Settling the Communication Complexity of Combinatorial Auctions with Two Subadditive Buyers

Tomer Ezra, Michal Feldman
Computer Science Department
Tel-Aviv University
Tel-Aviv, Israel
Email: tomer.ezra@gmail.com, mfeldman@tau.ac.il

Eric Neyman
Computer Science Department
Columbia
New York, NY
Email: eric.neyman@columbia.edu

Inbal Talgam-Cohen
Computer Science Department
Technion
Haifa, Israel
Email: italgam@cs.technion.ac.il

Matt Weinberg
Computer Science Department
Princeton University
Princeton, NJ
Email: smweinberg@princeton.edu

***Abstract*—We study the communication complexity of welfare maximization in combinatorial auctions with m items and two players with subadditive valuations. We show that outperforming the trivial $1/2$-approximation requires exponential communication, settling an open problem of Dobzinski, Nisan and Schapira [STOC'05, MOR'10] and Feige [STOC'06, SICOMP '09].**

To derive our results, we introduce a new class of subadditive functions that are "far from" fractionally subadditive (XOS) functions, and establish randomized communication lower bounds for a new "near-EQUALITY" problem, both of which may be of independent interest.

I. INTRODUCTION

Combinatorial auctions have been a driving force of Algorithmic Game Theory (AGT) since its inception: how should one allocate goods among interested parties? In a combinatorial auction, a central designer has a set M of m indivisible goods to allocate, and each of n players has a valuation (set) function $v_i : 2^M \to \mathbb{R}_+$, which is private (known only to player i). We wish to partition the items in a way that maximizes the *social welfare* $\sum_i v_i(S_i)$, where S_i denotes the items received by player i in the partition.

This fundamental problem has received significant attention in various models: with or without incentives, with or without restrictions on valuations, with or without computational limits on the players, etc. In this paper we prove standard communication lower bounds when two players have *subadditive* valuations.[1] That is, our lower bounds rule out the existence of good mechanisms even when players honestly follow the intended protocol, are computationally unbounded, and are assumed to have subadditive valuation functions. Subadditive set functions are central to Algorithmic Game Theory as the frontier of complement-freeness (e.g. [FFGL13], [RW15], [CM16], [CZ17]), and are natural mathematical objects of study in algorithm design broadly (e.g. [BDF+12], [RS17]).

Previous work: The study of combinatorial auctions through the lens of communication complexity was pioneered by Blumrosen, Nisan, and Segal, motivated by the fact that communication lower bounds sidestep challenging debates on what can be assumed about players' behavior [Nis00], [BNS07], [NS06]. A high-level overview of this early literature appears in Section I-B.

The current state-of-the-art is fairly remarkable: Without any restrictions on the valuations, a $\max\{1/n, 1/O(\sqrt{m})\}$-approximation to the optimal welfare can be achieved in $\text{poly}(n, m)$ communication [LS11]; and this is tight [NS06]. For *fractionally subadditive (XOS)* valuations (a strict subclass of subadditive),[2] a $(1 - (1 - 1/n)^n)$-approximation can be achieved in $\text{poly}(n, m)$ communication [Fei09]; and this is tight [DNS10]. For subadditive valuations, a $1/2$-approximation can be achieved in $\text{poly}(n, m)$ communication [Fei09]; and no better than a $(1/2 + 1/2n)$-approximation can be achieved in $\text{poly}(n, m)$ communication [DNS10], so this is tight as $n \to \infty$.

The two-player case: While Feige's $1/2$-approximation is tight as $n \to \infty$, the gap between $1/2$ and $3/4$ for the two-player case was posed as an open problem in [Fei09], [DNS10]. At first glance, it seems quite unusual for the $n = 2$ case to be singled out in this way when the asymptotics are resolved. Yet in our context, there is a substantial difference in the merits of a $1/2$-approximation for $n = 2$, and a $1/2$-approximation for $n > 2$: Feige's $1/2$-approximation for $n > 2$ employs an incredibly sophisticated LP rounding, but for $n = 2$ the same guarantee is achieved by multiple trivial algorithms.

For concreteness, consider the following simple $1/2$-

[1] A valuation function is subadditive if for all S, T $v(S \cup T) \leq v(S) + v(T)$.

[2] A valuation is fractionally subadditive if it can be written as a maximum over additive functions.

2575-8454/19/$31.00 ©2019 IEEE
DOI 10.1109/FOCS.2019.00025

approximation algorithms: (a) allocate all of the items in M to a uniformly random player; or (b) allocate each item independently to a uniformly random player; or (c) ask each player to report $v_i(M)$ and award M to the highest bidder. Note that methods (a) and (b) are particularly trivial in that they are completely oblivious to the valuations, and methods (a) and (c) are particularly trivial in that they maintain their guarantee even without subadditivity. All three can trivially be made into truthful auctions — methods (a) and (b) don't even solicit input and are therefore truthful, and (c) is simply a second-price auction on the grand bundle M.

By the above discussion, resolving the gap between $1/2$ and $1/2+1/2n = 3/4$ for $n = 2$ is not just a question of determining the optimal constant. The question is really whether or not it is possible to achieve *any non-trivial guarantees* for two subadditive valuations. As Feige notes, "*finding matching upper and lower bounds for this case may lead to new insights about subadditive functions.*"[3] The main result of this paper rules out non-trivial guarantees, while shedding new light on subadditivity and developing new randomized communication lower bounds — see Section I-A.

Main Result (Informal)**.** *For two subadditive players, a $1/2$-approximation of the maximum social welfare is optimal among all protocols with subexponential communication.*

Implications: Before overviewing our construction and discussing extensions of our main result beyond subadditivity, we wish to highlight immediate implications for combinatorial auctions with *strategic* bidders, via two recent reductions that underscore the significance of the $n = 2$ case.

(1) **The power of truthful vs. non-truthful communication-efficient protocols:** The central driving theme of algorithmic mechanism design is understanding the relative power of truthful vs. non-truthful "efficient" protocols [NR01]. For combinatorial auctions, when "efficient" refers to $\text{poly}(n, m)$-communication, no separation is known to exist — for any valuation class nor any number of players. This is despite significant gaps in the state-of-the-art approximation ratios, as described in Section I-B. Recent work of [Dob16b] provides a deep structural connection between truthful communication-efficient combinatorial auctions and their *simultaneous*, non-truthful communication-efficient counterparts, *specifically when* $n = 2$.[4] This immediately proposes extensive study of the $n = 2$ case to search for the first separation for truthful vs. non-truthful combinatorial auctions. On this front, our result proves that in fact no separation exists for subadditive players, as the aforementioned trivial protocols (now proved to be optimal) are also truthful.

(2) **Price of anarchy of simple mechanisms:** The central measure by which the performance of non-truthful combinatorial auctions is quantified in strategic settings is the *price of anarchy (PoA)*. In the Bayesian version (BPoA), this is defined as the worst-case ratio between the expected welfare of the worst equilibrium on one hand, and the optimal expected welfare on the other. The PoA and BPoA of various simple combinatorial auction formats have been studied extensively in recent years. For subadditive valuations, simultaneous first-price auctions are known to have BPoA at least $1/2$ [FFGL13]. The work of [CKST16] shows this is tight even for two players. Can auction formats *other than first-price* do better? Roughgarden [Rou14] shows that the key to simultaneously bounding BPoA over all auction formats is communication lower bounds. Applying his framework to our new lower bound reveals that there is no auction format with sub-doubly-exponentially many strategies and BPoA better than $1/2$. We conclude that simultaneous first-price auctions are optimal among this class of auctions, for all values of n.

A. Results and Intuition

Main Theorem. *Any (randomized) protocol that guarantees a $1/2 + 6/\log_2(m)$-approximation to the optimal welfare for two monotone subadditive bidders requires communication $\Omega(e^{\sqrt{m}})$.*

In this section we provide intuition for the main proof steps. The first step (Section III) is the construction of a new class of subadditive functions. Given previous work on fractionally-subadditive valuations, if our new class is to demonstrate a stronger hardness result than $3/4$, clearly it cannot be fractionally-subadditive. Feige [Fei09] shows how to construct such subadditive, non-fractionally-subadditive valuations from instances of set cover, but this is the only known general such construction in the literature.[5] Moreover, to move away from $3/4$, our new class must be "far" from fractionally-subadditive.

Our first contribution is to identify a key property that every $v(\cdot)$ in our class must have on top of being far from fractionally-subadditive in order to demonstrate hardness stronger than $1/2 + \varepsilon$: for all subsets T, $v(T) + v(\bar{T})$ must belong to the range $[v(M), (1 + 3\varepsilon)v(M)]$. In other words, $v(\cdot)$ must essentially appear *additive at the large scale*. In Section III we show that without this property for (say) player 1's valuation, there exists a simple protocol guaranteeing a $(1/2+\varepsilon)$-approximation (specifically, the best of allocating either T or $\bar{T}$ uniformly at random to player 1 and the rest to player 2, or allocating everything to player 2). The necessity of this property means we cannot use Feige's set cover construction as is. The main idea behind our new construction is then to hard-code $v(T) + v(\bar{T}) = v(M)$ for all T into the appropriate far-from-fractionally-subadditive instances, while carefully ensuring that subadditivity is not violated. This construction itself may be of independent

[3] See Conclusions Section in [Fei09].

[4] There are also implications when $n > 2$, but not quite as strong as for $n = 2$.

[5] While every valuation v for which $v(S) \in \{1, 2\}$ for any $S \neq \emptyset$ is subadditive, such functions trivially admit a $3/4$-approximation and so don't serve as a useful starting point.

interest due to the scarcity of subadditive classes that are far from fractionally-subadditive.

In Section IV, we show that the new construction is rich enough to encode the EQUALITY problem.[6] The property $v(T) + v(\bar{T}) = v(M)$ turns out to be extremely convenient as it immediately implies that $\text{OPT}(v_1, v_2) = v_1(M)$ whenever $v_1(\cdot) = v_2(\cdot)$. To complete the reduction from EQUALITY, we establish a doubly-exponentially-large subset $\mathcal{V}$ of valuations from our class such that for every nonequal $v(\cdot) \neq w(\cdot) \in \mathcal{V}$, it holds that: (a) $v(M) = \ell$; and (b) $\text{OPT}(v, w) \approx 2\ell$. By the convenient property, beating a $1/2$-approximation when both players have valuations from this subset is exactly deciding whether or not $v_1(\cdot) = v_2(\cdot)$.

Finally, we prove our full lower bound for randomized protocols in Section V. There, we first very briefly overview why our construction is unlikely to admit a reduction from canonical problems known to require exponential randomized communication (such as DISJOINTNESS or GAP-HAMMING-DISTANCE). Instead, we propose a new "near-EQUALITY" problem that we call EXIST-FAR-SETS, and directly prove that it requires exponential randomized communication via the information complexity approach of [BJKS04], [Bra12], [BGPW13]. To our knowledge, this is the first application of these modern communication complexity tools within Algorithmic Game Theory. The randomized hardness of EXIST-FAR-SETS may therefore be of independent interest.

Extensions beyond subadditive: Our main result concerns subadditive valuations. In Section VI, we consider the space between fractionally-subadditive and subadditive valuations, and between their respective $3/4$- and $1/2$-approximations. We apply a natural parameterization of this space called the Maximum-over-Positive-Hypergraphs (MPH) hierarchy [FFI+15]. Fractionally-subadditive valuations are equivalent to MPH-1 (level 1 in the hierarchy), and all subadditive functions lie in MPH-$m/2$ (level $m/2$ in the hierarchy). For any $2 \leq k \leq m/2$, we establish a protocol for welfare maximization as follows:

Theorem. *There exists a protocol with poly(m) communication that guarantees $1/2 + 1/O(\log k)$ of the optimal welfare for 2 bidders whose valuations are subadditive and MPH-k.*

Our protocol is an *oblivious* rounding of the well-known configuration linear program (LP).[7] The key property we use to establish our guarantee for MPH-k is a generalization of a result of [Fei09] for MPH-1 (see Lemma VI.4). Our results may be useful for obtaining a better understanding of the space between fractionally-subadditive and subadditive set functions in the future.[8]

We conclude by noting the following interesting corollary from combining the theorem on MPH-k with our main result: since the main result proves that a $(1/2 + 6/\log m)$-approximation is impossible with subexponential communication, and as all subadditive functions are MPH-$m/2$, our protocol and lower bound must be tight *even up to lower-order terms*.

B. Related Work

Communication complexity of combinatorial auctions: The works most related to ours concern the standard communication complexity of combinatorial auctions. The tables below summarize prior work for various valuation classes. While the $n = 2$ table is most relevant for the present paper, the general n table is included for reference. Note that no separate row is needed for hardness of truthful communication, because no such results are known (aside from general communication hardness).[9]

For context, it is worth noting that all referenced (truthful or not) communication protocols take one of two forms. The first is via solving a particular LP relaxation (called the configuration LP) and rounding the fractional optimum [FV10], [Fei09], [LS11]. The second is via mechanisms which randomly sample a fraction of bidders to gather statistics, then run a posted-price mechanism on the remaining bidders [Dob16a], [Dob07]. Both classes of mechanisms require bidders to communicate *demand queries*. That is, bidders are asked questions of the form: "For item prices $p_1, \ldots, p_m$, which set of items maximizes $v_i(S) - \sum_{j \in S} p_j$?" All of the aforementioned protocols/mechanisms make polynomially many demand queries, and have further polynomial-time overhead.

Recent work of [Dob16b] proves a surprising connection between two-player truthful combinatorial auctions, and two-player *simultaneous* (non-truthful) protocols. In particular, any separation between the approximation guarantees achievable by communication-efficient protocols and communication-efficient simultaneous protocols would constitute the first separatation between truthful and non-truthful communication-efficient combinatorial auctions. Such separations were already known for large n [ANRW15], [Ass17], but not for $n = 2$ (and therefore aren't relevant to Dobzinski's framework). As such, the $n = 2$ setting is now receiving extra attention, although the desired separation still

[6] In EQUALITY, Alice and Bob are each given a k-bit string as input, and are asked to decide whether their strings are equal or not. EQUALITY is known to require communication of k to solve deterministically, but to admit efficient randomized protocols.

[7] That is, while communication is indeed needed to optimally solve the configuration LP, no further communication is necessary in order to round the resulting solution. See [FFT16] for further discussion on the merits of oblivious versus non-oblivious rounding.

[8] For example, to claim lower bounds on the MPH levels of specific functions: since our lower bound construction does not admit a $(1/2 + 6/\log m)$-approximation in subexponential communication, it establishes the existence of a constant C such that it is provably not in MPH-Cm.

[9] Note that [DN15] consider "multi-unit" auctions, where $v(S) = v(T)$ whenever $|S| = |T|$. Such valuations can be described with m numbers, and therefore the VCG-auction is communication/computationally-efficient. They show, however, that a $(> 1/2)$-approximation can be achieved significantly faster (in poly$(\log(m))$ communication) without incentives, while a $1/2$-approximation is the best achievable in poly$(\log(m))$ communication by a truthful auction.

$n = 2$	Submodular	XOS	Subadditive	General
Communication hardness	$\frac{17}{18}$ [DV13]	$\frac{3}{4}$ [DNS10]	$\frac{3}{4}$ [DNS10]	$\frac{1}{2}$ [NS06]
Communication protocol	$\frac{13}{17}$ [FV10]	$\frac{3}{4}$ [Fei09]	$\frac{1}{2}$ [Trivial]	$\frac{1}{2}$ [Trivial]
Truthful comm. protocol	$\frac{1}{2}$ [Trivial]	$\frac{1}{2}$ [Trivial]	$\frac{1}{2}$ [Trivial]	$\frac{1}{2}$ [Trivial]

General n	Submodular	XOS	Subadditive	General
Comm. hardness	$1 - \frac{1}{2e}$ [DV13]	$1 - (1 - \frac{1}{n})^n$ [DNS10]	$\frac{1}{2} + \frac{1}{2n}$ [DNS10]	$\max\{\frac{1}{n}, \frac{1}{\Omega(\sqrt{m})}\}$ [NS06]
Comm. protocol	$1 - \frac{1}{e} + 10^{-5}$ [FV10]	$1 - (1 - \frac{1}{n})^n$ [Fei09]	$\frac{1}{2}$ [Fei09]	$\max\{\frac{1}{n}, \frac{1}{\sqrt{2m}}\}$ [LS11]
Truthful comm.	$\frac{1}{O(\sqrt{\log m})}$ [Dob16a]	$\frac{1}{O(\sqrt{\log m})}$ [Dob16a]	$\frac{1}{O(\log m \log\log m)}$ [Dob07]	$\max\{\frac{1}{n}, \frac{1}{\sqrt{2m}}\}$ [LS11]

remains elusive [BMW18]. In particular, note that no non-trivial approximation guarantees are known for two-player combinatorial auctions via truthful communication-efficient protocols, even when the valuations are submodular.

