantee an $O(m^{\frac{1}{k+1}})$-approximation.

Again, this is in contrast to interactive protocols that can achieve a factor 2 approximation [11] (with polynomially many rounds of communication). The lower bound shows that interaction is necessary to solve the economic calculation problem in combinatorial auctions. We show that if a small amount of interaction is allowed then one can get significantly better results:

**Theorem:** For every $k \geq 1$ there is a randomized protocol that obtains an $O(k \cdot m^{\frac{1}{k+1}})$-approximation in $k$ rounds, where at each round each player sends $\text{poly}(m, n) \log m$ rounds we get a poly-logarithmic approximation to the welfare.

**Open Questions**

In our opinion the most intriguing open question is to determine the possible approximation ratio achievable by simultaneous combinatorial auctions with submodular or even gross-substitutes players that are allowed to send $\text{poly}(m, n)$ bits. Our $O(m^{\frac{1}{k+1}})$-algorithm is clearly applicable for both settings. We do know that exactly solving the problem for gross-substitutes valuations requires exponential communication (details appear in the full version of the paper) although when interaction is allowed polynomial communication suffices.

Another natural open question is to prove lower bounds on the approximation ratio achievable by $k$-round protocols. Our bounds only hold for $k = 1$. Furthermore, how good is the approximation ratio that can be guaranteed when incentives are taken into account? Can a truthful $k$-round algorithm guarantee poly-logarithmic approximation in $O(\log n)$ rounds for XOS valuations?

For the bipartite matching setting we leave open the question of developing algorithms for weighted bipartite matching. In addition, our $k$-round algorithms are randomized; developing deterministic $k$-round algorithms even for the unweighted case is also of interest. In Appendix A we further discuss more open questions related to the communication complexity of bipartite matching and its relation to the computational complexity of bipartite matching.

Finally, we studied the matching problem in the framework of simultaneous communication complexity. A fascinating future direction is to study other classic combinatorial optimization problems (e.g., minimum cut, packing and covering problems, etc.) using the lenses of simultaneous communication complexity.

2. PRELIMINARIES

Combinatorial Auctions. In a combinatorial auction we have a set $N$ of players (|$N$| = $n$) and a set $M$ of different items (|$M$| = $m$). Each player $i$ has a valuation function $v_i : 2^M \rightarrow \mathbb{R}$. Each $v_i$ is assumed to be normalized ($v_i(\emptyset) = 0$) and non decreasing. The goal is to maximize the social welfare, that is, to find an allocation of the items to players ($A_1, \ldots, A_n$) that maximizes the welfare: $\Sigma_i v_i(A_i)$. A valuation function is subadditive if for every two bundles $S$ and $T$, $v(S) + v(T) \geq v(S \cup T)$. A valuation $v$ is additive if for every bundle $S$ we have that $v(S) = \Sigma_{i \in S} v_i([1, n])$. A valuation $v$ is XOS if there exist additive valuations $a_1, \ldots, a_n$ such that for every bundle $S$, $v(S) = \max_{i \in S} a_i(S)$. Each $a_i$ is a clause of $v$.

Matching. Here the goal is to find a maximum matching in an undirected bipartite graph $G = (V_1, V_2, E)$, |$V_1$| = |$V_2$| = $n$. Each player $i$ corresponds to vertex $i \in V_1$, and is only aware of edges of the form $(i, j)$ (j \in V_2 since the graph is bipartite). The neighbor set of $i$ is $S_i = \{j \in E | (i, j) \in E\}$. The goal is to maximize the number of matched pairs. When convenient we will refer to vertices on the left as unit demand bidders and the vertices on the right as goods. Under this interpretation the neighbor set of player $i$ is simply the set of goods that he is interested in.

Chernoff Bounds. Let $X$ be a random variable with expectation $\mu$. Then, for any $\delta > 0$: $P(X > (1 + \delta)\mu) < \left(\frac{e^{\delta}}{1 + \delta}\right)^{\mu} = \left(\frac{1}{e^{1 + \delta}}\right)^{\mu} = \left(\frac{1}{e^{(1 + \delta)^{\frac{1}{n+4}}}}\right)^{\mu}$. For $\delta > e^2$ we can loosely bound this expression by: $\frac{1}{e^{\mu}}$.

3. LOWER BOUNDS FOR BIPARTITE MATCHING

In this section we state lower bounds on the power of algorithms for bipartite matching. Proofs appear in the full version of the paper. The first one deals with the power of deterministic algorithms:

**Theorem 3.1** (Deterministic Algorithms Lower Bound). The approximation ratio of any deterministic simultaneous algorithm for matching that uses at most $l$ bits per player is no better than $\frac{n}{8l + 4 \log(n)}$. In particular, for any fixed $\epsilon > 0$ and $l = n^{\epsilon}$ the approximation ratio is $\Omega(n^{1-\epsilon})$. 

235
The second theorem gives a lower bound on the power of randomized algorithms:

**Theorem 3.2 (Randomized Algorithms Lower Bound).** Fix \( \epsilon > 0 \). The approximation ratio of every algorithm in which each player sends a message of size \( l \leq n^{1/2} - \frac{\epsilon}{2} \) is at most \( n^\epsilon \), for every \( \alpha \leq \frac{1}{2} - \epsilon \).

The next proposition (easy proof appears in the full version of the paper) shows that both lower bounds are essentially tight. In particular, this implies a proven gap between the power of deterministic and randomized algorithms.

**Proposition 3.3.**

1. There exists a deterministic simultaneous algorithm that provides an approximation ratio of \( \max(2, \frac{\log n}{\sqrt{n}}) \) using \( l \) bits per player.
2. There exists a simultaneous randomized algorithm that provides an expected approximation ratio of \( O(\sqrt{n}) \).