Related results on combinatorial auctions: As previously referenced, combinatorial auctions are studied via other complexity lenses as well. The most popular alternative is the value-queries model, or standard computational complexity. That is, each bidder is capable only of querying their valuation function on a given set (value query), or has access to the explicit (poly-sized) circuit which computes a value query. In both models, a tight $(1 - 1/e)$-approximation is known for submodular valuations [Von08], [MSV08], [DV12b], and a tight $\Theta(1/\sqrt{m})$-approximation is known for XOS and subadditive valuations [DNS10]. To reconcile these latter impossibility results with the above-referenced positive results, observe that it generally requires $\exp(m)$ value queries (or is NP-hard with explicit circuit access) to compute a demand query. Unlike the communication model, strong separations between guarantees of truthful and non-truthful mechanisms are known in these models [PSS08], [BDF$^+$10], [BSS10], [DSS15], [Dob11], [DV12a], [DV12b]. It is worth noting that some of these approaches also yield communication lower bounds for the restricted class of Maximal-in-Range/VCG-based protocols [BDF$^+$10], [DSS15], or for the "combinatorial public projects" problem [PSS08]. For further details of these results, see [DSS15, Table 1].

Communication complexity in AGT: Beyond combinatorial auctions, communication complexity has appeared in the study of matching markets [GNOR15], [ABKS17], fair division [BN19], voting theory [CS05], and most recently equilibrium computation [RW16], [BR17], [GR18], [BDN19]. The last sequence of works is notable for introducing the AGT community to recent developments in "lifting theorems." In a similar vein, the present paper applies recent developments in information complexity.

C. Summary

We study the communication complexity of welfare maximization in two-player combinatorial auctions. Our main result is that the trivial $1/2$-approximation is in fact optimal among all (possibly randomized) protocols with subexponential communication. Our key innovations are: (1) a new class of subadditive functions far from fractionally-subadditive with an "additive-at-the-large-scale" property, and (2) a new "near-EQUALITY" problem with high randomized communication complexity, EXIST-FAR-SETS (both of which may be of independent interest). In addition to resolving an open question of [Fei09], [DNS10], our results establish the following corollaries: (a) There is no gap between the approximation ratios achievable by truthful and not-necessarily-truthful mechanisms with $\text{poly}(m)$ communication for two subadditive players; (b) For any number of subadditive players, simultaneous first-price auctions achieve the optimal price of anarchy (i.e., $1/2$) among all auctions with sub-doubly-exponentially-many strategies. We additionally develop a $(1/2 + 1/O(\log k))$-approximation whenever both players are subadditive and MPH-k. Therefore, our lower bound is tight even up to lower order terms (i.e., $1/2 + 1/O(\log m)$ is achievable in $\text{poly}(m)$ communication, but no better).

II. PRELIMINARIES

We consider the following problem. There is a set M of m items. Alice and Bob each have a valuation function $A(\cdot)$ and $B(\cdot)$, respectively that takes as input subsets of M and outputs an element of $\mathbb{R}_+$. Moreover, $A(\cdot)$ and $B(\cdot)$ are both *monotone* ($v(X \cup Y) \geq v(X)$ for all X, Y) and *subadditive* ($v(X \cup Y) \leq v(X) + v(Y)$). Alice and Bob wish to communicate as little as possible about their valuation functions in order to find a welfare-maximizing allocation (that is, the X maximizing $A(X) + B(M \setminus X)$). Formally, we study the following decision problem – observe that this is a promise problem for which if the input does not satisfy the promise, any output is considered correct.

Definition II.1 (WELFARE-MAXIMIZATION(m, α))**.** WELFARE-MAXIMIZATION *is a communication problem between Alice and Bob:*

- ***Alice's Input:*** *$A(\cdot)$, a monotone subadditive function over $2^{[m]}$; and a target C.*
- ***Bob's Input:*** *$B(\cdot)$, a monotone subadditive function over $2^{[m]}$; and a target D.*
- ***Promise:*** *$C = D$. Also, there either exists an $S \subseteq [m]$ satisfying $A(S) + B(\overline{S}) \geq C$, or for all $S \subseteq [m]$, $A(S) + B(\overline{S}) < \alpha C$.*
- ***Output:*** *1 if $\exists S \subseteq [m]$, $A(S) + B(\overline{S}) \geq C$; 0 if $\forall S \subseteq [m]$, $A(S) + B(\overline{S}) < \alpha C$.*

We will sometimes drop the parameter m when it is irrelevant. We will also refer to any protocol solving WELFARE-MAXIMIZATION(m, α) *as an α-approximation for* WELFARE-MAXIMIZATION(m).

Also of interest is the corresponding search problem, which instead asks Alice and Bob to find an X maximizing $A(X) + B(\overline{X})$ (and an α-approximation is a protocol guaranteeing a Y satisfying $A(Y) + B(\overline{Y}) \geq \max_{X \subseteq M}\{A(X) + B(\overline{X})\}$). It is easy to see that any poly(m)-communication protocol for the search problem implies a poly(m)-communication protocol for the decision problem (with an extra round of communication). As such, we will prove all lower bounds against the decision problem (as they immediately imply to search as well), and develop all protocols for the search problem (as they immediately imply to decision as well).

III. MAIN CONSTRUCTION

In this section, we present our base construction. In subsequent sections, we show how to leverage this construction to derive our lower bounds. We begin by considering a collection of subsets $\mathcal{S} = \{S_1, \ldots, S_k\}$ where each $S_i \subseteq M$, and defining a useful property that $\mathcal{S}$ may possess. Throughout this section, let ℓ denote an even integer ≥ 4. Most proofs are deferred to Appendix A.

Definition III.1 (ℓ-sparse)**.** We say that $\mathcal{S}$ is *ℓ-sparse* if for all $T_1, \ldots, T_{\ell-1} \in \mathcal{S}$, $\cup_j T_j \neq M$.

That is, $\mathcal{S}$ is ℓ-sparse if there do *not* exist $\ell - 1$ elements of $\mathcal{S}$ such that their union is the entire ground set M. We now follow [Fei09], [BR11] in defining a class of valuation functions parameterized by a collection of sets. Specifically, let $\mathcal{S} = \{S_1, \ldots, S_k\}$ be an ℓ-sparse collection. For $X \subseteq M$, define

$$\sigma_{\mathcal{S}}(X) := \begin{cases} \min\left\{|Y| : Y \subseteq [k], X \subseteq \bigcup_{i \in Y} S_i\right\}, & \text{Case 1;} \\ \max\{\ell, k\}, & \text{O/W;} \end{cases}$$

where Case 1 occurs if X is covered by $\mathcal{S}$, that is, if $X \subseteq \bigcup_{i \in [k]} S_i$. So $\sigma_{\mathcal{S}}(X)$ is the smallest number of sets from $\mathcal{S}$ whose union contains X, or some large number $\max\{\ell, k\}$ if there are no such sets. Let us now revisit the barrier to using this construction as-is, before defining our new construction.

Lemma III.1. *Let $\varepsilon \leq 1/2$, and let there exist a set T and player i for which $v_i(T) + v_i(\bar{T}) \geq (1 + \varepsilon)v_i(M)$. Then at least one of the following three allocations achieves a $1/2 + \varepsilon/3$-approximation to* WELFARE-MAXIMIZATION(m)*: (a) give to player i the set T and to the other $\bar{T}$, (b) give to player i the set $\bar{T}$ and to the other player T, (c) give the other player all items. Therefore, a $1/2 + \varepsilon/3$-approximation to* WELFARE-MAXIMIZATION(m) *exists with poly(m) communication (simply check all three allocations and pick the best).*

Proof: Without loss of generality, let $i = 1$. Then the three allocations, respectively, achieve welfare $v_1(T) + v_2(\bar{T}), v_1(\bar{T}) + v_2(T)$, and $v_2(M)$. Observe further that choosing uniformly at random between the first two guarantees welfare:

$$\begin{aligned} & 1/2 \cdot \left(v_1(T) + v_2\left(\bar{T}\right) + v_1\left(\bar{T}\right) + v_2(T)\right) \\ \geq\ & 1/2\left((1+\varepsilon)\, v_1(M) + v_2(M)\right) \\ \geq\ & 1/2 \cdot \text{OPT}(v_1, v_2) + \varepsilon v_1(M). \end{aligned}$$

Therefore, if $v_1(M) \geq \text{OPT}(v_1, v_2)/3$, the best of the first two guarantees guarantees a $1/2 + \varepsilon/3$ approximation. On the other hand, if $v_1(M) \leq \text{OPT}(v_1, v_2)/3$, then we have:

$$\begin{aligned} & \text{OPT}(v_1, v_2) \leq v_1(M) + v_2(M) \\ \leq\ & \text{OPT}(v_1, v_2)/3 + v_2(M) \\ \implies\ & v_2(M) \geq 2/3 \cdot \text{OPT}(v_1, v_2). \end{aligned}$$

So in either case, the best of the three allocations guarantees a $1/2 + \varepsilon/3$ approximation. ∎

Lemma III.1 implies that any hard instance must appear additive at the large scale, but $\sigma_{\mathcal{S}}(\cdot)$ generally lacks this property (see Figure 1). Our new construction essentially hard-codes this property onto $\sigma_{\mathcal{S}}(\cdot)$. We now define our new valuation function $f^{\ell}_{\mathcal{S}}(\cdot)$:

(a) If $\sigma_{\mathcal{S}}(X) < \frac{\ell}{2}$, then define $f^{\ell}_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X)$ and $f^{\ell}_{\mathcal{S}}(\overline{X}) = \ell - \sigma_{\mathcal{S}}(X)$.
(b) For any X whose value is not defined in (a), $f^{\ell}_{\mathcal{S}}(X) = \frac{\ell}{2}$.

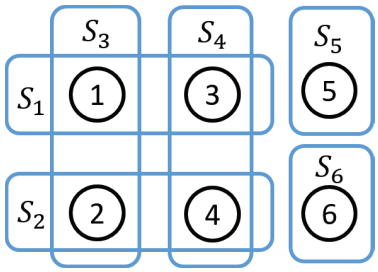

Figure 1: Example of ℓ-sparse $\mathcal{S}$. There are $m = 6$ items, $k = 6$ sets, and $\ell = 4$. The set $\mathcal{S} = \{S_1, \ldots, S_6\}$ is depicted by the rectangles. $\mathcal{S}$ is ℓ-sparse, as no collection of 3 sets in $\mathcal{S}$ covers $[m]$.

It is not immediately clear that $f^{\ell}_{\mathcal{S}}(\cdot)$ is well-defined; indeed, if $\sigma_{\mathcal{S}}(X)$ and $\sigma_{\mathcal{S}}(\overline{X})$ are both $< \frac{\ell}{2}$, then $f^{\ell}_{\mathcal{S}}(X)$ is doubly defined. Fortunately, this can never occur when $\mathcal{S}$ is ℓ-sparse.

Lemma III.2. *If $\mathcal{S}$ is ℓ-sparse, then $f^{\ell}_{\mathcal{S}}(\cdot)$ is well-defined.*

Now, we would like to prove that $f^{\ell}_{\mathcal{S}}$ is monotone and subadditive whenever $\mathcal{S}$ is ℓ-sparse (Corollary III.4). The following facts about $f^{\ell}_{\mathcal{S}}$ and $\sigma_{\mathcal{S}}$ highlight the key steps in the proof.

Lemma III.3. *Let $\mathcal{S}$ be ℓ-sparse. Then:*

(1) $\sigma_{\mathcal{S}}$ is monotone and subadditive.
(2) For all X, $f^{\ell}_{\mathcal{S}}(X) = \ell - f^{\ell}_{\mathcal{S}}(\overline{X})$.

(3) If $\sigma_{\mathcal{S}}(X) < \frac{\ell}{2}$ or $f^{\ell}_{\mathcal{S}}(X) < \frac{\ell}{2}$, then $f^{\ell}_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X)$.
(4) If $f^{\ell}_{\mathcal{S}}(X) > \frac{\ell}{2}$ then $f^{\ell}_{\mathcal{S}}(X) = \ell - \sigma_{\mathcal{S}}(\overline{X})$.
(5) For all X, $f^{\ell}_{\mathcal{S}}(X) \leq \sigma_{\mathcal{S}}(X)$.

Corollary III.4. *If $\mathcal{S}$ is ℓ-sparse, then $f^{\ell}_{\mathcal{S}}(\cdot)$ is monotone and subadditive.*

Functions of the form $f^{\ell}_{\mathcal{S}}(\cdot)$ will form the basis of our lower bound constructions, which we overview in the following sections. Figure 1 demonstrates one illustrative example to help parse the construction: In this example, it holds that $\sigma_{\mathcal{S}}(\{2\}) = 1 < \ell/2$, so $f^{\ell}_{\mathcal{S}}(\{2\}) = 1$. It follows that $f^{\ell}_{\mathcal{S}}(\{1,3,4,5,6\}) = \ell - \sigma_{\mathcal{S}}(\{2\}) = 3$. As another example, $\sigma_{\mathcal{S}}(\{1,4,5\}) = 3 \geq \ell/2$ and $\sigma_{\mathcal{S}}(\{2,3,6\}) = 3 \geq \ell/2$, so $f^{\ell}_{\mathcal{S}}(\{1,4,5\}) = f^{\ell}_{\mathcal{S}}(\{2,3,6\}) = \ell/2 = 2$. Finally, observe that $\sigma_{\mathcal{S}}(\{1,4,5\}) + \sigma_{\mathcal{S}}(\{2,3,6\}) = 6 > 4 = \sigma_{\mathcal{S}}(M)$, so trying to use $\sigma_{\mathcal{S}}(\cdot)$ directly would violate Lemma III.1, and the modification to $f^{\ell}_{\mathcal{S}}(\cdot)$ is necessary for any traction.

IV. Deterministic Protocols for Subadditive Valuations

The construction in Section III gets us most of the way towards our deterministic lower bound. The remaining step is a reduction from EQUALITY. To briefly remind the reader, Alice receives input $a \in \{0,1\}^k$, and Bob receives input $b \in \{0,1\}^k$. Their goal is to output yes if $a_i = b_i$ for all $i \in [k]$, and no otherwise. It is well-known (see, e.g., [KN97]) that any deterministic protocol for EQUALITY requires communication $\geq k$.

Theorem IV.1. *For any even integer $\ell \in [4, \log_2(m)]$, any deterministic communication protocol that guarantees a $(1/2+1/\ell)$-approximation to* WELFARE-MAXIMIZATION(m) *requires communication $\exp(m/\ell \cdot 2^{\ell})$. In particular, a guarantee of $1/2 + \varepsilon$ requires communication $e^{\varepsilon m/2^{1/\varepsilon}} = e^{\Omega(m)}$, and a guarantee of $1/2 + 2/\log(m)$ requires communication $e^{m^{\Omega(1)}}$.*

Before proceeding with our construction, we'll need one more property of collections of sets:

Definition IV.1 ([KS73])**.** A collection $\mathcal{S} = \{S_1, \ldots, S_k\}$ is *ℓ-independent* if $\{T_1, \ldots, T_k\}$ is ℓ-sparse whenever $T_i \in \{S_i, \overline{S_i}\}$.

In other words, $\mathcal{S}$ is ℓ-independent if we can choose either S_i or $\overline{S_i}$ independently, for each i, and form an ℓ-sparse collection no matter our choices. For example, while $\mathcal{S}$ in Figure 1 is ℓ-sparse for $\ell = 4$, it is not ℓ-independent, since, e.g., $\bar{S}_5 \cup \bar{S}_6$ covers $[m]$. We now proceed with our reduction, which relies on the existence of large ℓ-independent collections (such collections are known to exist; at the end of this section we give a precise statement and a proof appears in Appendix A for completeness).

Proposition IV.2. *Let $\mathcal{S}$ be an ℓ-independent collection with $|\mathcal{S}| = k$. Then any deterministic communication protocol that guarantees a $(1/2 + 1/2\ell-3)$-approximation to the optimal welfare for two monotone subadditive bidders requires communication at least k.*

Proof: Let $\mathcal{S} = \{S_1, \ldots, S_k\}$ be ℓ-independent. For each i, define $S^1_i := S_i$, and $S^0_i := \overline{S_i}$. Now, consider an instance of EQUALITY where Alice is given a and Bob is given b. Alice will create the valuation function $f^{\ell}_{\mathcal{A}}$, where $\mathcal{A} := \{S^{a_1}_1, \ldots, S^{a_k}_k\}$ (i.e. Alice builds $\mathcal{A}$ by taking either S^1_i or S^0_i, depending on a_i). Bob will create the valuation function $f^{\ell}_{\mathcal{B}}$, where $\mathcal{B} := \{S^{b_1}_1, \ldots, S^{b_k}_k\}$. Observe first that $f^{\ell}_{\mathcal{A}}(\cdot)$ and $f^{\ell}_{\mathcal{B}}(\cdot)$ are indeed well-defined, monotone, and subadditive as $\mathcal{S}$ is ℓ-independent (and therefore $\mathcal{A}$ and $\mathcal{B}$ are both ℓ-sparse).

Observe that if $a = b$, then $\mathcal{A} = \mathcal{B}$ and moreover $f^{\ell}_{\mathcal{A}}(\cdot) = f^{\ell}_{\mathcal{B}}(\cdot)$. So immediately by part (2) of Lemma III.3, the maximum possible total welfare is ℓ (indeed, any partition of the items gives welfare ℓ). On the other hand, if there exists an i such that $a_i \neq b_i$ (without loss of generality say that $a_i = 1$ and $b_i = 0$), we claim that welfare $2\ell - 2$ is achievable. To see this, consider the allocation which awards $\overline{S_i}$ to Alice and S_i to Bob. Indeed, $f^{\ell}_{\mathcal{A}}(S_i) = 1$ (as $S_i \in \mathcal{A}$), so $f^{\ell}_{\mathcal{A}}(\overline{S_i}) = \ell - 1$. Similarly, $f_B(\overline{S_i}) = 1$, so $f_B(S_i) = \ell - 1$, achieving total welfare $2(\ell - 1)$.[10]

So assume for contradiction that a deterministic $\frac{1}{2} + \frac{1}{2\ell-3} > \frac{\ell}{2\ell-2}$-approximation exists to the optimal welfare for 2 monotone subadditive bidders with communication $< k$. Then such a protocol would solve EQUALITY with communication $< k$ by the reduction above, a contradiction. ∎

Finally, in the next lemma we show how large k can be while guaranteeing an ℓ-independent collection of size k to exist. This suffices to complete the proof of Theorem IV.1. The lemma is based on a known existential construction using the probabilistic method, which we repeat for the sake of completeness in Appendix A (explicit constructions of comparable guarantees also exist [Alo86]).

Lemma IV.3. *For all m, $x > 1$, and $\ell = \log_2(m) - \log_2(x)$, there exists a ℓ-independent collection of subsets of $[m]$ of size $k = e^{x/\ell}$.*

Proof of Theorem IV.1: Combine Proposition IV.2 and Lemma IV.3, and the observation that $2\ell - 3 \geq \ell$ whenever $\ell \geq 4$. The "in particular" parts of the statement follow first by taking $\ell = 1/\varepsilon$ (implying $\ell = \log_2(m) - \log_2(m/2^{1/\varepsilon})$ and $k = e^{\varepsilon m/2^{1/\varepsilon}}$, and then by taking $\ell = \log_2(m)/2$ (implying $\ell = \log_2(m) - \log_2(\sqrt{m})$) and $k = e^{2\sqrt{m}/\log(m)}$. ∎

[10] As an aside, note that welfare exceeding $2\ell - 2$ is not possible, as Alice and Bob each value all non-empty sets at least at 1, and therefore value all strict subsets of M at most at $\ell - 1$ by (2). Therefore, the only way Alice or Bob could have value exceeding $\ell - 1$ is to get all of M, meaning that the other player receives value 0.

V. Randomized Protocols for Subadditive Valuations

The construction in Section IV carries the high-level intuition for our randomized lower bound. However, we clearly cannot reduce from EQUALITY and get a randomized lower bound, as EQUALITY admits randomized communication-efficient protocols. As such, we will instead directly show that a certain "near-EQUALITY" problem requires exponential randomized communication. Our proof uses the information complexity approach popularized in [BJKS04], [Bra12], [BGPW13]. In order to introduce these tools to the AGT community, we will provide a complete exposition starting from the basics.

Let's first be clear about what a randomized protocol looks like. Alice and Bob have access to a public infinite string of perfectly random bits, r. All messages sent by (e.g.) Alice may therefore depend on her input, any messages sent by Bob, and r. At the end of the protocol, Alice and Bob will guess yes or no, and the answer should be correct with probability $2/3$.[11] The protocol is only "charged" communication for actual messages sent, and not for randomness used. The main result of this section is as follows:

Theorem V.1 (Randomized hardness)**.** *Any randomized protocol that guarantees a* $(1/2 + 6/\log(m))$*-approximation to* WELFARE-MAXIMIZATION(m) *requires communication complexity* $\Omega(e^{\sqrt{m}})$.

Let's now understand the issue with our previous construction. Aside from the fact that the previous proof clearly does not extend to a randomized lower bound, the construction itself admits a good randomized algorithm. Specifically, let $\mathcal{S}$ be some ℓ-independent set, and let exactly one of $\{S, \overline{S}\}$ be in $\mathcal{A}$ for all $S \in \mathcal{S}$ (and also let exactly one of $\{S, \overline{S}\}$ be in $\mathcal{B}$). Let Alice have valuation $f^\ell_{\mathcal{A}}(\cdot)$ and Bob have valuation $f^\ell_{\mathcal{B}}(\cdot)$. The problem is that Alice and Bob are still just trying to determine whether or not $\mathcal{A} = \mathcal{B}$ (that is, if $\mathcal{A} = \mathcal{B}$, then the optimal welfare is at most ℓ. If not, then the optimal welfare is $2\ell - 2$). Since $\mathcal{A}$ and $\mathcal{B}$ are both subsets of 2^M, the randomized algorithm for EQUALITY for inputs of size 2^m works.

The natural idea to try next is to reduce from a problem like DISJOINTNESS instead (for which randomized protocols indeed require exponential communication). Let's see one natural attempt from our previous construction and why it fails (just for intuition, we will not exhaustively repeat this for all possible reductions). Again let $\mathcal{S}$ denote an ℓ-independent collection, and again consider any instance (A, B) of DISJOINTNESS of size k (recall that A and B are bitstrings of length k and DISJOINTNESS asks to decide whether or not there exists an index i with $A_i = B_i = 1$). A first attempt at a reduction might be to let $\mathcal{A}$ contain S_i for all i such that $A_i = 1$, and $\mathcal{B}$ contain all $\overline{S_i}$ such that $B_i = 1$ (but $\mathcal{A}$ will never contain $\overline{S_i}$, and $\mathcal{B}$ will never contain S_i). Indeed, with this construction if there exists any index i with $A_i = B_i = 1$, the optimal welfare will be $2\ell - 2$ (give Alice $\overline{S_i}$, and Bob S_i). Unfortunately, even if there does not exist an index for which $A_i = B_i = 1$, the welfare can still be $\ell - 1 + \ell/2$. To see this, consider any index i for which $A_i = 1$. Then $f^\ell_{\mathcal{A}}(\overline{S_i}) = \ell - 1$. Moreover, as $S_i \notin \mathcal{B}$, and $\mathcal{S}$ is ℓ-independent, $f^\ell_{\mathcal{B}}(S_i) = \ell/2$ (because $\mathcal{S}$ is ℓ-independent, neither S_i nor $\overline{S_i}$ can be covered with fewer than $\ell - 1$ of the other sets in $\mathcal{B}$. As such, $f^\ell_{\mathcal{B}}(S_i) = f^\ell_{\mathcal{B}}(\overline{S_i}) = \ell/2$). So while in the "yes" case, the welfare is indeed $2\ell - 2$ just like our previous reduction, the welfare in the "no" case will be $3\ell/2 - 1$ as opposed to ℓ, proving only that a $3/4$-approximation requires exponential randomized communication (which is already known). Of course this is not a formal claim that no reduction from DISJOINTNESS is possible, but provides some intuition for why searching for one (or from GAP-HAMMING-DISTANCE, etc.) is likely not the right approach.

The issue is that our construction is getting much of its mileage from the fact that $v(S) + v(\overline{S}) = \ell$ for all S, and reducing from any problem except EQUALITY fails to make use of this. So the plan for our new construction is to observe that if $v(\cdot)$ and $w(\cdot)$ are *almost* the same (in a precise sense defined shortly), then we can still claim that $v(S)+w(\overline{S}) \approx \ell$ for all S.

The main idea of our construction is as follows: consider still an ℓ-independent set $\mathcal{S}$. For each $S_i \in \mathcal{S}$, rather than adding either S_i or $\overline{S_i}$ to $\mathcal{A}$, we will add either $S_i \cup \{j\}$ or $\overline{S_i} \cup \{j\}$, where j is a uniformly random element of M (and ditto for $\mathcal{B}$). Adding this random element to each set barely changes the welfare, but makes it significantly harder for Alice and Bob to figure out whether their valuations are nearly identical or not. We now proceed with the construction, followed by a complete proof.

Definition V.1. Two ordered collections of subsets $\mathcal{X} = \langle X_1, \ldots, X_k \rangle; \mathcal{Y} = \langle Y_1, \ldots, Y_k \rangle$ of M are *ℓ-compatible* if

(1) $|X_i| = |Y_i| = \frac{m}{2} + 1$ for all i.
(2) Either $|X_i \triangle Y_i| = 2$ or $|X_i \cap Y_i| = 2$ for all i.[12]
(3) $X_1, \ldots, X_k$ are ℓ-sparse, as are $Y_1, \ldots, Y_k$.
(4) For any subset $S \subseteq M$ of size less than $\frac{\ell}{2}$, at least one of $X_1, \ldots, X_k$ contains S, as does at least one of $Y_1, \ldots, Y_k$.

The main idea is as follows: for any ℓ-compatible $\mathcal{X}, \mathcal{Y}$, consider the valuation functions $f^\ell_{\mathcal{X}}(\cdot)$ and $f^\ell_{\mathcal{Y}}(\cdot)$. If for some i, $|X_i \cap Y_i| = 2$, this roughly corresponds to the "not equal" case in the previous construction, and welfare near 2ℓ is achievable. If instead, for all i $|X_i \triangle Y_i| = 2$, this roughly corresponds to the "equal" case in the previous construction, and welfare near ℓ is the best achievable. We

[11] As usual, the bound of $2/3$ is arbitrary, as any protocol with success probability $\geq 1/2 + 1/\text{poly}(m)$ can be repeated independently $\text{poly}(m)$ times to achieve a protocol with success probability $2/3$ (and then further repeated to achieve success $1 - 1/\exp(m)$).

[12] $\triangle$ represents symmetric difference, $|X_i \cup Y_i| - |X_i \cap Y_i|$.

first state this formally, and then follow with a proof that randomized protocols require exponential communication to distinguish these two cases. The proof of Lemma V.2 appears in Appendix A.

Lemma V.2. *Let $\mathcal{X}, \mathcal{Y}$ be ℓ-compatible. If for some i, $|X_i \cap Y_i| = 2$, then welfare $2(\ell-1)$ is achievable between $f^\ell_{\mathcal{X}}(\cdot)$ and $f^\ell_{\mathcal{Y}}(\cdot)$. Otherwise, the maximum achievable welfare is at most $\ell+1$.*

A. FAR-SETS and EXIST-FAR-SETS

Towards proving our lower bound, we'll define the following two problems, which may themselves be of independent interest, at least within the AGT community, as a "near-EQUALITY" problem which requires exponential randomized communication. Below, note that both FAR-SETS and EXIST-FAR-SETS are promise problems: if the input doesn't satisfy any of the stated conditions, arbitrary output is considered correct.

Definition V.2 (FAR-SETS(m))**.** FAR-SETS is a communication problem between Alice and Bob:

- **Alice's Input:** $X \subseteq M$, with $|X| = m/2+1$.
- **Bob's Input:** $Y \subseteq M$, with $|Y| = m/2+1$.
- **Promise:** Either $|X \triangle Y| = 2$ or $|X \cap Y| = 2$.
- **Output:** 0 if $|X \triangle Y| = 2$; 1 if $|X \cap Y| = 2$.

Definition V.3 (EXIST-FAR-SETS(m, k, ℓ))**.** EXIST-FAR-SETS is a communication problem between Alice and Bob:

- **Alice's Input:** $\mathcal{X} = \langle X_1, \ldots, X_k \rangle$. Each $X_i \subseteq M$.
- **Bob's Input:** $\mathcal{Y} = \langle Y_1, \ldots, Y_k \rangle$. Each $Y_i \subseteq M$.
- **Promise:** $\mathcal{X}$ and $\mathcal{Y}$ are ℓ-compatible.
- **Output:** $\bigvee_{i \in [k]}$ FAR-SETS(X_i, Y_i). Observe that if the EXIST-FAR-SETS promise above is satisfied, then by definition the FAR-SETS promise is satisfied for all (X_i, Y_i) (but not necessarily vice versa — the EXIST-FAR-SETS promise is strictly stronger due to ℓ-sparsity).

Observe that EXIST-FAR-SETS is *exactly* the problem referenced in Lemma V.2 (we state this formally below). Therefore, the goal of this section is to lower bound the randomized communication complexity of EXIST-FAR-SETS.

Corollary V.3. *Let $C(m, k, \ell)$ be such that every randomized communication protocol which solves any given instance of* EXIST-FAR-SETS*(m, k, ℓ) with probability at least $2/3$ has communication complexity at least $C(m, k, \ell)$. Then every randomized communication protocol which solves any given instance of* WELFARE-MAXIMIZATION$(m, \frac{1}{2} + \frac{1}{\ell-1})$ *with probability at least $2/3$ has randomized communication complexity at least $C(m, k, \ell)$.*

Our plan of attack follows Braverman's lower bound on the randomized communication complexity of DISJOINTNESS [Bra12]. The end result is Theorem V.18, which states that EXIST-FAR-SETS requires communication exponential in k.

B. Information Theory Preliminaries

Here, we provide some basic facts about information theory and information complexity. These are the standard preliminaries one would find in a paper on information complexity (e.g. [BGPW13]). Below, when we refer to a distribution μ, we use $\mu(\omega)$ where ω is in the support of μ to denote the probability of ω according to μ. All logarithms taken in this section are base-2. Also for this section, all distributions and random variables are supported on a finite set Ω. If $\mu(\omega) = 0$ for some $\omega \in \Omega$, we let $0 \cdot \log_2(1/0) := \lim_{x \to 0} x \cdot \log_2(1/x) = 0$.