## 4. Algorithms for Bipartite Matching

We provide two algorithms that guarantee significantly better approximation ratios using a small number of rounds. We show that \( O(\frac{\log n}{\sqrt{n^\epsilon}}) \) rounds suffice to get a \((1+\epsilon)\) approximation. In the full version of the paper we present an algorithm that provides an approximation ratio of \( O(n^{1/2+\epsilon}) \) in \( k \) rounds. This shows that even a constant number of rounds suffices to get much better approximation ratios than what can be achieved by simultaneous algorithms.

### 4.1 A \((1+\epsilon)\)-Approximation for Bipartite Matching in \( O(\frac{\log n}{\sqrt{n^\epsilon}}) \) Rounds

The algorithm is based on an auction where each player competes at every point on one item that he demands the most at the current prices. Therefore, it will be easier for us to imagine the players as having valuations. Specifically, each player \( i \) is a unit demand bidder with \( v_i(j) = 1 \) if \( j \in S_i \) and \( v_i(j) = 0 \) otherwise.

**The Algorithm**

1. For every item \( j \) let \( p_j = 0 \).
2. Let \( N_i \) be the set of all players.
3. In every round \( r = 1, \ldots, \frac{\log n}{2\epsilon} \):
   (a) For each player \( i \in N_r \), let the demand of \( i \) be \( D_i = \arg \min_{j, p_j < 1, j \in S_i} p_j \). This is the subset of \( S_i \) for which the price of each item is minimal and smaller than 1.
   (b) Each player \( i \in N_r \) selects uniformly at random an item \( j_i \in D_i \) and reports its index.
   (c) Go over the players in \( N_r \) in a fixed arbitrary order. If item \( j \) was not yet allocated in this round, player \( i \) receives it and the price \( p_j \) is increased by \( \delta \). In this case we say that player \( i \) is committed to item \( j \). A player \( i \) that was committed to \( j \) in the previous round (if such exists) now becomes uncommitted.
   (d) Let \( N_{r+1} \) be the set of uncommitted players at the end of round \( r \).

4.2 An \( O(\frac{\log n}{\sqrt{n^\epsilon}}) \)-Approximation for XOS Valuations

Our algorithm is very similar to the classical auction algorithms except for two seemingly small changes. However, quite surprisingly, these changes allow us to substantially reduce the communication cost. The first change is to ask all the players to report an item of their demand set simultaneously (instead of sequentially). This change alone is not enough as in the worst case many players might report the same item and hence the number of rounds might still be \( \Omega(\frac{n}{2^k}) \). Hence we ask each player to report a random item of his demand instead. In the full version we prove the following theorem:

**Theorem 4.1.** After \( O(\frac{\log n}{\sqrt{n^\epsilon}}) \) rounds the algorithm above provides an approximation ratio of \((1 + \epsilon)\).

It is worth noting a different version of the auction algorithm which was discussed in [8]. In this version at every round each player reports its entire demand set (simultaneously with the other players), then a minimal set of over demanded items is computed and only their prices are increased. While the number of rounds for this algorithm might be small the communication cost of each round can be linear in \( n \).

## 5. A Lower Bound for Subadditive Combinatorial Auctions

We now move to discuss combinatorial auctions with subadditive bidders. In particular, in this section we prove our most technical result:

**Theorem 5.1.** No randomized simultaneous protocol for combinatorial auctions with subadditive bidders where each bidder sends sub-exponentially many bits can approximate the social welfare to within a factor of \( m^{1/2-\epsilon} \), for every constant \( \epsilon > 0 \).

We will actually prove the lower bound using only XOS valuations, a subclass of subadditive valuations. We present a distribution over the inputs and show that any deterministic algorithm in which each bidder sends sub-exponentially many bits cannot approximate the expected social welfare to within a factor of \( m^{1/2-\epsilon} \) for this specific distribution (where expectation is taken over the distribution). This implies, by Yao’s principle, that there is no randomized algorithm that achieves an approximation ratio of \( m^{1/2-\epsilon} \) for every constant \( \epsilon > 0 \).

Our hard distribution which we denote by \( D \) is the following:

- \( n = k^3 \) players, \( m = k^3 + k^4 \) items.
- Each player \( i \) gets a family \( F_i \) of size \( t = e^{2k} \) of sets of \( k \) items. The valuation of player \( i \) is defined to be: \( v_i(S) = \max_T \epsilon_T S \subseteq T \cap F_i \). Observe that this is an XOS valuation where each set \( S \subseteq F_i \) defines a clause in which all items in \( S \) have value 1 and the rest of the items have value 0.
- The families \( F_i \) are chosen, in a correlated way, as follows: first, a center \( C \) of size \( k^3 \) is chosen at random; then for each player \( i \) a petal \( P_i \) of size \( k^2 \) is chosen at random from the complement of \( C \). Now, for each player \( i \), the family \( F_i \) is chosen as follows: one set \( T_i \) of size \( k \) is chosen at random from \( P_i \) and \( t - 1 \) sets of size \( k \) are chosen at random from \( C \cup P_i \).
- The players do not know \( C \) nor do they know which of the sets was chosen from \( P_i \). We may assume without loss of generality that each player \( i \) knows the set \( C \cup P_i \).
Each player sends, deterministically, simultaneously with the others, at most $l$ bits of communication just based on his input $F_i$. A referee that sees all the messages chooses an allocation $A_1, \ldots, A_k$ of the $m$ items to the $n$ players (only based on the messages), with the $A_i$’s being disjoint sets. We assume without loss of generality that all items are allocated.

In order to prove that no deterministic algorithm can obtain a good approximation for instances drawn from $D$, we show that to get a good approximation we must identify $T_i$ for almost all of the players. This would have been easy had each player could have distinguished between the items in $C$ and $P_i$, but this information is missing. We show that for the central planner to successfully identify even a single $T_i$ player $i$ has to send exponentially many bits. Formally, this is done by reducing the two-player “set seeking” problem that we define below to the multi-player combinatorial auction problem. The main technical challenge is to prove the hardness of the set seeking problem.