Definition V.1 (Entropy)**.** *Let μ be a probability distribution over a finite set Ω. The (Shannon) entropy of μ, denoted by $H(\mu)$, is defined as $H(\mu) := \sum_{\omega \in \Omega} \mu(\omega) \log(\frac{1}{\mu(\omega)})$. If A is a random variable distributed according to μ_A, we also write $H(A) := H(\mu_A)$.*

Definition V.2 (Conditional Entropy)**.** *Let A and B be two random variables supported on a finite set Ω. Then the conditional entropy of A, conditioned on B is $H(A|B) := \sum_{b \in \Omega} \Pr[B = b] \cdot H(A|B = b)$.*[13]

Observe that as $H(\cdot)$ is a strictly concave function, $H(A|B) \leq H(A)$ for all B (with equality iff A and B are independent).

Fact V.4. *$H(A, B) = H(A) + H(B|A)$. Here, $H(A, B)$ denotes the entropy of the random variable (A, B).*

Fact V.4 above intuitively says that the entropy of a tuple of random variables is equal to the entropy of the first, plus the entropy of the second conditioned on the first. Note that if A and B are independent, then the joint entropy $H(A, B) = H(A) + H(B)$.

Definition V.3 (Mutual Information)**.** *For two random variables A, B, the mutual information between A and B, denoted by $I(A; B)$ is: $I(A; B) := H(A) - H(A|B) = H(B) - H(B|A)$.*

Definition V.4 (Conditional Mutual Information)**.** *For three random variables A, B, C, the mutual information between A and B, conditioned on C is denoted by $I(A; B|C)$, and $I(A; B|C) := \sum_{c \in \Omega} \Pr[C = c] \cdot I(A|C = c; B|C = c)$.*

Fact V.5 (Chain Rule)**.** *Let A, B, C, D be random variables. Then $I(A, B; C|D) = I(A; C|D) + I(B; C|A, D)$.*

Fact V.5 above intuitively says that the information learned about (A, B) from C (conditioned on D) can be broken down into the information learned about A from C (conditioned on D), plus the information learned about B from C (now conditioned on A in addition to D).

Definition V.5 (KL Divergence)**.** *We denote by $\mathbb{D}(A||B)$ the Kullback-Leibler Divergence between A and B, which is defined as $\mathbb{D}(A||B) := \sum_{\omega \in \Omega} \Pr[A = \omega] \cdot \log \frac{\Pr[A=\omega]}{\Pr[B=\omega]}$.*

[13] To be clear, by $H(A|B = b)$ we mean the entropy of the random variable A, when drawn conditioned on the event $B = b$.

Fact V.6. *For any random variables A, B, C, $I(A; B|C) = \mathbb{E}_{b,c}[\mathbb{D}(A_{bc}||A_c)]$. Here, A_{bc} denotes the random variable A conditioned on $B = b, C = c$, and A_c denotes the random variable A conditioned on $C = c$.*

Fact V.7 (Pinsker's Inequality)**.** *For any pair of random variables A, B of finite support, $||P - Q||_1 \leq \sqrt{2\mathbb{D}(P||Q)}$. Here, $||P - Q||_1 := \sum_{\omega \in \Omega} |\Pr[P = \omega] - \Pr[Q = \omega]|$.*

Definition V.6 (Information Complexity)**.** *The (internal) Information Complexity of a communication protocol π with respect to a distribution μ over pairs (X, Y) of inputs is defined as follows. Let $\Pi(X, Y)$ denote the random variable which is the transcript produced when Alice and Bob participate in protocol π with inputs (X, Y), when (X, Y) are drawn from μ. Then $IC_\mu(\pi) := I(\Pi(X, Y); Y|X) + I(\Pi(X, Y); X|Y)$.*

Above, the "transcript" refers to all communication between Alice and Bob (including the order bits were sent, who sent them, etc., *and including any public randomness*) when participating in protocol π. In particular, it is always possible to glean the output produced by π from viewing the transcript (but possibly additional information). Informally, the Information Complexity captures the amount of information Alice learns about Bob's input from participating in π (given that she already knows her own input, the public randomness, and that their joint input is drawn from μ), plus what Bob learns about Alice's input. Intuitively, it should be impossible for a protocol to convey C bits of information without C bits of communication. Indeed, this is the case:

Lemma V.8 ([BR14])**.** *For any distribution μ and protocol π, $IC_\mu(\pi) \leq CC(\pi)$ (where $CC(\pi)$ denotes the worst-case number of bits communicated during protocol π on any input).*

We conclude with a few more basic facts about communication protocols. Lemma V.9 below captures one key difference between communication protocols and algorithms with access to the entire input. Lemma V.9 below refers to *private randomness*, which are random bits which are known only to Alice (but not Bob) and vice versa. Such bits are also not counted towards the communication cost of the protocol (unless Alice wishes to share her private randomness with Bob).

Lemma V.9. *Let $P_\pi(\cdot, \cdot, \cdot)$ be a function where $P_\pi(z, X, Y)$ denotes the probability that Alice with input X and Bob with input Y produce transcript z when participating in a protocol π* over the randomness of any private randomness used *(as public randomness is already accounted for in the transcript). Then there exist functions $Q_\pi(\cdot, \cdot)$ and $R_\pi(\cdot, \cdot)$ such that $P_\pi(z, X, Y) = Q_\pi(z, X) \cdot R_\pi(z, Y)$.*

The proof of Lemma V.9 is straightforward. Essentially, $Q_\pi(z, X)$ is the probability that Alice *doesn't* deviate from transcript z with input X, conditioned on Bob communicating according to transcript z so far. Similarly, $R_\pi(z, Y)$ denotes the probability that Bob doesn't deviate from transcript z with input Y, conditioned on Alice communicating according to transcript z so far. These probabilities are well-defined because Alice must choose her future messages based only on the transcript so far (including the public randomness) and her input X (and Bob's must be only on the messages so far and Y), as well as her private randomness. Once confirming that these probabilities are well-defined, it is easy to see that indeed $P_\pi(z, X, Y) = Q_\pi(z, X) \cdot R_\pi(z, Y)$.

Finally, the lemma below states that lower bounds on the information complexity of any protocol that *only* uses private randomness also lower bound the information complexity of any protocol which uses public randomness. This initially may seem counterintuitive, since the opposite is true for communication. Both simple claims below have "approximate" versions in the other directions (discussed in the cited references), but we only use the easy directions.

Lemma V.10 (Folklore, see [New91], [BG14], [BBK+16])**.** *Let π be a protocol, and μ be a distribution over inputs. Then:*

- *If π uses private randomness, there exists a protocol π' using public randomness with exactly the same output as π, and $CC(\pi') \leq CC(\pi)$. But maybe $IC_\mu(\pi') \gg IC_\mu(\pi)$.*
- *If π uses public randomness, there exists a protocol π' using private randomness with exactly the same output as π, and $IC_\mu(\pi') \leq IC_\mu(\pi)$. But maybe $CC(\pi') \gg CC(\pi)$.*

Both claims in Lemma V.10 follow by simple reductions. For the first bullet, simply use all odd bits of the public randomness string as private randomness for Alice, and all even bits of the public randomness string as private randomness for Bob. Then the output of the protocol is identical, and the communication has not changed. However, Bob now *knows* Alice's private randomness, so the protocol may reveal significantly more information than previously (one example to have in mind is that perhaps the protocol has Alice output one uniformly random bit of her input. With private randomness, Bob learns very little about Alice's input upon seeing the bit. With public randomness, Bob learns exactly one bit of Alice's input). For the second bullet, simply use Alice's private randomness as the public randomness. That is, whenever the protocol requests random bits, Alice outputs these bits from her private random string. These bits are completely independent of her input, and therefore reveal no additional information. However, the communication might become enormous, as the randomness is now being directly communicated, and counts towards the communication cost. To use Lemma V.10, our lower bounds proceed by first lower bounding the information complexity of any protocol for EXIST-FAR-SETS with private randomness, using the second bullet of Lemma V.10 to lower bound the information complexity of any protocol for EXIST-

FAR-SETS with public randomness,[14] using Lemma V.8 to lower bound the communication complexity of EXIST-FAR-SETS with public randomness (if desired, we could then use the first bullet of Lemma V.10 to further lower bound the communication complexity of EXIST-FAR-SETS with private randomness). The point is just that exponential communication is required with either public or private randomness.

C. From EXIST-FAR-SETS to FAR-SETS

In this section, we show how to lower bound the randomized communication complexity of EXIST-FAR-SETS, provided we have *any* lower bound on the information complexity of (certain instances of) FAR-SETS. This section follows a similar path as Section 7.2 of [Bra12].[15] To get started, we need some additional notation for promise problems.

Definition V.7 (Promise problem)**.** *Let f be some function mapping $\{0,1\}^m \times \{0,1\}^m \to \{0,1\}$, and let $P \subseteq \{0,1\}^m \times \{0,1\}^m$. Then the communication problem* solving f under promise P *refers to the communication problem which requires Alice with input A and Bob with input B to output $f(A,B)$ whenever $(A,B) \in P$ (and they may provide arbitrary output otherwise).*

Definition V.8 (b-compatible inputs)**.** *Say that an input (X,Y) is b-compatible with respect to f and P if $(X,Y) \in P$ and $f(X,Y) = b$.*

Definition V.9 ((k,z)-safe distributions)**.** *Say that a distribution μ over $\{0,1\}^m \times \{0,1\}^m$ together with a promise $P \subseteq \{0,1\}^m \times \{0,1\}^m$ are (k,z)-*safe with respect to promise $P^* \subseteq (\{0,1\}^m \times \{0,1\}^m)^k$ *if for all $j \in [k]$, and all $(X_j,Y_j) \in P$, the probability that $\langle (X_1,Y_1),\ldots,(X_k,Y_k)\rangle \in P^*$ is at least z, when (X_i,Y_i) are drawn i.i.d. from μ for all $i \neq j$.*

Below is the main tool from information complexity that we'll use, which provides a method for proving communication lower bounds for k-dimensional communication problems via information complexity lower bounds on a related 1-dimensional problem (plus some technical assumptions to handle the promises).

Theorem V.11 (Follows from [Bra12])**.** *Let f be some function mapping $\{0,1\}^m \times \{0,1\}^m \to \{0,1\}$, and let $P \subseteq \{0,1\}^m \times \{0,1\}^m$ be some promised set of inputs. Let F^* be defined so that*

$$F^*((X_1,\ldots,X_k),(Y_1,\ldots,Y_k)) = \bigvee_{i \in k} f(X_i,Y_i).$$

Let also $P^* \subseteq (\{0,1\}^m \times \{0,1\}^m)^k$ be some promised set of inputs such that $P^* \subseteq P^k$ (that is, every coordinate of an element of P^* is in P).

Let also μ be any distribution over inputs that are 0-compatible with respect to f and P. Let also μ and P be (k,z)-safe with respect to P^*.

Then, if any protocol π that solves f under promise P with probability at least q has $IC_\mu(\pi) \geq c$, any randomized protocol π^* that solves F^* under promise P^* with probability at least q/z must have $CC(\pi^*) \geq kc$.

Proof: We use the same reduction proposed in Section 7.2 of [Bra12] for DISJOINTNESS. The analysis requires a little extra work to accommodate the promises. Assume for contradiction that there exists a randomized protocol π^* that solves F^* under promise P^* with probability at least q/z and $CC(\pi^*) < kc$. We prove the contrapositive of the theorem statement by providing a randomized protocol π that solves f under promise P with probability at least q, with $IC_\mu(\pi) < c$. Here is the reduction (recall that π^* is the assumed protocol for F^* under promise P^*, and we will use π to refer to the designed protocol):

- Alice and Bob are given input $(X,Y) \in P$.
- Alice and Bob use the shared randomness to select an $i \in [k]$ uniformly at random. Let $X_i := X$ and $Y_i := Y$.
- Alice and Bob use shared randomness to publicly sample $X_1,\ldots,X_{i-1}$ i.i.d. from μ_X, and $Y_{i+1},\ldots,Y_k$ i.i.d. from μ_Y. Here, μ_X denotes the marginal of μ restricted to X (ditto for μ_Y).
- Alice privately samples X_j from $\mu_X|Y_j$ for all $j > i$. That is, Alice samples X_j from μ_X, conditioned on Y_j. Bob privately samples Y_j from $\mu_Y|X_j$ for all $j < i$.
- Alice and Bob run protocol π^* on input $(X_1,\ldots,X_k;Y_1,\ldots,Y_k)$ (refer to this as $(\mathcal{X},\mathcal{Y})$ for notational simplicity) and output the answer.

Let's first observe the following about the correctness of the above protocol: First, maybe $(\mathcal{X},\mathcal{Y}) \notin P^*$. In this case, we have no guarantees about the success of the protocol (because π^* may behave arbitrarily). However, because (μ,P) are (k,z)-safe with respect to P^*, we know that for all (X,Y), the resulting $(\mathcal{X},\mathcal{Y})$ is in P^* with probability at least z.

Moreover, observe that, conditioned on $(\mathcal{X},\mathcal{Y}) \in P^*$, protocol π^* is correct with probability at least q/z by definition. Finally, we observe that because μ is 0-compatible with respect to f and P, $f(X_j,Y_j) = 0$ for all $j \neq i$. Therefore, $f(X,Y) = \bigvee_{j=1}^k f(X_j,Y_j) = F^*(\mathcal{X},\mathcal{Y})$. So whenever protocol π^* is correct on $(\mathcal{X},\mathcal{Y})$, our protocol π is correct on (X,Y). **At this point, we may conclude that our protocol is correct with probability at least $z \cdot q/z = q$.** The remaining step is to compute its information complexity with respect to μ.

First, we wish to point out that indeed the *communication complexity* of π can be quite high. Indeed, it is blowing up what should be a one-dimensional problem into a k-

[14] Indeed, the second bullet of Lemma V.10 states that if there exists a protocol for EXIST-FAR-SETS with public randomness and low IC then there exists such a protocol with private randomness, contradicting the first lower bound in the chain.

[15] Also, because of the additional promise of EXIST-FAR-SETS, we are unaware of prior work which could be cited as a black-box.

dimensional problem before solving it. Intuitvely, though, we'd like to say that only a $1/k$-fraction of this communication is actually being used to solve our original instance. It's hard to make this a formal statement, but this intuition can be made formal if we look at information complexity instead: In expectation, only a $1/k$-fraction of the *information our protocol learns* is relevant to the original instance.

Below, recall that $I(P;Q)$ denotes the mutual information of random variables P and Q and $I(P;Q \mid R)$ denotes the expected mutual information of P and Q conditioned on R. We will use $\Pi(X,Y)$ to denote the random variable corresponding to the transcript of the protocol π, when (X,Y) are drawn from μ, and will use $\Pi^*(\mathcal{X},\mathcal{Y})$ to denote the random variable corresponding to the transcript of π^* when inputs $(\mathcal{X},\mathcal{Y})$ are drawn from μ^k. Let's begin by writing the information that Alice learns about Y given X from our protocol π: $I(\Pi(X,Y);Y \mid X) = I(\Pi^*(\mathcal{X},\mathcal{Y});Y \mid i, X_1, \ldots, X_{i-1}, X, X_{i+1}, \ldots, X_k, Y_{i+1}, \ldots, Y_k)$.