The next couple of lemmas together gives us Theorem 5.1. The proof of the first lemma is easy and provided in the appendix. The second lemma is the heart of the lower bound.

**Lemma 5.2.** With very high probability, over this distribution $(D)$, there exists an allocation with social welfare $\sum_i v_i(A_i) = \Theta(k^3)$.

**Lemma 5.3.** Every deterministic protocol for the combinatorial auction problem with $l < t'$ produces an allocation with $\sum_i v_i(A_i) = t\Omega(\log^2 n)$ in expectation (over the distribution).

To prove Lemma 5.3 we first define a two-player “set seeking” problem and show its hardness (Subsection 5.1). Next, we reduce the two-player set seeking problem to our multi-player combinatorial auction problem (Subsection 5.2).

### 5.1 The Two Player “Set Seeking” Problem

The “Set Seeking” problem includes two players and $x = k^2 + k^3$ items. One of the players plays the role of the keeper and gets as an input a family of $t = e^{2k^2}$ sets $F$ where all the sets are of size $k$. The other player plays the role of the seeker and gets a set $P$ of size $k^2$. In this problem, first the keeper sends a message of at most $l$ bits (advice). Next, based on this message the seeker outputs some set $A \subseteq P$, $|A| \geq k$. The goal is to maximize $\max_{X \in F} \frac{|A \cap X|}{|A|}$.

We will analyze the performance of deterministic algorithms on a specific distribution $(D_2)$ for this problem which we now define. This distribution is based on choosing two sets, $F$ and $P$ which are chosen in correlation as follows:

1. A set $P$ of size $k^2$ is chosen uniformly at random from all items.
2. The set $F$ is constructed by choosing uniformly at random a special set $T_F$ of size $k$ from $P$ and additional $t - 1$ sets of size $k$ from all items.

**Lemma 5.4.** If $l \leq t'$ then there is no $k^{1-\epsilon}$-approximation for the set seeking problem.

**Proof.** Fix a message $m$ and let $A_m(P)$ denote the set $A$ that the seeker returns when his input is $P$ and the keeper sends a message $m$.

**Definition 5.5.** Fix a message $m$ and a set $P$, $|P| = k^2$. A set $S$ is $(m, P)$-compatible if $|A_m(P) \cap S| \geq \frac{|A_m(P)|}{k^{1-\epsilon}}$.

**Claim 5.6.** Fix a message $m$ and a set $P$, $|P| = k^2$. The probability that a set $S$ which is chosen uniformly at random from $P$ is $(m, P)$-compatible is at most $2e^{-k^2}$.

**Proof.** We fix a set $A_m(P) = A \subseteq P$ and compute the probability that the intersection of this set with a set $S$ of size $k$ chosen uniformly at random contains $\frac{|A|}{k^{1-\epsilon}}$ items. The probability of each element in $S$ to be in $A \cap S$ is exactly $\frac{|A|}{k^{1-\epsilon}}$. Thus we expect that $|A \cap S| \approx \frac{|A|}{k}$. We now use Chernoff bounds to make this precise.

Consider constructing the following random set $T$: $k$ items are selected so that each item is chosen uniformly at random among the $k^2$ items of $P$. $T$ is similar to the way $S$ is constructed, except that it possibly contains less than $k$ items as there is some positive chance that some item in $P$ will be selected twice. We conservatively assume that every item that was selected twice is in $A$. Thus, if we bound the probability that $|A \cap T|$ is too large, we also bound the probability that $|A \cap S|$ is too large.

We first bound the probability that more than $k^\epsilon$ items are selected at least twice to $T$. Since there are at most $k$ items in $T$, in the $k^\epsilon$th item that we have the select, the probability that we will choose an already-selected item is at most $\frac{k}{k^2} = \frac{1}{k}$. By the Chernoff bounds, the probability that more than $k^\epsilon$ items are selected twice to $T$ is at most: $\frac{1}{ek^\epsilon}$.

To compute the expected intersection with $A$, assume independent variables, each variable $y_i$ is true with probability $\frac{|A|}{k^{1-\epsilon}}$ and false with probability $1 - \frac{|A|}{k^{1-\epsilon}}$. Observe that $Y = \sum_i y_i$ is distributed exactly as $|T \cap A|$. The expectation of $Y$ is $\frac{|A|}{k}$. Now by Chernoff bounds we get that: $\Pr[Y > k^\epsilon \cdot \frac{|A|}{k}] \leq \left(\frac{1}{ek^\epsilon}\right)^{\frac{|A|}{k}} \leq \frac{1}{ek^\epsilon}.$ The last transition holds by the assumption that $|A| \geq k$.

We have that with probability of at most $\frac{2}{ek^\epsilon}$ we have that $|A \cap T| \leq 2e^{-k^\epsilon}$. By our discussion above, with at most the same probability we have that $|A \cap S| \leq 2e^{-k^\epsilon}$. □

**Definition 5.7.** Let $F$ be a family of sets of size $k$, $|F| = t = e^{2k^2}$. We say that a message $m$ is $F$-good if $\Pr[|A_m(P) \cap T| \geq \frac{|A_m(P)|}{k^{1-\epsilon}}] \geq e^{-\frac{1}{2}k^2}$, where $T$ is chosen uniformly at random from $F$ and $P$, $|P| = k^2$, contains $T$ and $k^2 - k$ items chosen uniformly at random from the rest of the items.

**Claim 5.8.** For every message $m$, the probability that $m$ is $F$-good is at most $p' = e^{\left(-\frac{1}{2}k^2\right)e^{k^2}}$ where the sets in $F$ are chosen uniformly at random.

**Proof.** Fix a message $m$. Consider $F$ where the sets in $F$ are chosen uniformly at random. We first compute the probability that for a single $T \in F$ we have that $|A_m(P) \cap T| \geq \frac{|A_m(P)|}{k^{1-\epsilon}}$. Observe that every set $T \in F$ in this setting can be thought of as chosen uniformly at random from a fixed set $P$. Thus that probability is the same as the probability that $T$ is $(m, P)$-compatible, which is $e^{-k^2}$ by Claim 5.6.
Now we would like to compute the probability that \( m \) is \( F \)-good, that is the probability that there exist at least \( e^{-\frac{1}{2}k^2} \cdot |F| = e^{\frac{1}{2}k^2} \) sets \( T \in F \) such that \( |A_m(P) \cap T| \geq \frac{|A_m(P)|}{k^{1-\epsilon}} \). The expected number of such sets is \( e^{-k^2} \cdot |F| = e^{k^2} \). By the Chernoff bounds \((\mu = e^{k^2}, \delta = e^{\frac{1}{2}k^2})\) this probability is at most \( p' \).

We can now finish the proof. Choose \( F \) at random. For every \( m \) the probability that \( m \) is \( F \)-good is at most \( p' \). The message length is \( l \), and the total number of messages is therefore at most \( 2^l \). Thus, by the union bound, the probability that there exists some message \( m \) which is \( F \)-good is at most \( 2^l \cdot p' \leq 2^{\epsilon^2} \cdot p' = 2^{\epsilon^2} \cdot e^{-k^2} \cdot \left(2e^{-\frac{1}{2}k^2}\right)^{\frac{1}{2}k^2} < e^{-\frac{1}{2}k^2} \), where the transition before the last uses the fact that \( \epsilon < 1 \). Hence, with probability at least \( 1 - e^{-\frac{1}{2}k^2} \) every message \( m \) is not \( F \)-good for the randomly chosen \( F \). This in turn implies that for a family \( F \) and for every message \( m \) the probability for \( P \) and \( T \) chosen as in Definition 5.7 that \( T \) is \((m, P)\)-compatible is at most \( e^{-\frac{1}{2}k^2} \). Thus, with probability at least \( \frac{1}{2} - e^{-\frac{1}{2}k^2} \) we have that \( |A_m(P) \cap T| \leq \frac{|A_m(P)|}{k^{1-\epsilon}} \). This implies that with probability \( 1 - e^{-\frac{1}{2}k^2} - e^{-\frac{1}{2}k^2} \) the approximation ratio is at most \( k^{1-\epsilon} \). Even if with probability \( e^{-\frac{1}{2}k^2} + e^{-\frac{1}{2}k^2} \) the approximation ratio is 1, the expected approximation ratio is at most \( \frac{1}{2} k^{1-\epsilon} \).

### 5.2 The Reduction (Set Seeking \( \rightarrow \) Combinatorial Auctions with XOS Bidders)

Given the hardness of the set seeking problem, we will be able to derive our result for combinatorial auctions using the following reduction:

**Lemma 5.9.** Any protocol for combinatorial auctions on distribution \( D \) that achieves approximation ratio better than \( m^{\frac{1}{4} - \epsilon} \) where the message length of each player is \( l \) can be converted into a protocol for the two-player set seeking problem on distribution \( D_2 \) achieving an approximation ratio of \( k^{1-\epsilon} \) with the same message length \( l \).

**Proof.** In the proof we fix an algorithm for the multi-player combinatorial auction problem and analyze its properties.

**Definition 5.10.** Fix an algorithm for the XOS problem and consider the distribution \( D \). We say that player \( i \) is good if \( E[|A_i \cap T_i|] \geq \max \{ \frac{|A_i|}{k^{1-\epsilon}}, k^k \} \).

To prove the lemma we first show that if none of the players are good then the algorithms approximation ratio is bounded by \( m^{\frac{1}{4} - \epsilon} \). Else, there exists at least a single player which is good. In this case we show how the algorithm for combinatorial auctions can be used to get a good approximation ratio for the set seeking problem.

**Claim 5.11.** If none of the players are good then the expected approximation ratio is at most \( m^{\frac{1}{4} - \epsilon} \).

**Proof.** To give an upper bound on the expected social welfare we assume that the \( k^3 \) items in the center are always allocated to players that demand them. We now compute an upper bound on the contribution of the remaining \( k^k \) items to the expected social welfare. Observe that since none of the players is good, each player contributes at most \( \frac{|A_i|}{k^{1-\epsilon}} + k^k \) to the expected social welfare (of the \( k^k \) items). Hence the expected social welfare achieved by the algorithm is at most \( \sum \frac{|A_i|}{k^{1-\epsilon}} + k^k + |C| \leq \frac{k^4}{k^{1-\epsilon}} + k^k + k^k \leq 3k^{1-\epsilon} \). This implies that the approximation ratio of the algorithm is \( k^{1-\epsilon} \leq m^{\frac{1}{4} - \epsilon} \).

**Claim 5.12.** If there exists a good player then there exists an algorithm for the two-player set seeking problem that guarantees an approximation ratio of \( k^{1-\epsilon} \) with the same message length.

**Proof.** Let player \( i \) be the good player. Recall that \( D \) is the distribution for the multi-player combinatorial auction problem and that \( D_2 \) is the distribution defined for the set seeking problem. We denote by \( E_{D_1}[\cdot] \) and \( E_{D_2}[\cdot] \) expectations taken over the distributions \( D \) and \( D_2 \) respectively. We show that there exists an algorithm for the set seeking problem achieving expected approximation ratio of \( k^{1-\epsilon} \) on \( D_2 \).