Let's parse the above statement before proceeding. It is essentially saying that what Alice learns about Y from $\Pi(X,Y)$ given that she already knows X is exactly the same as what she learns from the random variable $\Pi(\mathcal{X},\mathcal{Y})$ (because these random variables are identical), except now Alice already knows $\mathcal{X}$ and $Y_{i+1}, \ldots, Y_k$ (because they were sampled publicly). Now we can perform some manipulations based on the facts from Section V-B. For the first step, we'll separate out the conditioning on i, as i is independent of all other random variables (we'll also start replacing $X_1, \ldots, X_{i-1}, X, X_{i+1}, \ldots, X_k$ with $\mathcal{X}$ to save space):

$$\begin{aligned} & I(\Pi^*(\mathcal{X},\mathcal{Y});Y|i,\mathcal{X},Y_{i+1},\ldots,Y_k) \\ =& \sum_{i=1}^{k} \frac{1}{k} \cdot I(\Pi^*(\mathcal{X},\mathcal{Y});Y|\mathcal{X},Y_{i+1},\ldots,Y_k) \\ =& \frac{1}{k} \cdot \sum_{i=1}^{k} I(\Pi^*(\mathcal{X},\mathcal{Y});Y_i|\mathcal{X},Y_{i+1},\ldots,Y_k). \end{aligned}$$

Above, the second equality is simply relabeling (X,Y) as (X_i,Y_i), as they are identically distributed and independent of all other random variables. From here, we can repeatedly apply the chain rule. Specifically, recall that the chain rule implies that

$$I(\Pi^*(\mathcal{X},\mathcal{Y});Y_k|\mathcal{X}) + I(\Pi^*(\mathcal{X},\mathcal{Y});Y_{k-1}|\mathcal{X},Y_k) = I(\Pi^*(\mathcal{X},\mathcal{Y});Y_{k-1},Y_k|\mathcal{X}).$$

More generally, for any i, $I(\Pi^*(\mathcal{X},\mathcal{Y});Y_i,\ldots,Y_k|\mathcal{X}) + I(\Pi^*(\mathcal{X},\mathcal{Y});Y_{i-1}|\mathcal{X},Y_i,\ldots,Y_k) = I(\Pi^*(\mathcal{X},\mathcal{Y});Y_{i-1},\ldots,Y_k|\mathcal{X})$. As such, we get that:

$$\frac{1}{k} \cdot \sum_{i=1}^{k} I(\Pi^*(\mathcal{X},\mathcal{Y});Y_i \mid \mathcal{X},Y_{i+1},\ldots,Y_k) = \frac{1}{k} \cdot I(\Pi^*(\mathcal{X},\mathcal{Y});\mathcal{Y} \mid \mathcal{X}).$$

We may now conclude that $I(\Pi;Y|X) = \frac{1}{k} \cdot I(\Pi^*(\mathcal{X},\mathcal{Y});\mathcal{Y}|\mathcal{X})$. The exact same argument swapping the roles of X and Y yields that $I(\Pi(X,Y);X \mid Y) = \frac{1}{k} \cdot I(\Pi^*(\mathcal{X},\mathcal{Y});\mathcal{X} \mid \mathcal{Y})$. These are the key claims: even though the communication of the protocol π may be huge, the information complexity is small. In a formal sense, only a $1/k$ fraction of the information conveyed throughout protocol π is conveyed about the specific indices where we placed (X,Y). From here, we now conclude:

$$\begin{aligned} IC_\mu(\pi) &= I(\Pi(X,Y);X \mid Y) + I(\Pi(X,Y);Y \mid X) \\ &= \frac{1}{k}(I(\Pi^*(\mathcal{X},\mathcal{Y});\mathcal{X} \mid \mathcal{Y}) + I(\Pi^*(\mathcal{X},\mathcal{Y});\mathcal{Y} \mid \mathcal{X})) \\ &= \frac{1}{k} IC_{\mu^k}(\pi^*) \leq \frac{1}{k} CC(\Pi). \end{aligned}$$

Recall that we assumed for for contradiction that $CC(\Pi) < ck$, and therefore we would conclude that $IC_\mu(\pi) < c$. The contrapositive proves the theorem statement. ■

D. The Information Complexity of FAR-SETS

Theorem V.11 reduces our search for a communication lower bound on EXIST-FAR-SETS to a search for an information complexity lower bound on FAR-SETS. In this section, we prove such a bound. To get started, let's first define the distribution μ for which we'll prove our information complexity lower bound:

Definition V.4. Let μ denote the uniform distribution over all pairs of sets (X,Y) of size $\frac{m}{2}+1$ such that $|X \triangle Y| = 2$.

Recall that intuitively, the goal of this section is to prove that any protocol that solves FAR-SETS correctly with probability bounded away from $1/2$ *on all instances* must result in Alice learning *some* information about Y (or Bob learning some information about X) when this protocol is run on instances (X,Y) drawn from μ. Note that it is crucial to assume that the protocol is correct on *all* instances, and not just those drawn from μ (as all instances drawn from μ can be correctly answered by communicating nothing and outputting 0).

Theorem V.12. *Let π be a randomized protocol (with public or private randomness) that solves* FAR-SETS *correctly with probability greater than 0.8 on every input (that satisfies the promise). Then $IC_\mu(\pi) > \frac{1}{4m^5}$.*

Proof: For simplicity of notation, assume that m is a multiple of 4 (if not, it is just an issue of being more careful with indices). For further simplicity of notation, let $n = m/2$. The main idea of the proof is to derive a contradiction from the following two arguments. First, if $IC_\mu(\pi)$ is negligible, then there exists a "chain" of sets $S_1, \ldots, S_n$ such that $|S_1 \cap S_n| = 2$; and for all i, $|S_i \triangle S_{i+1}| = 2$, Alice with input S_i must not be able to effectively distinguish between when Bob has S_{i+1} or S_{i-1} just from the transcript of π.

On the other hand, if π solves FAR-SETS on *every* instance (that satisfies the promise), then when Alice has input S_1, Alice must be able to effectively distinguish between when Bob has S_2 and S_n just from the transcript

of π (because the correct output is different in these two cases, and the distribution of transcripts must therefore be noticeably different). Due to the nature of communication protocols, these two claims will turn out to be contradictory. These are the key steps in the proof approach; we now begin formally. In the lemma statement below, for two random variables X, Y with finite support we let $||X - Y||_1 := \sum_z |\Pr[X = z] - \Pr[Y = z]|$.

Proposition V.13. *Let π be any private-randomness protocol that correctly solves* FAR-SETS *on every instance (satisfying the promise) with probability at least* 0.8*. Then no chain of sets $S_1, \ldots, S_n$ satisfies the following properties simultaneously:*

- $|S_i| = m/2 + 1$ *for all* i.
- $|S_1 \cap S_n| = 2$ *and for all* $1 \leq i < n$, $|S_i \triangle S_{i+1}| = 2$.
- *For all odd* $1 < i < n$, $||\Pi(S_i, S_{i-1}) - \Pi(S_i, S_{i+1})||_1 \leq 1/m^2$.
- *For all even* $1 < i < n$, $||\Pi(S_{i-1}, S_i) - \Pi(S_{i+1}, S_i)||_1 \leq 1/m^2$.

Proof: Assume for contradiction that such a chain of sets exists. By Lemma V.9, there exist functions $P(\cdot, \cdot)$ and $Q(\cdot, \cdot)$ such that $\Pr[\Pi(X, Y) = z] = P(z, X) \cdot Q(z, Y)$. We thus have (simply by expanding the bottom two hypotheses).

$$\sum_z |P(z, S_i) \cdot (Q(z, S_{i-1}) - Q(z, S_{i+1}))| \leq \frac{1}{m^2} \ (i \text{ odd});$$

$$\sum_z |Q(z, S_i) \cdot (P(z, S_{i-1}) - P(z, S_{i+1}))| \leq \frac{1}{m^2} \ (i \text{ even}).$$

For notational simplicity for the remainder of the proof, we will denote by $a_i(z) := P(z, S_i)$ when i is even, and $a_i(z) := Q(z, S_i)$ when i is odd. Observe that the above equations are then simply:

$$\sum_z |a_i(z)(a_{i-1}(z) - a_{i+1}(z))| \leq \frac{1}{m^2}, \quad 1 < i < n.$$

Finally, let $s(z) := \sum_{1<i<n} |a_i(z)(a_{i-1}(z) - a_{i+1}(z))|$. The following lemma bounds $a_1(z)(a_2(z) - a_n(z))$ in terms of $s(z)$. Recall that $a_1(z) \cdot a_2(z) = \Pr[\Pi(S_1, S_2) = z]$ and $a_1(z) \cdot a_n(z) := \Pr[\Pi(S_1, S_n) = z]$. So the lemma is bounding some term having to do with the difference between $\Pi(S_1, S_2)$ and $\Pi(S_1, S_n)$ in terms of the sums of differences between $\Pi(S_i, S_{i-1})$ and $\Pi(S_i, S_{i+1})$.

Lemma V.14. *For all z, we have $a_1(z)(a_2(z) - a_n(z)) \leq ms(z)$.*

Proof: To ease notational burden through this proof, we will drop the parameter z (since the proof is independent of z). In particular, we will use terms a_i throughout the proof and s, where $s := \sum_{1<i<n} |a_i \cdot (a_{i-1} - a_{i+1})|$. Observe that $a_i(a_{i-1} - a_{i+1}) \leq s$ for all i.

If $a_1 a_2 \leq ms$, then the lemma statement follows trivially, as $a_1, a_n \geq 0$. So now consider when that $a_1 a_2 > ms$. In this case, we will define new $b_1, \ldots, b_n$ for which $b_i \leq a_i$ for all i, and analyze these instead.

To this end, define $b_1 = a_1$, $b_2 = a_2$, and for $2 < i \leq n$ define $b_i := b_{i-2} - s/b_{i-1}$. Observe that the b_is satisfy the following equality (to see this, substitute $b_{i-1} - s/b_i$ for b_{i+1}):

$$b_i b_{i+1} = b_i b_{i-1} - s. \tag{1}$$

Since $b_1 b_2 = a_1 a_2 > ms$, this means that for all $i < n$ we have $b_i b_{i+1} > (m - i + 1)s > 0$ (recall that $n = \frac{m}{2}$). Since b_1 and b_2 are strictly positive, we conclude that all b_i are strictly positive as well. Now, we claim that $b_i \leq a_i$ for all i. Indeed, this is true for $i = 1, 2$. We now prove this by induction for $i > 2$. Assume for inductive hypothesis that $b_j \leq a_j$. Then $b_{j+1} = b_{j-1} - \frac{s}{b_j} \leq a_{j-1} - \frac{s}{a_j} \leq a_{j+1}$. The last step follows from the equation $a_j(a_{j-1} - a_{j+1}) \leq s$.

Now, we wish to prove further properties of the b_is towards our end goal. We show the following inequality by induction:

$$b_1(b_2 - b_{2i}) \leq s\left(\frac{m}{m-1} + \frac{m}{m-3} + \cdots + \frac{m}{m-(2i-3)}\right). \tag{2}$$

Proof of Equation 2:

Observe first for $i = 2$ that we have:

$$b_1(b_2 - b_4) = b_1\left(b_2 - \left(b_2 - \frac{s}{b_3}\right)\right) = \frac{b_1}{b_3} \cdot s = s \cdot \frac{b_1}{b_1 - \frac{s}{b_2}} = s \cdot \frac{b_1 b_2}{b_1 b_2 - s} = s \cdot \frac{a_1 a_2}{a_1 a_2 - s}.$$

Now, $\frac{a_1 a_2}{a_1 a_2 - s}$ is decreasing in $a_1 a_2$, and we assumed already that $a_1 a_2 > ms$, so we have

$$b_1(b_2 - b_4) = s \cdot \frac{a_1 a_2}{a_1 a_2 - s} < s \cdot \frac{ms}{ms - s} = s \cdot \frac{m}{m-1}.$$

This proves the base case ($i = 2$). Now assume for inductive hypothesis that the inequality holds for $i - 1$. We then have:

$$\begin{aligned} b_1(b_2 - b_{2i}) &= b_1(b_2 - b_{2(i-1)}) + b_1(b_{2(i-1)} - b_{2i}) \\ &= b_1(b_2 - b_{2(i-1)}) + b_1 \cdot \frac{s}{b_{2i-1}}. \end{aligned}$$

The last step here follows from Equation 1. From here we continue with:

$$\begin{aligned} & b_1(b_2 - b_{2(i-1)}) + b_1 \cdot \frac{s}{b_{2i-1}} \\ = \ & b_1(b_2 - b_{2(i-1)}) + b_1 \cdot \frac{s}{b_{2i-3} - \frac{s}{b_{2i-2}}} \\ = \ & b_1(b_2 - b_{2(i-1)}) + b_1 \cdot \frac{b_{2i-2} s}{b_{2i-3} b_{2i-2} - s}. \end{aligned}$$

From here, we now apply Equation (1) to the term in the denominator $2i - 4$ times. That is, we successively replace $b_j b_{j+1}$ with $b_j b_{j-1} - s$, $2i - 4$ times. This leaves us with:

$$b_1(b_2 - b_{2i}) = b_1(b_2 - b_{2(i-1)}) + b_1 \cdot \frac{b_{2i-2} s}{b_1 b_2 - (2i-3)s}. \tag{3}$$

Now, it is easy to see from the definition of b_is that for all j, $b_{j+2} \leq b_j$. As such, we also have $b_{2i-2} \leq b_2$. Finally,

note that the denominator $(b_1b_2-(2i-3)s)$ is positive (since we are in the case that $b_1b_2 = a_1a_2 > ms$), so we can now write:

$$\begin{aligned} b_1(b_2 - b_{2i}) &\leq b_1(b_2 - b_{2(i-1)}) + s \cdot \frac{b_1b_2}{b_1b_2 - (2i-3)s} \\ &\leq b_1(b_2 - b_{2(i-1)}) + s \cdot \frac{ms}{ms - (2i-3)s} \\ &\leq s\left(\frac{m}{m-1} + \frac{m}{m-3} + \cdots + \frac{m}{m-(2i-3)}\right). \end{aligned}$$

The penultimate step above follows from the fact that $\frac{b_1b_2}{b_1b_2-(2i-3)s}$ decreasing in b_1b_2 and we have assumed that $b_1b_2 > ms$. The last step follows by the inductive hypothesis. This completes the proof of Equation (2). ■

Now we return to the proof of Lemma V.14 with Equation (2) in hand. Now, since $a_n \geq b_n$, we have:

$$a_1(a_2 - a_n) = b_1(b_2 - a_n) \leq b_1(b_2 - b_n) \leq ms\left(\frac{1}{m-1} + \frac{1}{m-3} + \cdots + \frac{1}{m-\left(\frac{m}{2}-3\right)}\right).$$

In the right-most term, observe that there are $m/4 - 1$ terms in the sum, each of which are at most $2/m$. As such, the total sum of these terms is ≤ 1. Therefore, the right-hand-side above is at most ms, and $a_1(a_2 - a_n) \leq ms$, as desired. This concludes the proof of Lemma V.14. ■

Now, we return to the proof of Proposition V.13. Recall that the purpose of Lemma V.14 is to claim that the random variables $\Pi(S_1, S_2)$ and $\Pi(S_1, S_n)$ are not that different (which would contradict that π is correct with probability at least 0.8 on both (S_1, S_2) and (S_1, S_n)). Observe by Lemma V.14 that:

$$\begin{aligned} &\sum_{z, a_2(z) \geq a_n(z)} a_1(z)\left(a_2(z) - a_n(z)\right) \\ \leq\ & m\left(\sum_{z, a_2(z) \geq a_n(z)} s(z)\right) \leq m\sum_z s(z) \\ =\ & m\sum_{i=2}^{n-1}\sum_z |a_i(z)(a_{i-1}(z) - a_{i+1}(z))| \\ \leq\ & \frac{m(n-2)}{m^2} \leq \frac{1}{2}. \end{aligned}$$

Above, recall that $n := m/2$, and that we have assumed for contradiction that $\sum_z |a_i(z)(a_{i-1}(z) - a_{i+1}(z)| \leq 1/m^2$ for all i. (Recall that $\sum_z |a_i(z)(a_{i-1}(z) - a_{i+1}(z)| := ||\Pi(S_i, S_{i-1}) - \Pi(S_i, S_{i+1})||_1$ when i is odd, or $||\Pi(S_{i-1}, S_i) - \Pi(S_{i+1}, S_i)||_1$ when i is even, both of which are assumed to be $\leq 1/m^2$ in the proposition statement.)