Let the keeper take the role of player \( i \) in the multi-player algorithm and the seeker play the roles of the rest of the \( n - 1 \) players. More precisely, the keeper first sends player \( i \)'s message to the seeker. This is possible as the input of the keeper is identical to the input of the players in the multi-player problem. Next, the seeker simulates the messages of the remaining players and run the algorithm internally. This simulation is possible by the assumption that in the multi-player algorithm all messages are sent simultaneously. The number of items in the combinatorial auction will be \( k^3 + k^k \), where the \( x \) items of the set seeking problem will correspond to some set \( X \) of size \( x \) of items in the combinatorial auction. We first show that given the information of the seeker and keeper is drawn in a correlated way from \( D_2 \), they have enough information to simulate the correlated distribution \( D \).

The input of the keeper is defined in a straightforward way, where each set in \( F \) defines a clause in the XOS valuation. All items in these clauses are subsets of \( X \). The seeker constructs the valuations of the other \( n - 1 \) players as follows: the items that are in \( X \setminus P \) form the center. Next the seeker chooses uniformly at random for each player \( j \) a petal \( P_j \) of the \( k^4 \) items not in the center, a set \( T_j \subseteq P_j \) of size \( k^3 \) and additional \( t - 1 \) sets of size \( k^k \) from \( C \cup P_j \). Observe that the distribution of valuations constructed this way is identical to \( D \). The inherent reason for this is that for player \( i \) the distribution of \( P_i \) and \( T_i \) is identical to the distribution of \( P \) and \( T \) in \( D_2 \) as in both cases \( P \) and \( T \) are chosen uniformly at random from a set of size \( k^3 + k^k \) and \( T \) (or \( T_i \)) is chosen uniformly at random from \( P \) (or \( P_i \)). In other words, the distribution \( D_2 \) on \( P \) and \( T \) is identical to distribution \( D \) projected on \( P_i \), \( P_i \) and \( T_i \).

We now observe that since \( i \) is a good player we have that \( E_{D_2}[|A_i \cap T_i|] \geq \max \{ \frac{|A_i|}{k^{1-\epsilon}}, k^k \} \). We show that this implies an algorithm achieving expected approximation ratio of \( k^{1-\epsilon} \) for the two-player set seeking problem. The algorithm works as follows: we first perform the reduction above, and therefore the distribution we are analyzing is \( D \). Now, if player \( i \) was assigned a bundle \( A_i \) of size at least \( k \), the algorithm returns \( A = A_i \). Else, the algorithm returns a bundle \( A \) that contains \( A_i \) and additional \( k - |A_i| \) arbitrary items.

Thus, we have that \( E_{D_2}[|A \cap T_i|] \geq \frac{|A|}{k^{1-\epsilon}} \), as we made sure that \( |A| \geq k \) implying that \( \frac{|A|}{k^{1-\epsilon}} \geq k^k \).

We claim that the expected approximation ratio of the algorithm on \( D_2 \) is \( k^{1-\epsilon} \). As the distribution \( D_2 \) on \( P \) and \( T \) is identical to distribution \( D \) projected on \( P_i \), \( P_i \) and \( T_i \), we have that \( E_{D_2}[|A \cap T_i|] \geq \frac{|A|}{k^{1-\epsilon}} \). Thus, \( E_{D_2}[|A \cap T|] \geq \frac{1}{k^{1-\epsilon}} \). This in turn implies that \( \max_{S \subseteq F} |A \cap S| \geq \frac{1}{k^{1-\epsilon}} \) and since the optimal solution has
a value of 1 the expected approximation ratio is $k^{1-t^*}$. □

From the last two claims we get that either the algorithm for the combinatorial auction problem does not guarantee a good approximation ratio, or that we have constructed an efficient protocol for the set seeking problem.

6. ALGORITHMS FOR SUBADDITIVE COMBINATORIAL AUCTIONS

We design algorithms for a restricted special case we refer to as “t-restricted” instances (see definition below). We will show however that the existence of a simultaneous algorithm for t-restricted instances implies a simultaneous approximation algorithm for subadditive bidders with almost the same approximation ratio.

**DEFINITION 6.1.** Consider an XOS valuation $v(S) = \max_i a_i(S)$, where each $a_i$ is an additive valuation. $v$ is called binary if for every $a_i$ and item $j$ we have that $a_i\{j\} \in \{0, \mu\}$, for some $\mu$.

**DEFINITION 6.2.** An instance of combinatorial auctions with binary XOS valuations (all with the same $\mu$, for simplicity and without loss of generality $\mu = 1$) is called t-restricted if there exists an allocation $(A_1, \ldots, A_n)$ such that all the following conditions hold:

1. For every $i$, $v_i(A_i) = |A_i|$.
2. For every $i$, either $|A_i| = t$ or $|A_i| = 0$.
3. $t$ is a power of 2.
4. $\sum_i v_i(A_i) \geq \frac{OPT}{2 \log m}$.

**PROPOSITION 6.3.** If there exists a simultaneous algorithm for t-restricted instances that provides an approximation ratio of $\alpha$ where each bidder sends a message of length $l$, then there exists a simultaneous algorithm for subadditive bidders that provides an approximation ratio of $O(\alpha \cdot \log^2 m)$ where each bidder sends a message of length $O(l \cdot \log^3 m)$.

The proof of Proposition 6.3 can be found in the full version of the paper.

6.1 A Simultaneous $O(m^{\frac{1}{2}})$-Approximation

We show that simultaneous algorithms can achieve better approximation ratios than those that can be obtained by sketching the valuations. Specifically, we prove that:

**THEOREM 6.4.** There is a deterministic simultaneous algorithm for combinatorial auctions with subadditive bidders where each player sends $\text{poly}(m, n)$ bits that guarantees an approximation ratio of $O(m^{\frac{1}{2}})$.

Given Proposition 6.3, we may focus only on designing algorithms for t-restricted instances. The algorithm for t-restricted instances is simple:

1. Each player reports a maximal set of disjoint bundles $S_i$ such that for every bundle $S \in S_i$, $|S| = \frac{t}{2}$ and $v_i(S) = |S|$.
2. For each $i$, let $v'_i$ be the following XOS valuation: $v'_i(S) = \max_{T \in S_i} |T \cap S|$.
3. Output $(T_1, \ldots, T_n)$ – the best allocation\(^3\) with respect to the $v'_i$’s.

Notice that the size of the message that each player sends is $\text{poly}(m)$. Furthermore, for each bundle $S$ and bidder $i \ v_i(S) \geq v'_i(S)$. We will show that the best allocation with respect to the $v'_i$’s provides a good approximation with respect to the original valuations $v_i$’s.

I.e., $\sum_i v_i(T_i) = O(m^{\frac{1}{2}})$.

The proof considers three different allocations and shows that each allocation provides a good approximation for a different regime of parameters. See the full version of the paper for formal statements and proofs.

- The best allocation (with respect to the $v'_i$’s) in which each player receives at most one item. We show that this provides an $O(t)$ approximation with respect to the $v_i$’s.
- Each player $i$ is allocated the fraction of the bundle $T \in S_i$ that maximizes $|T \cap A_i|$. We show that this allocation guarantees an approximation ratio of $O(l)$, for some $l$ related to the $l_i$’s.
- The third allocation is constructed randomly (even though our algorithm is deterministic): each player $i$ chooses at random a bundle $S_i$, and each item $j$ is allocated to some player that $j$ is in his randomly selected bundle, if such exists. Let $n'$ be the number of nonempty bundles in $(A_1, \ldots, A_n)$. We show that the expected approximation of this allocation is $O(n'/l)$, for the same $l$ as above.

By choosing the best of these three allocations we get the desired approximation ratio:

**CLAIM 6.5.** Suppose that we have three allocations $B^1, B^2$ and $B^3$ such that: $\sum_i v_i(A_i) = O(t)$, $\sum_i v_i(A_i) = O(l)$ and $\sum_i v_i(A_i) = O(n'/l)$. Then, $\sum_i v_i(A_i) = O(m^{\frac{1}{2}})$.

**PROOF.** By the first two claims, we get an approximation ratio of $O(m^{\frac{1}{2}})$ whenever $l \leq m^{\frac{1}{2}}$ or $t \leq m^{\frac{1}{2}}$. Hence, we now assume that $l, t \geq m^{\frac{1}{2}}$. Now observe that when $t \geq m^{\frac{1}{2}}$ and $n' \leq m^{\frac{1}{2}}$, since there can be at most $m/t$ players that receive non-empty (disjoint) bundles in any allocation where the size of each non-empty bundle is at least $t$. We therefore have that $\frac{n'}{l} \leq m^{\frac{1}{2}}$ and the lemma follows by the third claim. □

6.2 A t-Round Algorithm

We now develop an algorithm that guarantees an approximation ratio of $O(m^{\frac{1}{2t}})$ for combinatorial auctions with subadditive valuations in $k$ rounds. In each of the rounds each player sends $\text{poly}(m)$ bits. We provide an algorithm for t-restricted instances (see Section 6.1 for a definition). By Proposition 6.3 this implies an algorithm with almost the same approximation ratio for general subadditive valuations.

The Algorithm (for t-restricted instances)
1. Let $N_1 = N, U_1 = M$ and $U_{i+1} = M$.

\(^3\)As stated, this algorithm uses polynomial communication but may not run in polynomial time since finding the optimal solution with explicitly given XOS valuations is NP hard. If one requires polynomial communication and time, an approximate solution may be computed using any of the known constant ratio approximation algorithms. The analysis remains essentially the same, with a constant factor loss in the approximation ratio.
2. In every round $r = 1, \ldots, k$:
   (a) Each player reports the maximal set of disjoint bundles $S_{r,i}$ such that for every bundle $S \in S_{r,i}$:
       $|S| \leq \frac{|S|}{m}$ and $v(S) = |S|$. 
   (b) Go over the players in $N_r$ in an arbitrary order. For every player $i$ for which there exists a bundle $S \in S_{r,i}$
       such that at least $\frac{1}{m}$ of its items were not allocated yet, allocate player $i$ the remaining unallocated items of $S$.
   (c) Let $N_{r+1} \subseteq N_r$ be the set of players that were not allocated items at round $r$ or before.
   (d) Let $U_{r+1} \subseteq U_r$ be the set of items that were not allocated at round $r$ or before.
   (e) Let $U_{r+1} = (\cup_{S \in S_{r,i}} S) \cap U_{r+1}$.

In the full version we prove the following theorem:

**Theorem 6.6.** For every $k \leq \log m$, there exists an algorithm for $t$-restricted instances that provides an approximation ratio $O(k \cdot m^{\frac{t+1}{t}})$ in $k$ rounds where each player sends $\text{poly}(m, n)$ bits. In particular, when $k = O(\log m)$ the approximation ratio is $O(\log m)$. As a corollary, there exists a $k$-round approximation algorithm for subadditive valuations that provides an approximation ratio of $O(k \cdot m^{\frac{t+1}{t}} \cdot \log^3 m)$.