Invoking that $a_1(z)a_2(z) = \Pr[\Pi(S_1, S_2) = z]$, and $a_1(z)a_n(z) = \Pr[\Pi(S_1, S_n) = z]$, we further conclude that:

$$\begin{aligned} &\sum_{z, \Pr[\Pi(S_1,S_2)=z] \geq \Pr[\Pi(S_1,S_n)=z]} |\ \Pr[\Pi(S_1, S_2) = z] \ - \\ &\qquad \Pr[\Pi(S_1, S_n) = z]\ | \\ =\ &\sum_{z, a_2(z) \geq a_n(z)} a_1(z)(a_2(z) - a_n(z)) \leq \frac{1}{2}. \end{aligned}$$

But note that $\sum_z \Pr[\Pi(S_1, S_2) = z] = \sum_z \Pr[\Pi(S_1, S_n) = z] = 1$, which means that:

$$\begin{aligned} &\sum_{z, \Pr[\Pi(S_1,S_2)=z] \geq \Pr[\Pi(S_1,S_n)=z]} |\ \Pr[\Pi(S_1, S_2) = z] \ - \\ &\qquad \Pr[\Pi(S_1, S_n) = z]\ | \ = \\ &\sum_{z, \Pr[\Pi(S_1,S_2)=z] < \Pr[\Pi(S_1,S_n)=z]} |\ \Pr[\Pi(S_1, S_2) = z] \ - \\ &\qquad \Pr[\Pi(S_1, S_n) = z]\ |, \end{aligned}$$

so in fact both sums are $\leq 1/2$. We may now conclude that:

$$||\Pi(S_1, S_2) - \Pi(S_1, S_n)||_1 \leq 1. \tag{4}$$

Now, we are ready to wrap up the proof of Proposition V.13, as Equation 4 asserts that the random variables $\Pi(S_1, S_2)$ and $\Pi(S_1, S_n)$ are too similar in order for protocol π to be correct with probability at least 0.8 on both (S_1, S_2) and also (S_1, S_n).

Indeed, let Z_b be the set of transcripts that output b. Then because π is correct with probability at least 0.8, $\Pi(S_1, S_n)$ assigns mass at least 0.8 to $z \in Z_1$, and $\Pi(S_1, S_2)$ assigns mass at most 0.2. So terms in Z_1 alone contribute at least 0.6 to the difference. In addition, $\Pi(S_1, S_n)$ assigns mass at most 0.2 to $z \in Z_0$, and $\Pi(S_1, S-2)$ assigns mass at least 0.8. So terms in Z_0 contribute at least 0.6 to the difference. Therefore, because π is correct, we must have $||\Pi(S_1, S_2) - \Pi(S_1, S_n)||_1 \geq 1.2$, contradicting our conclusion above.

This concludes the proof of Proposition V.13. To recap: we first showed how to relate the statistical difference between $\Pi(S_1, S_2)$ and $\Pi(S_1, S_n)$ to the statistical distance between adjacent pairs $\Pi(S_i, S_{i+1})$ and $\Pi(S_{i+2}, S_{i+1})$ (Lemma V.14). Lemma V.14 required a decent amount of math, the bulk of which is in proving Equation (2). With Lemma V.14, we then obtained a contradiction: the statistical difference between $\Pi(S_1, S_2)$ and $\Pi(S_1, S_n)$ cannot be small if π is to possibly be correct on both (S_1, S_2) and (S_1, S_n), because the answers must be different with good probability.

■

Proposition V.13 claims that no chains of the proposed form can exist for any private-randomness protocol which solves all instances of FAR-SETS with probability 0.8. Now we will claim that any protocol which has low information complexity with respect to μ must have some chain of the proposed form. To this end, we first give a few definitions:

Definition V.5. Let $n = \frac{m}{2}$ as before.

- A *link* is an ordered pair (X, Y) of sets of size $n+1$ such that $|X \triangle Y| = 2$.
- A *chain* is a sequence of sets $T_1, \ldots, T_n$ such that (T_i, T_{i+1}) is a link for $1 \leq i < n$, and $|T_1 \cap T_n| = 2$.
- Given a link (X, Y), recall that $\Pi(X, Y)$ is the random variable denoting the transcript when π is run on input (X, Y). Further define $\Pi(X, ?)$ to be the random variable which first samples a uniformly random Y such that (X, Y) is a link (note that this is simply sampling a uniformly random element of X to remove, and a uniformly random element $\notin X$ to add), and then samples $\Pi(X, Y)$. Define $\Pi(?, Y)$ similarly.
- A link (X, Y) is *broken* if $\mathbb{D}(\Pi(X, Y) \parallel \Pi(X, ?)) > \frac{1}{8m^4}$ or if $\mathbb{D}(\Pi(X, Y) \parallel \Pi(?, Y)) > \frac{1}{8m^4}$. Here, $\mathbb{D}(\cdot \parallel \cdot)$ represents Kullback-Leibler divergence from Definition V.5.
- A chain $T_1, \ldots, T_n$ is *broken* if for some odd i, (T_i, T_{i+1}) or (T_i, T_{i-1}) is a broken link. These $n-1$ links are the chain's *structural links*.

Proposition V.15. *Let π be a protocol with $IC_\mu(\pi) \leq \frac{1}{8m^4 N}$. Then at most a $1/N$ fraction of links are broken.*

Proof: Let's first compute $I(\Pi(X, Y); X | Y)$ when (X, Y) are drawn from μ. In order to draw from μ, we may first draw X uniformly at random from all sets of size $n+1$, and then draw Y uniformly at random among all sets of size $n+1$ with $|X \triangle Y| = 2$ (note that there are $(n+1)(n-1)$ such sets: pick an element in X to kick out and an element $\notin X$ to add). So we get that $I(\Pi(X, Y); X \mid Y)$ is equal to:

$$\frac{1}{\binom{m}{n+1}} \sum_{\substack{T \subseteq M \\ |T| = n+1}} I(\Pi(X, Y); X \mid Y = T) =$$

$$\frac{1}{\binom{m}{n+1}} \sum_{\substack{T \subseteq M \\ |T| = n+1}} \sum_{\substack{S \text{ where} \\ (S,T) \\ \text{is a link}}} \frac{1}{(n+1)(n-1)} \mathbb{D}(\Pi(S, T) \parallel \Pi(?, T)).$$

The first step above is just expanding the definition of conditional mutual information. The second step requires some further explanation. First, note that the number of S such that (S, T) is a link is $(n+1)(n-1)$ (one of the $n+1$ elements of T needs to be kicked out and one of the $n-1$ elements of T needs to be put in to form S), and each of these sets are equally likely to be drawn from μ. The second step is again just unraveling the definition of mutual information via Fact V.6.

Now, the above sum can be written to directly sum over all links. That is:

$$I(\Pi(X, Y); X \mid Y) =$$
$$\frac{1}{\binom{m}{n+1} \cdot (n+1)(n-1)} \cdot \sum_{\text{links } (S,T)} \mathbb{D}(\Pi(S, T) || \Pi(?, T)).$$

Identical math concludes that $I(\Pi(X, Y); Y \mid X) = \frac{1}{\binom{m}{n+1} \cdot (n+1)(n-1)} \cdot \sum_{\text{links } (S,T)} \mathbb{D}(\Pi(S, T) || \Pi(S, ?))$. This means that $IC_\mu(\pi)$ is equal to $\frac{1}{\binom{m}{n+1} \cdot (n+1)(n-1)} D(\Pi(S, T) \parallel \Pi(?, T)) + D(\Pi(S, T) \parallel \Pi(S, ?))$. Observe that this is simply the average over all links of $D(\Pi(S, T) \parallel \Pi(?, T)) + D(\Pi(S, T) \parallel \Pi(S, ?))$. So if a $1/N$ fraction of all links are broken, then the average of this quantity over all links is at least $\frac{1}{8m^4 N}$, meaning that $IC_\mu(\pi) \geq \frac{1}{8m^4 N}$. ∎

Now, we claim that if most links are not broken, there must exist an entire chain which is not broken.

Lemma V.16. *If fewer than $\frac{2}{m}$ links are broken, there exists an unbroken chain.*

Proof: We first claim that the number of chains is $\frac{m!}{2}$. Indeed, consider the structure of a chain $T_1, \ldots, T_n$. Observe that there are exactly two items that persist in $\cap_i T_i$. Moreover, each T_i adds a unique element to T_{i-1} (that previously wasn't added) and removes another one (which previously wasn't removed). So if we order the m elements so that the first two elements are in $\cap_i T_i$, and the next $n-1$ elements are added in $T_2, T_3, \ldots, T_n$, and the next $n-1$ elements are removed in $T_2, \ldots, T_n$, this defines a chain. The chain defined is invariant under flipping the order of the first two elements, but modulo this, each ordering defines a unique chain.

As counted previously, the total number of links is $\binom{m}{n+1}(n-1)(n+1)$: for a link (X, Y) there are $\binom{m}{n+1}$ choices for X and $(n-1)(n+1)$ choices for Y conditioned on this. Each chain consists of $n-1$ structural links, so the number of pairs $(\mathcal{C}, L)$ such that L is a link in chain $\mathcal{C}$ is $\frac{m!}{2}(n-1)$. By symmetry, every link is a structural link of the same number of chains, so for a given link, the number of chains which contain it as a structural link is:

$$\frac{m!(n-1)}{2\binom{m}{n+1}(n-1)(n+1)} = \frac{m!}{2\binom{m}{n+1}(n+1)}.$$

Thus, for all $\frac{m!}{2}$ chains to be broken, there must be at least $m!/2$ pairs $(\mathcal{C}, L)$ such that $\mathcal{C}$ is a chain and L is a broken structured link of $\mathcal{C}$. By the above counting, this means there must be at least $\binom{m}{n+1}(n+1)$ broken links. But this is $\frac{1}{n-1} = \frac{1}{m/2-1} > 2/m$ fraction of all of the links. So if fewer than this are broken, there must exist an unbroken chain. ∎

Now, we want to claim that an unbroken chain exactly satisfies the hypotheses of Proposition V.13. This is the last step in wrapping up the proof of Theorem V.12.

Lemma V.17. *Let $S_1, \ldots, S_n$ be an unbroken chain. Then for all odd $1 < i < n$, $||\Pi(S_i, S_{i-1}) - \Pi(S_i, S_{i+1})||_1 \leq 1/m^2$, and for all even $1 < i < n$, $||\Pi(S_{i-1}, S_i) - \Pi(S_{i+1}, S_i)||_1 \leq 1/m^2$.*

Proof: Because the chain is unbroken, we get that for all odd $i > 1$, $\mathbb{D}(\Pi(S_i, S_{i+1} \parallel \Pi(S_i, ?))$ and $\mathbb{D}(\Pi(S_i, S_{i-1}) \parallel \Pi(S_i, ?))$ are both at most $\frac{1}{8m^4}$. Now, Pinsker's inequality allows us to conclude that $||\Pi(S_i, S_{i+1}) - \Pi(S_i, ?)||_1 \leq \sqrt{\frac{2}{8m^4}} = \frac{1}{2m^2}$ for all i, and also that $||\Pi(S_i, S_{i+1}) -$

$\Pi(?, S_{i+1})||_1 \leq \frac{1}{2m^2}$ for all i. By the triangle inequality, we therefore conclude that for odd i:

$$||\Pi(S_i, S_{i+1}) - \Pi(S_i, S_{i-1})||_1 \leq 1/m^2$$

And for even i:

$$||\Pi(S_{i+1}, S_i) - \Pi(S_{i-1}, S_i)||_1 \leq 1/m^2.$$

■

Now, we are finally ready to conclude the proof of Theorem V.12. We have two contradictory lines of argument: on one hand, if π uses private randomness and is correct on every input with probability at least 0.8, then Proposition V.13 combined with Lemma V.17 claims that no unbroken chain can exist. However, Proposition V.15 combined with Lemma V.16 claims that if $IC_\mu(\pi) < \frac{1}{4m^5}$, then there is an unbroken chain. So no private-randomness protocol can simultaneously have $IC_\mu(\pi) < \frac{1}{4m^5}$ and correctly solve FAR-SETS on every input with probability at least 0.8. From Lemma V.10 it also follows that no protocol, whether making use of public randomness or not, can simulatenously have $IC_\mu(\pi) < \frac{1}{4m^5}$ and correctly solve FAR-SETS on every input with probability at least 0.8. ■

E. Wrapping everything up

Now, we can complete the proof of Theorem V.1 by combining everything together. We first need to combine Theorem V.12 and Theorem V.11 to conclude that randomized protocols for EXIST-FAR-SETS require exponential communication.

Theorem V.18. *For all m, ℓ, k such that $\ell \leq (1-c)\log_3(m)$ for some constant $c > 0$ and $k \in (m, exp(\frac{m^c}{2\log_3(m)}))$, any randomized protocol π that solves* EXIST-FAR-SETS *with probability at least $2/3$ on all instances (which satisfy the promise) has $CC(\pi) = \Omega(k/m^5)$.*

Proof: We already have a distribution μ for which any protocol π solving FAR-SETS with probability at least 0.8 on all inputs has $IC_\mu(\pi) > \frac{1}{4m^5}$. It is also clear that EXIST-FAR-SETS$(\mathcal{X}, \mathcal{Y}) = \bigvee_{i\in[k]}$ FAR-SETS(X_i, Y_i). So we just need to check the details with respect to the promises. In particular, we just need to see for what values of (k, z), μ is and the FAR-SETS promise are (k, z)-safe with respect to the EXIST-FAR-SETS promise.

Observe that conditions (1) and (2) of the EXIST-FAR-SETS promise are trivially satisfied. So we just need to check conditions (3) and (4). Note that these conditions depend only on Alice's sets and, separately, Bob's sets, and not on how Alice's and Bob's sets interact. Now, the EXIST-FAR-SETS promise is invariant under permutations of M, and also under permutations of the indices (i.e. if $(X_1, Y_1), \ldots, (X_k, Y_k)$ satisfy the EXIST-FAR-SETS promise, then so do $(X_{\sigma(1)}, Y_{\sigma(1)}), \ldots, (X_{\sigma(k)}, Y_{\sigma(k)})$ for any permutation σ from $[k]$ to $[k]$). Therefore, we may treat Alice as having i.i.d. subsets of M of size $\frac{m}{2} + 1$, and likewise with Bob. We are interested in bounding the probability that Alice's sets do not satisfy conditions (3) and (4); then, a simple union bound will give us the probability that the EXIST-FAR-SETS promise is satisfied.

Lemma V.19. *Let $\mathcal{X} = \langle X_1, \ldots, X_k\rangle$ be drawn so that each X_i is an i.i.d. uniformly random set of size $m/2+1$. For any x, let $\ell := \log_3(m) - \log_3(x)$, and let $k \leq e^{\frac{x}{2\ell}}$. Then with probability at least $1 - e^{-x/2} - ke^{-\Omega(m)}$, $\mathcal{X}$ is ℓ-sparse.*

Proof: Consider the following roundabout way to draw X_i: first place each element of M in X_i' independently with probability $2/3$. Then, if $|X_i'| \geq m/2+1$, let X_i be a random subset of X_i' of size $m/2+1$. If any $|X_i'| < m/2+1$, abort the entire process and consider it a failure. Observe that $X_i \subseteq X_i'$.

For a fixed item j, and fixed set L of indices with $|L| = \ell$, the probability that $j \in \cup_{i\in L} X_i'$ is $1-(1/3)^\ell$. Because these events are independent, the probability that $M \subseteq \cup_{i\in L} X_i'$ is exactly $(1 - (1/3)^\ell)^m$. Taking a union bound over all $\binom{k}{\ell}$ subsets we get that the probability the collection is *not* ℓ-sparse, conditioned on not failing initially, is at most:

$$\begin{aligned}\binom{k}{\ell}\left(1 - 1/3^\ell\right)^m &\leq k^\ell \exp\left(\frac{-m}{3^\ell}\right)\\ &= \exp\left(\ell\ln(k) - x\right) \leq \exp(-x/2).\end{aligned}$$

Finally, observe that the expected number of items in each X_i' is $2m/3$. So the probability that a single X_i' fails to have $m/2+1$ elements is $e^{-\Omega(m)}$ by the Chernoff bound. Taking a union bound over all k X_i's accounts for the additional $ke^{-\Omega(m)}$ term. ■

Lemma V.20. *Let $\mathcal{X} = \langle X_1, \ldots, X_k\rangle$ be drawn so that each X_i is an i.i.d. uniformly random set of size $m/2+1$. Then with probability at least $1 - e^{-k/3^\ell} - e^{-\Omega(m)}$, for all sets T of $|T| = \ell$, there exists an $X_i \supseteq T$.*

Proof: Let's again consider the following roundabout way to draw X_i: first place each element of M in X_i' independently with probability $1/3$. Then if $|X_i'| \leq m/2+1$, let X_i be a random superset of X_i' of size $m/2+1$. If any $|X_i'| > m/2+1$, abort the entire process and consider it a failure. Observe that $X_i \supseteq X_i'$.