7. REFERENCES

[1] Moshe Babaioff, Brendan Lucier, Noam Nisan, and Renato Paes Leme. On the efficiency of the Walrasian mechanism. CoRR, abs/1311.0924, 2013.
[2] Ashwinkumar Badanidiyuru, Shahar Dobzinski, Hu Fu, Robert Kleinberg, Noam Nisan, and Tim Roughgarden. Sketching valuation functions. In SODA, 2012.
[3] Maria-Florina Balcan, Florin Constantin, Satoru Iwata, and Lei Wang. Learning valuation functions. Journal of Machine Learning Research - Proceedings Track, 23:4.1–4.24, 2012.
[4] Maria-Florina Balcan and Nicholas J. A. Harvey. Learning submodular functions. In STOC, 2011.
[5] Dimitri P Bertsekas. The auction algorithm: A distributed relaxation method for the assignment problem. Annals of operations research, 14(1):105–123, 1988.
[6] Kshipra Bhawalkar and Tim Roughgarden. Welfare guarantees for combinatorial auctions with item bidding. In SODA, 2011.
[7] G. Christodoulou, A. Kovacs, and M. Schapira. Bayesian combinatorial auctions. Automata, Languages and Programming, pages 820–832, 2010.
[8] Gabrielle Demange, David Gale, and Marilda Sotomayor. Multi-item auctions. The Journal of Political Economy, pages 863–872, 1986.
[9] Shahar Dobzinski and Michael Schapira. An improved approximation algorithm for combinatorial auctions with submodular bidders. In SODA, 2006.
[10] Danny Dolev and Tomás Feder. Multi-party communication complexity. In Foundations of Computer Science, 1989. 30th Annual Symposium on, pages 428–433. IEEE, 1989.
[11] Uriel Feige. On maximizing welfare where the utility functions are subadditive. In STOC, 2006.
[12] Uriel Feige and Jan Vondrák. Approximation algorithms for allocation problems: Improving the factor of 1-1/e. In FOCS, 2006.
[13] Michal Feldman, Hu Fu, Nick Gravin, and Brendan Lucier. Simultaneous auctions are (almost) efficient. In Proceedings of the 45th annual ACM Symposium on theory of computing, pages 201–210. ACM, 2013.
[14] 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.
[15] Avinatan Hassidim, Haim Kaplan, Yishay Mansour, and Noam Nisan. Non-price equilibria in markets of discrete goods. In ACM Conference on Electronic Commerce, pages 295–296, 2011.
[16] F. A. Hayek. The use of knowledge in society. American Economic Review, 35(4):519–530, 1945.
[17] John E Hopcroft and Richard M Karp. An n^5/2 algorithm for maximum matchings in bipartite graphs. SIAM Journal on computing, 2(4):225–231, 1973.
[18] 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.
[19] Michael Kapralov. Better bounds for matchings in the streaming model. arXiv preprint arXiv:1206.2269, 2012.
[20] 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.
[21] Richard M Karp, Umesh Vazirani, and Vijay V Vazirani. An optimal algorithm for on-line bipartite matching. In Proceedings of the twenty-second annual ACM symposium on Theory of computing, pages 352–358. ACM, 1990.
[22] O. Lange and F. Taylor. On the Economic Theory of Socialism. University of Minnesota Press, 1938.
[23] B. Lehmann, D. Lehmann, and N. Nisan. Combinatorial auctions with decreasing marginal utilities. Games and Economic Behavior, 55(2):270–296, 2006.
[24] Renato Paes Leme, Vasilis Syrgkanis, and Éva Tardos. The curse of simultaneity. In ITCS, pages 60–67, 2012.
[25] 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.
[26] Marcin Mucha and Piotr Sankowski. Maximum matchings via gaussian elimination. In FOCS, pages 248–255, 2004.
[27] Ketan Mulmuley, Umesh 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.
[28] Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. Journal of Economic Theory, 129(1):192–224, 2006.
[29] Alvin E Roth and Marilda A Oliveira Sotomayor. Two-sided matching: A study in game-theoretic modeling and analysis. Number 18. Cambridge University Press, 1992.
[30] Jan Vondrák. Optimal approximation for the submodular welfare problem in the value oracle model. In STOC, 2008.
[31] Raphael Yuster. Maximum matching in regular and almost regular graphs. Algorithmica, 66(1):87–92, 2013.
APPENDIX

A. COMMUNICATION COMPLEXITY OF BIPARTITE MATCHING

In this appendix we discuss a proposed communication complexity investigation of the bipartite matching problem. This model is essentially the same as that used in our investigation of bipartite matching in the rest of the paper but focusing on the exact problem rather than on approximations, and proposing the study of communication as of itself rather than merely as an abstraction of market processes.

We focus on the open problem(s) and shortly mention some related models where communication bottlenecks for matching have been investigated and give a few pointers to different such threads where the interested reader may find many more references.

A.1 The Model and Problem

There are \(n\) items and \(n\) players. Each player \(i\) holds a subset \(S_i\) of items that he is interested in. I.e. we have a bipartite graph with \(n\) left vertices (players) and \(n\) right vertices (items), and have a player in our model for each left vertex, a player that knows the set of neighbors of his vertex in the bipartite graph (but there are no players associated with the items.) The goal of these players is to find a maximum matching between items and players, i.e. that each player is assigned a single item \(j_i \in S_i\) with no items assigned to multiple players \(j_i \neq j_{i'}\) for \(i \neq i'\).

Communication Model

The players engage in a fixed communication protocol using broadcast messages. Formally they take turns writing on a common "blackboard." At every step in the protocol, the identity of the next player \(i\) to write a bit on the blackboard must be completely determined by the contents of the blackboard, and the message written by this player \(i\) must be determined by the contents of the blackboard as well as his own private input \(S_i\). Whether the protocol terminates at a given point must be completely determined by the contents of the blackboard, and at this point the output matching must be solely determined by the contents of the blackboard. This model is completely equivalent to a decision tree, each query can be an arbitrary function depending only on a single player’s information \(S_i\). The measure of complexity here is the total number of bits communicated.

Rounds

The communication model above allows an arbitrary order of communication. Of interest are also various limited orders: oblivious (the order of speaking is fixed independently of the input), and the simplest special case of it, one-way communication where the players speak in the fixed order of player 1, player 2, etc. We will focus on speaking in rounds: in each "round" each of the \(n\) players writes a message on the blackboard, a message that may depend on his own input as well as the messages of the others in previous rounds (i.e. on the contents of the blackboard after the previous round). The measures of complexity here are the number of rounds and the total number of bits communicated. The special case of a single round is called a simultaneous protocol.

Open Problems

How much communication is needed for finding a maximum matching? What if we are limited to \(r\) rounds? These questions apply both to deterministic and to randomized protocols. The tradeoff between communication and approximation is of course also natural to explore.

What is Known

The trivial upper bound for communication is \(n^2\) since players can all simultaneously send their full input. The trivial lower bound is \(\Omega(n \log n)\) as this is the number of bits necessary to represent the output matching (and every matching may need to be given as output).

Significantly, the non-deterministic (and co-non-deterministic) communication complexity is also \(O(n \log n)\): to verify that a given matching is maximum size it suffices to add a Hall-theorem blocking set, or alternatively a solution for the dual. Specifically, a specification of a set of "high-price" items, so that \(1\) only allocated items are high-price \(2\) all players that are not allocated a low price item are only interested in high-price ones. The fact that the non-deterministic complexity is low means that "easy" lower bounds techniques such as fooling-sets or cover-size bounds will not suffice for giving good lower bounds.

Interestingly, an \(O(n^{\frac{3}{5}} \log n)\) upper bound can be obtained by adapting known algorithms to this framework: First, the auction algorithm described in Section 4 gives a \((1 - \delta)\)-approximation using \(O(n \log n / \delta)\) communication. When we choose \(\delta = 1 / \sqrt{n}\) this means that we get a matching that is at most smaller than the optimal one by an additive \(\sqrt{n}\). We can thus perform \(\sqrt{n}\) more augmenting path calculations to get an optimal matching. Each augmenting path calculation requires only \(O(n \log n)\) bits of communication: it requires finding a path in a graph on the players that has a directed edge between player \(i\) and player \(j\) whenever \(i\) is interested in an item that is currently allocated to \(j\). The goal here is to find a path from any player that is not allocated an item to any player that is interested in an unallocated item. A breadth first search with the blackboard serving as the queue requires writing every vertex at most once on the blackboard, at most \(O(n \log n)\) communication.

We do not know any better upper bound, nor do we know a better than \(O(n^2)\) upper bound for even \(n\) rounds. Our lower bounds for matching provide an \(\Omega(n^2)\) lower bound for simultaneous protocols, and a \(n^{1+\Omega(1/\log \log n)}\) lower bound for one-way communication follows from [14]. We don’t know any lower bound better than \(\Omega(n \log n)\) for general protocols or even for 2-round protocols.

Algorithmic Implications

We believe that studying the bipartite matching under this model may be a productive way of understanding the general algorithmic complexity of the problem. A major open problem is whether bipartite matching has a \((\text{nearby})\)-linear time algorithm: \(O(n^{2+o(1)})\) time for dense graphs (and maybe \(O(n^{1+o(1)})\) for graphs with \(m\) edges). The best deterministic running time known (for the dense case) is the 40-year old \(O(n^{2.5})\) algorithm of [17], with a somewhat better randomized \(O(n^\omega)\) algorithm known [26] (where \(\omega = 2.3...\) is the matrix multiplication exponent). For special cases like regular or near-regular graphs nearly linear times are known (e.g. [31]). In parallel computation, a major open problem is whether bipartite matching can be solved in parallel poly-logarithmic time (with a polynomial amount of processors). Randomized parallel algorithms for the problem [27, 20] have been known for over 25 years.

On the positive side, it is "likely" that any communication protocol for bipartite matching that improves on the currently known \(O(n^{1.5})\) complexity will imply a faster than the currently known \(O(n^{2.5})\) algorithm. This is not a theorem, however the computational complexity needed to send a single bit in a communication protocol is rarely more than linear in the input held by the player sending the bit. Most often each bit is given by a very simple com-
In our model, the complexity of bipartite matching is completely open, and in particular the communication complexity of the decision problem of whether the input graph has a perfect matching is open with no known non-trivial, $\omega(n)$, lower bounds or non-trivial, $o(n^{1.5})$ upper bounds.

**Streaming and Semi-streaming**

One of the main applications of communication complexity is to serve as lower bounds for “streaming” algorithms, those are algorithms that go over the input sequentially in a single pass (or in few passes), while using only a modest amount of space. The model of communication complexity required for such lower bounds is that of a one-way single-round private-channel protocol where in step $i$ player $i$ sends a message to player $i+1$. (For $r$-pass variants of streaming algorithms, we will have $r$ such rounds of one-way communication.) The lower bounds mentioned above thus imply that no streaming algorithm that uses $o(n)$ space can get even a constant factor-approximation of the maximum matching, even with $O(1)$ rounds. A greedy algorithm gets $1/2$-approximation in a single round using $O(n)$ space, and slight improvements in the approximation factor using linear space are possible, e.g. using the online matching algorithm of [21]. In $r$ passes and nearly-linear space, [19] gets an $1 - O(1/\sqrt{r})$ approximation. Streaming algorithms that use linear or near-linear space are usually called semi-streaming algorithms and lower bounds for them are usually derived by looking at the information transfer between the “first half” and the “second half” of the input data and proving a significantly super-linear lower bound on the one-way two-party communication. This was done in [14] who give a $n^{1+\Omega(1/\log \log n)}$ lower bound for improving the $2/3$ approximation. To the best of our knowledge no better lower bound is known even for getting an exact maximum matching.

**Distributed Computing**

In this model the input graph is also the communication network. I.e. players can communicate with each other only over links that are edges in the input graph, and the interest is the number of rounds needed. For this to make sense in a bipartite graph we need to also have processors for the right-vertices of the graph (and thus every edge is known by the two processors it connects.) It is not hard to see that to get a perfectly maximal matching in this model transfer of information across the whole diameter of the graph may be needed, which may require $\Omega(n)$ rounds of communication, but in [25] a protocol is exhibited that gives a $\left(1 - \frac{1}{\sqrt{n}}\right)$-approximation in $O(\log n / \sqrt{n})$ rounds.