For a fixed set T of size ℓ, The probability that X_i' contains T is just $1/3^\ell$. So the probability that no X_i' contains T is $(1 - 1/3^\ell)^k \leq \exp(-k/3^\ell)$. Again, the probability of failure is at most $ke^{-\Omega(m)}$, resulting in the lemma statement. ■

Now to wrap up, we observe that the probability that the EXIST-FAR-SETS promise is *not* satisfied is at most $2(ke^{-\Omega(m)} + e^{-k/3^\ell} + e^{-x/2})$, where $\ell := \log_3(m) - \log_3(x)$ and $k \leq e^{\frac{x}{2\ell}}$. Here, the factor of 2 comes from a union bound from the events that Alice's sets fail condition (3) or (4) and that Bob's sets fail condiiton (3) or (4).

When $x = m^c$, for any $c < 1$, we get that $\ell := (1 - c)\log_3(m)$, and $k \leq e^{\frac{m^c}{2(1-c)\log_3(m)}}$. As such, $ke^{-\Omega(m)} =$

$o(1)$, $e^{-x/2} = o(1)$, and furthermore $e^{-k/3^\ell} = o(1)$ as long as $k = \Omega(m)$.

Theorem V.11 combined with Theorem V.12 then implies that any protocol π for EXIST-FAR-SETS which succeeds on all inputs (satisfying the promise) with probability at least $0.8/(1 - o(1))$ has $CC(\pi) \geq \frac{k}{4m^5}$. To replace $0.8/(1 - o(1))$ with $2/3$ in the Theorem statement, observe that if we had a protocol with success probability $2/3$, we could repeat it independently a constant number of times and take a majority to get a protocol with success probability $0.8/(1-o(1))$. ∎

Proof of Theorem V.1: The proof of Theorem V.1 now follows immediately from Theorem V.18 and Corollary V.3. In particular, by plugging in $\ell = \log_3(m)/3$, we may take $c = 2/3$, and $k = 4m^5 \cdot e^{\sqrt{m}}$ in Theorem V.18 to conclude that any randomized protocol guaranteeing a $(1/2+\frac{6}{\log(m)})$-approximation for WELFARE-MAXIMIZATION with probability at least $2/3$ on all instances requires communication $\Omega(e^{\sqrt{m}})$.

∎

VI. FROM XOS TO SUBADDITIVE: THE MPH HIERARCHY

In this section, we explore the space between fractionally subadditive and subadditive functions via the *MPH hierarchy*.

A. MPH Preliminaries

Let's first review the definitions.

Definition VI.1 (Positive-Hypergraph [ABDR12])**.** *A valuation function $v(\cdot)$ is PH-k if there exists a non-negative set function $w(\cdot)$ such that for all S, $v(S) = \sum_{T \subseteq S, |T| \leq k} w(T)$.*

Observe that PH-1 functions are exactly additive functions.

Definition VI.2 (Maximum-over-Positive-Hypergraphs [FFI$^+$15])**.** *A valuation function $v(\cdot)$ is MPH-k if there exists a collection $\mathcal{F}$ of PH-k valuation functions such that for all S, $v(S) = \max_{f \in \mathcal{F}}\{f(S)\}$.*

Observe that MPH-1 functions are exactly fractionally subadditive (XOS) functions.

The next observation follows directly from the definitions.

Observation VI.1. *Let $v(\cdot)$ be MPH-k. Then for every item set S, there exists a non-negative set function $w_S(\cdot)$ such that (i) $v(S) = \sum_{T \subseteq S} w_S(T)$; (ii) $v(S') \geq \sum_{T \subseteq S'} w_S(T)$ for every $S' \neq S$; and (iii) $w_S(T) = 0$ for every set T of size $|T| > k$.*

Proof: By definition, $v(\cdot)$ is a maximum over PH-k functions. Let $f_S(\cdot)$ be the PH-k function that is the argmaximizer at S, i.e., $v(S) = f_S(S)$; then clearly $v(S') \geq f_S(S')$ for every S'. Because $f_S(\cdot)$ is PH-k, by definition there exists a non-negative set function $w(\cdot)$ for which all conditions are satisfied. ∎

The following lemma lower-bounds the MPH level of a set function $v(\cdot)$ based on the ratio between the sum of marginal contributions of the items to the grand bundle and the grand bundle's value.

Lemma VI.2. *Let $v(\cdot)$ be any set function. Then $v(\cdot)$ is not MPH-k for $k < \frac{\sum_j v(M) - v(M \setminus \{j\})}{v(M)}$.*

Proof: Let k be the level of $v(\cdot)$ in the MPH hierarchy (i.e., $v(\cdot)$ is MPH-k). Then by Observation VI.1 there exists a non-negative set function $w(\cdot)$ such that (i) $\sum_T w(T) = v(M)$; (ii) $v(M \setminus \{j\}) \geq \sum_{T \subseteq M \setminus \{j\}} w(T)$ for all j; and (iii) $w(T) = 0$ for every T with $|T| > k$. Summing both sides of (ii) over all items j yields:

$$\begin{aligned}
\sum_{j \in M} v(M \setminus \{j\}) &\geq \sum_j \sum_{T \subseteq M \setminus \{j\}} w(T) \\
&= \sum_T w(T)(m - |T|) \\
&\geq \sum_T w(T)(m - k) \\
&= v(M)(m - k).
\end{aligned}$$

Rearranging the inequality yields $kv(M) \geq \sum_j v(M) - v(M \setminus \{j\})$, as required. ∎

[FFI$^+$15] show that every monotone valuation function is MPH-m, and that there exist subadditive functions that are not in MPH-k for any $k < m/2$. In the Appendix we show that this is tight; i.e., that every monotone subadditive valuation function is MPH-$\lceil \frac{m}{2} \rceil$ (Proposition A.1).

Clearly, for any $k > 1$, there exist MPH-k functions that are not complement-free (i.e., subadditive). Indeed, even MPH-2 functions exhibit complementarities. Since we are interested in exploring the space of functions between XOS and subadditive, our results in this section will be for "subadditive MPH-k" functions, which belong to both classes simultaneously. The extra subadditive assumption is necessary for our results (Proposition VI.12).

B. Our Results for Subadditive MPH-k

A preliminary question to address is: what is the MPH level of the subadditive functions constructed in Section III? It turns out to be quite high:

Proposition VI.3. *For all $\mathcal{S}, \ell$, such that $f^\ell_{\mathcal{S}}(\cdot)$ is well-defined, $f^\ell_{\mathcal{S}}(\cdot)$ is not MPH-k for $k < m/\ell$.*

Proof: For every item j, $f^\ell_{\mathcal{S}}(M) - f^\ell_{\mathcal{S}}(M \setminus \{j\}) \geq 1$, and $f^\ell_{\mathcal{S}}(M) = \ell$. Applying Lemma VI.2 completes the proof. ∎

Proposition VI.3 raises the possibility to improve upon the $1/2$-approximation ratio for subadditive functions which reside in lower MPH levels. The main result of this section is a $(\frac{1}{2} + \Omega(\frac{1}{\log k}))$-approximation in poly$(m)$ communication for WELFARE-MAXIMIZATION when Alice and Bob have valuations that are both subadditive and MPH-k. This approximation guarantee is tight for sufficiently large k by direct application of our previous bounds.

Our main technical contribution in this section is an *oblivious* rounding protocol for the configuration LP when two bidders both have subadditive MPH-k valuations. The protocol's performance gradually degrades with the level k, starting from 0.625 for $k = 2$ (Section VI-D) and behaving like $\frac{1}{2} + \Omega(\frac{1}{\log k})$ in general (Section VI-E). An alternative and simpler protocol with guarantee $\frac{1}{2}+\Omega(\frac{1}{\log m})$ was developed independently for subadditive valuations by Dobzinski (which makes use of the fact that XOS valuations pointwise $1/H_m$-approximate subadditive functions).[16] It is an interesting open question to determine whether the simpler version can be extended to subadditive MPH-k valuations with the $\frac{1}{2} + \Omega(\frac{1}{\log k})$ guarantee (for instance, by proving that XOS valuations pointwise $1/H_k$-approximate subadditive MPH-k functions; note, however, that the "subadditive" is necessary for this claim to possibly be true).

We remark that the obliviousness of the protocol is known to be without loss, by the results of [FFT16]: since the class of MPH-k valuations is closed under convex combinations, by [FFT16] there exists an oblivious rounding scheme that achieves an approximation guarantee matching the integrality gap of the LP. In Section VI-F we show the matching integrality gap of $\frac{1}{2} + \Omega(\frac{1}{\log k})$ for sufficiently large k.We also remark that our rounding-based technique necessarily fails for MPH-k valuations that are not subadditive. Indeed, even for two MPH-2 valuations, the integrality gap may be as large as $\frac{1}{2}$ (Proposition VI.13).

C. Notation and Key Lemma

Throughout this section, we overload the notation $v(S)$ as follows. When S is a random set drawn from distribution D, we use $v(S)$ to denote $\mathbb{E}_{S\sim D}[v(S)]$. Also, if X denotes either Alice or Bob, we use $\overline{X}$ to denote the other player (i.e. if X = Alice, then $\overline{X}$ = Bob, and vice versa).

The following key lemma extends a well-known result of [Fei06] for XOS valuations to the MPH-k hierarchy.

Lemma VI.4. *Let $v(\cdot)$ be an MPH-k function, and let $S \subseteq M$ be a subset of items. If T is a random set such that every $U \subseteq S$ with $|U| \leq k$ is contained in T with probability at least p, then $v(T) \geq p \cdot v(S)$.*

Proof: Let $w_S(\cdot)$ be as promised from Observation VI.1. Then $v(S) = \sum_{T\subseteq S} w_S(T)$, $v(U) \geq \sum_{T\in U} w_S(T)$ for all U, and $w_S(T) = 0$ for all T with $|T| > k$. Then we can conclude that:

$$\begin{aligned} v(T) &\geq \sum_{T',|T|\leq k} \Pr[T' \subseteq T] \cdot w_S(T') \\ &\geq p \sum_{T'\subseteq S',|T'|\leq k} w_S(T') = p \cdot v(S). \end{aligned}$$

The first inequality follows from the fact that nonzero weights only belong to sets T' of size at most k. The second follows because every subset of S of size k appears in T with probability at least p. ∎

It is also not hard to extend Lemma VI.4 to functions which are "close" to MPH-k.

Definition VI.3 (Pointwise β-approximation [DMSW15])**.** *A valuation function $v(\cdot)$ is pointwise β-approximated by a valuation class $\mathcal{W}$ if for any set $S \subseteq M$, there exists a valuation $w \in \mathcal{W}$ such that: (a) $\beta w(S) \geq v(S)$ and (b) for all $T \subseteq M$, $v(T) \geq w(T)$.*

Similarly, we say that a class $\mathcal{V}$ is pointwise β-approximated by a class $\mathcal{W}$ if all $v \in \mathcal{V}$ are pointwise β-approximated by $\mathcal{W}$.

Corollary VI.5. *Let $v(\cdot)$ be pointwise β-approximated by MPH-k functions, and let $S \subseteq M$ be a subset of items. If T is a random set such that every $U \subseteq S$ with $|U| \leq k$ is contained in T with probability at least p, then $v(T) \geq \frac{p}{\beta} v(S)$.*

Proof: Let $w(\cdot)$ be the MPH-k function which pointwise β-approximates $v(\cdot)$ at S. Then we know that: $v(T) \geq w(T) \geq p \cdot w(S) \geq \frac{p}{\beta} v(S)$. The outer two inequalities are by definition of pointwise approximation. The inner inequality is by Lemma VI.4. ∎

D. Protocol for Subadditive MPH-2 Valuations

Here, we describe our protocol specifically for subadditive MPH-2, as it conveys the main ideas. Our protocol will proceed as follows. First, we will solve the configuration LP relaxation (defined shortly) which finds the optimal fractional allocation. Then, we provide an oblivious rounding which takes the fractional solution to a distribution over allocations. Assuming that both $A(\cdot)$ and $B(\cdot)$ are subadditive MPH-k, the rounding will maintain at least a $\frac{1}{2} + \Omega(1/\log(k))$ fraction of the welfare.

Let's now recall the configuration LP (defined below for any n, we only use it for $n = 2$):

- Variables: $x_i(S)$, for all bidders i and subsets $S \subseteq M$.
- Constraint: $x_i(S) \geq 0$ for all i, S.
- Constraint: For all i, $\sum_S x_i(S) = 1$ (dual variable u_i).
- Constraint: For all j, $\sum_{S\ni j} \sum_i x_i(S) \leq 1$ (dual variable p_j).
- Maximizing: $\sum_{i,S} v_i(S) \cdot x_i(S)$.

It is clear that the configuration LP is indeed a fractional relaxation, as any integral allocation is feasible. Despite having exponentially many variables, there are only $n + m$ non-trivial constraints, and so the dual has only $n + m$ variables. We quickly remind the reader of the dual:

- Variables: u_i for all bidders i, p_j for all items j.
- Constraint: $u_i \geq v_i(S) - \sum_{j\in S} p_j$ for all bidders i and subsets $S \subseteq M$.
- Constraint: $p_j \geq 0$.
- Minimize: $\sum_j p_j + \sum_i u_i$.

We remind the reader that a separation oracle for the dual can be achieved via a *demand oracle* for each $v_i(\cdot)$

[16] This protocol, which is not yet in print, was brought to our attention in a personal correspondence with Shahar Dobzinski.

(recall that this oracle takes as input a set of prices $\vec{p}$ and outputs the set maximizing $v_i(S) - \sum_{j \in S} p_j$). So the dual can indeed be solved in polynomial communication (and indeed only requires demand queries of the valuations). Once the dual is solved, an optimal primal can be recovered as well (for further details, see [BN07]). Most state-of-the-art communication protocols (including all those referenced in Table **??**) are derived by first solving this LP, and then rounding. Indeed, our protocol follows this blueprint as well. Moreover, our rounding protocol is *oblivious*: while we of course need demand-query access to the valuations in order to find the optimal fractional solution, once we have this fractional solution our rounding scheme never looks at the valuations again. The same rounding argument is guaranteed to work for *all* subadditive MPH-2 valuations.

Oblivious rounding scheme for $k = 2$. Draw player X uniformly at random from {Alice, Bob} (also referred to as $\{1, 2\}$).

1) Let $\vec{x}$ denote the input fractional allocation.
2) Let D_X denote the distribution that takes set S with probability $x_X(S)$.
3) With probability $\lambda = 0.5$, draw S_X from D_X; give set S_X to player X; and give set $M \setminus S_X$ to player $\bar{X}$.
4) Otherwise (with probability $1 - \lambda = 0.5$), draw set S'_A from D_A and set S'_B from D_B; give set $S'_A \cap S'_B$ to player X; and give set $M \setminus (S'_A \cap S'_B)$ to player $\bar{X}$.

Proposition VI.6. *Let C denote the optimal value of the configuration LP. Then when Alice and Bob are both subadditive MPH-2, the expected welfare of the above oblivious rounding scheme is $\geq 0.625 \cdot C$.*

Proof: Let bundles S_A, S_B be independent random draws from distributions D_A, D_B, respectively. The first case of the oblivious rounding scheme achieves expected welfare of (in expectation over randomly sampling X from $U(\{1, 2\})$):

$$\begin{aligned} &\frac{1}{2}(v_A(S_A) + v_B(M \setminus S_A)) \quad + \\ &\frac{1}{2}(v_A(M \setminus S_B) + v_B(S_B)) \quad \geq \\ &\qquad \frac{1}{2}(v_A(S_A) + v_B(S_B)) \quad + \\ &\frac{1}{2}(v_A(S_A \setminus S_B) + v_B(S_B \setminus S_A)) \quad , \end{aligned} \tag{5}$$

where the inequality holds pointwise for every instantiation of S_A, S_B and follows from monotonicity of v_A, v_B. Similarly, the second case of the oblivious rounding scheme achieves expected welfare of:

$$\begin{aligned} &\frac{1}{2}(v_A(S_A \cap S_B) + v_B(S_A \cap S_B)) + \\ &\frac{1}{2}(v_A(M \setminus (S_A \cap S_B)) + v_B(M \setminus (S_A \cap S_B))). \end{aligned}$$

To analyze the latter case we use the following claim:

Claim VI.7. *Let v_X be MPH-2 and let bundles S_A, S_B be independent random draws from distributions D_A, D_B, respectively. Then $v_X(M \setminus (S_A \cap S_B)) \geq \frac{1}{2} v_X(M)$.*

Proof of Claim VI.7: For every pair of items $i, j \in M$, the probability that player X gets the pair when allocated the random bundle $M \setminus (S_A \cap S_B)$ is $\geq \frac{1}{2}$. To see this, observe that for player X not to get the pair, either i or j or both must be in $S_A \cap S_B$. By definition of S_A, S_B, the probability $\Pr[i \in S_A \cap S_B]$ is equal to $(\sum_{S \ni i} x_1(S)) \cdot (\sum_{S \ni i} x_2(S))$. We further know that $\sum_{S \ni i} x_1(S) + x_2(S) \leq 1$ by the constraints in the configuration LP. This means that $(\sum_{S \ni i} x_1(S)) \cdot (\sum_{S \ni i} x_2(S)) \leq 1/4$, and therefore $\Pr[i \in S_A \cap S_B] \leq 1/4$ as well. By applying the union bound we get that i and/or j are in $S_A \cap S_B$ with probability $\leq 1/4$. Applying Lemma VI.4 for MPH-2 valuations, player X's expected value for the random bundle $M \setminus (S_A \cap S_B)$ is thus at least $\frac{1}{2} v_X(M)$. ∎

By Claim VI.7 the second case achieves expected welfare of at least

$$\frac{1}{2}(v_A(S_A \cap S_B) + v_B(S_A \cap S_B)) + \frac{1}{4}(v_A(\Omega) + v_B(\Omega)). \tag{6}$$

Summing up the contributions from (5) and (6) weighted by their respective probabilities λ and $1 - \lambda$, the total expected welfare of the oblivious rounding scheme is at least:

$$\begin{aligned} &\frac{\lambda}{2}(v_A(S_A) + v_B(S_B)) + \\ &\frac{\lambda}{2}(v_A(S_A \setminus S_B) + v_B(S_B \setminus S_A)) + \\ &\frac{1-\lambda}{2}(v_A(S_A \cap S_B) + v_B(S_A \cap S_B)) + \\ &\frac{1-\lambda}{4}(v_A(M) + v_B(M)) \\ \geq \quad &\frac{\lambda}{2}(v_A(S_A) + v_B(S_B)) + \frac{\lambda}{2}(v_A(S_A) + v_B(S_B)) + \\ &\frac{1-\lambda}{4}(v_A(M) + v_B(M)) \qquad (7) \\ \geq \quad &\frac{5}{8}(f_A(S_A) + f_B(S_B)) = 0.625 \cdot C, \qquad (8) \end{aligned}$$

where (7) holds since λ was chosen such that $\frac{1-\lambda}{2} = \frac{\lambda}{2}$, and by subadditivity of v_A, v_B (importantly, note that (7) does *not* necessarily hold without subadditivity), and (8) holds since $\lambda = \frac{1}{2}$ and by monotonicity of v_A, v_B (both inequalities hold pointwise for every instantiation of S_A, S_B). This completes the proof of Proposition VI.6. ∎

To recap the proof of Proposition VI.6, the fact that the valuations are MPH-2 means that if a player "loses" in the second rounding case and is allocated the "leftovers", this player still gets at least half of her total value (see Claim VI.7). The fact that the valuations are subadditive means that allocating the bundle in contention $S_A \cap S_B$ with some probability p to player X, and allocating the bundle not in contention $S_X \setminus (S_A \cap S_B)$ with the same probability to the same player, is as good in terms of welfare as allocating S_X to player X with probability p (see (7)).

Corollary VI.8. *If Alice and Bob are β-pointwise approximated by subadditive MPH-2 valuations, then the expected welfare of the oblivious rounding scheme is $\geq (\frac{1}{2} + \frac{1}{8\beta})\omega$.*

Corollary VI.8 follows by replacing Lemma VI.4 by Corollary VI.5 in the proof of Proposition VI.6.

E. Protocol for Subadditive MPH-k Valuations

We now generalize the oblivious rounding scheme above to subadditive MPH-k. The main idea behind the approach is similar to our protocol for MPH-2: the "level 0" protocol that one should try first is to simply draw sets S_A and S_B independently from D_A, D_B. Prioritize awarding items in $S_A \cap S_B$ to a uniformly random player, and give the leftovers to the other. Of course, this protocol might fail to beat a $1/2$-approximation. But the only way it fails is if both Alice and Bob have high expected value for the intersection $S_A \cap S_B$. If Alice and Bob have high expected value for the intersection, then we can instead just recurse within $S_A \cap S_B$. The following protocol and subsequent proof makes this formal.

Oblivious rounding scheme for general $k \geq 2$. Draw player X uniformly at random from {Alice, Bob}. Set $r = \lceil \log \log k \rceil$.

1) Let $\vec{x}$ denote the input fractional allocation.
2) Let D_X denote the distribution that takes set S with probability $x_X(S)$.
3) With probability $\lambda = \frac{1}{2}$, draw S_X from D_X; give set S_X to player X; and give set $M \setminus S_X$ to player $\bar{X}$.
4) For $0 \leq q < r$, with probability $\lambda_q = \frac{\lambda}{2^{q+1}}$, draw 2^q sets $S_A^1, \ldots, S_A^{2^q}$ i.i.d. from D_A and 2^q sets $S_B^1, \ldots, S_B^{2^q}$ i.i.d. from D_B; give set $S^q = S_A^1 \cap \cdots \cap S_A^{2^q} \cap S_B^1 \cap \cdots \cap S_B^{2^q}$ to player X, and give set $M \setminus S^q$ to player $\bar{X}$.
5) Otherwise (with probability $\lambda_r = 1 - \lambda - \sum_{q=0}^{r-1} \lambda_q = \frac{1}{2^{r+1}}$), draw 2^r sets $S_A^1, \ldots, S_A^{2^r}$ i.i.d. from D_A and 2^r sets $S_B^1, \ldots, S_B^{2^r}$ i.i.d. from D_B; give $S^r = S_A^1 \cap \cdots \cap S_A^{2^r} \cap S_B^1 \cap \cdots \cap S_B^{2^r}$ to player X, and give $M \setminus S^r$ to player $\bar{X}$.

Theorem VI.9. *Let C denote the optimal value of the configuration LP. Then when Alice and Bob are both subadditive MPH-k, the expected welfare of the above oblivious rounding scheme is $\geq (\frac{1}{2} + \Omega(\frac{1}{\log k})) \cdot C$.*

Proof: The proof generalizes that of Proposition VI.6 and proceeds by analyzing the contribution from each case of the oblivious rounding scheme. The first case of the scheme achieves expected welfare of at least $\frac{1}{2}(v_A(S_A)+v_B(S_B))+\frac{1}{2}(v_A(S_A \setminus S_B)+v_B(S_B \setminus S_A))$ (identical to (5)). For every $0 \leq q \leq r$, the corresponding case of the scheme achieves expected welfare of at least

$$\frac{1}{2}(v_A(S^q)+v_B(S^q))+\frac{1}{2}(v_A(S^q \setminus S^{q+1})+v_B(S^q \setminus S^{q+1})). \tag{9}$$

To analyze the last case we use the following claim:

Claim VI.10. *Let v_X be an MPH-k valuation and let bundles $S_A^1, \ldots, S_A^{2^r}$ and $S_B^1, \ldots, S_B^{2^r}$ be independent random draws from distributions D_A and D_B, respectively. Let $S^r = S_A^1 \cap \cdots \cap S_A^{2^r} \cap S_B^1 \cap \cdots \cap S_B^{2^r}$. Then $v_X(M \setminus S^r) \geq (1 - \frac{k}{4^{2^r}})v_X(M)$.*

Proof of Claim VI.10: For every bundle of k items, the probability that player X gets this bundle when allocated the random bundle $M \setminus S^r$ is $\geq 1 - \frac{k}{4^{2^r}}$: For player X not to get the bundle, at least one of its items must be in S^r. By definition of S^r, the probability $\Pr[i \in S^r]$ is equal to $(\sum_{S \ni i} x_1(S))^{2^r} \cdot (\sum_{S \ni i} x_2(S))^{2^r}$, where $\sum_{S \ni i} x_1(S) + x_2(S) \leq 1$ because the solution is feasible for the configuration LP. This means that $\Pr[i \in S^r] \leq \frac{1}{4^{2^r}}$, and by applying the union bound we get that at least one of the items in the bundle are in S^r with probability $\leq \frac{k}{4^{2^r}}$. Applying Lemma VI.4 for MPH-k valuations, player X's expected value for the random bundle $M \setminus S^r$ is thus at least $(1 - \frac{k}{4^{2^r}})v_X(M)$. ∎

By Claim VI.10 the last case achieves expected welfare of at least

$$\frac{1}{2}(v_A(S^r) + v_B(S^r)) + \frac{1}{2}(1 - \frac{k}{4^{2^r}})(v_A(M) + v_B(M)). \tag{10}$$

We can now sum up the contributions from the first case, the intermediate case (9), and the last case (10) weighted by their respective probabilities of λ, λ_q for every intermediate case $0 \leq q < r$, and λ_r. Notice that $\lambda_{r-1} = \lambda_r$. The weighted sum of the last *two* cases is thus at least

$$\begin{aligned}
& \frac{\lambda_{r-1}}{2}(v_A(S^{r-1}) + v_B(S^{r-1})) + \\
& \frac{\lambda_{r-1}}{2}(v_A(S^{r-1} \setminus S^r) + v_B(S^{r-1} \setminus S^r)) \\
+ \quad & \frac{\lambda_r}{2}(v_A(S^r) + v_B(S^r)) + \frac{\lambda_r}{2}(1 - \frac{k}{4^{2^r}})(v_A(M) + v_B(M)) \\
\geq \quad & \frac{\lambda_{r-2}}{2}(v_A(S^{r-1}) + v_B(S^{r-1})) + \frac{1}{2^{r+2}}(1 - \frac{k}{4^{2^r}}) \cdot C \\
= \quad & \frac{\lambda_{r-2}}{2}(v_A(S^{r-1}) + v_B(S^{r-1})) + \Omega(\frac{1}{\log k}) \cdot C.
\end{aligned}$$

Above, the first inequality follows by first observing that $\lambda_r = \lambda_{r-1}$, $\lambda_{r-2} = 2\lambda_{r-1}$, and using subadditivity of $v_A(\cdot), v_B(\cdot)$ to combine the first three terms together. The last equality follows by observing that $k/4^k < 1/2$, and that $r = \lceil \log \log k \rceil$.

Now, we wish to continue by induction and sum the last τ cases for $\tau \leq r + 1$. We claim that the contribution from the last τ cases is at least:

$$\frac{\lambda_{r-\tau}}{2} \cdot (v_A(S^{r-\tau+1}) + v_B(S^{r-\tau+1})) + \Omega(1/\log k) \cdot C.$$

We have already proven the base case: this holds when plugging in $\tau = 2$. Assume for inductive hypothesis that the claim holds for τ, and we now prove it for $\tau + 1$. Then the

sum of the last $\tau+1$ terms is exactly:

$$\frac{\lambda_{r-\tau}}{2}\left(v_A(S^{r-\tau})+v_B(S^{r-\tau})\right)+$$
$$\frac{\lambda_{r-\tau}}{2}\left(v_A(S^{r-\tau}\setminus S^{r-\tau+1})+v_B(S^{r-\tau}\setminus S^{r-\tau+1})\right)+$$
$$\frac{\lambda_{r-\tau}}{2}\cdot\left(v_A(S^{r-\tau+1})+v_B(S^{r-\tau+1})\right)+\Omega(1/\log k)\cdot C \quad \geq$$
$$\frac{\lambda_{r-\tau-1}}{2}\left(v_A(S^{r-\tau})+v_B(S^{r-\tau})\right)+\Omega(1/\log k)\cdot C.$$

The inequality follows by definition of $\lambda_{r-\tau-1}:=2\lambda_{r-\tau}$, and because both $v_A(\cdot)$ and $v_B(\cdot)$ are subadditive. This proves the inductive step. This means that the last $r+1$ cases together contribute exactly (below, $\lambda_{-1}:=2\lambda_0=\lambda$:

$$\frac{\lambda_{-1}}{2}\cdot\left(v_A(S^0)+v_B(S^0)\right)+\Omega(1/\log k)\cdot C.$$

Now, together with the weighted contribution of the first case, which is at least

$$\frac{\lambda}{2}\left(v_A(S_A)+v_B(S_B)+v_A(S_A\setminus S_B)+v_B(S_B\setminus S_A)\right),$$

we get a total expected welfare of at least (recall that $S^0=S_A\cap S_B$):

$$\frac{\lambda}{2}\cdot\left(v_A(S_A)+v_A(S_B)+v_A(S_A\setminus S_B)+v_B(S_B\setminus S_A)\right)$$
$$+\frac{\lambda}{2}\cdot\left(v_A(S_A\cap S_B)+v_B(S_A\cap S_B)\right)+\Omega(1/\log k)\cdot C \quad \geq$$
$$\lambda\left(v_A(S_A)+v_B(S_B)\right)+\Omega(1/\log k)\cdot C \quad =$$
$$(1/2+\Omega(1/\log k))\cdot C.$$

This completes the proof of Theorem VI.9. ∎

Corollary VI.11. *If Alice and Bob have valuations that are pointwise β-approximated by subadditive MPH-k valuations, then the expected welfare of the oblivious rounding scheme is $\geq(\frac{1}{2}+\Omega(\frac{1}{\beta\log k}))\cdot C$.*

F. Integrality Gaps and Hardness

In this section, we briefly derive integrality gaps and communication lower bounds for subadditive MPH-k valuations based on previous constructions. We also show an integrality gap of $\frac{1}{2}$ for MPH-2 valuations that are not subadditive.

Proposition VI.12. *Let $k\geq 2$. The integrality gap of the configuration LP with two subadditive MPH-k bidders is $\frac{1}{2}+\Theta(\frac{1}{\log k})$.*

Proof Sketch: The rounding algorithm presented above witnesses that the integrality gap is $1/2+O(1/\log k)$, and the construction from Section III witnesses that the gap is $1/2+\Omega(1/\log k)$.

Specifically, let $S_1,\ldots,S_t$ be random sets of size $k/2$ that are all subsets of the same $K\subseteq M$ of $|K|=k$. Then with $t:=\sqrt{k}$, $\ell:=\log_2(k)/2$, $\{S_1,\ldots,S_t\}$ is ℓ-sparse with probability $1-1/\text{poly}(k)$. So consider the instance with $f^\ell_{\mathcal{S}}(\cdot)=v_1(\cdot)=v_2(\cdot)$. Then in this case, we know that the best achievable welfare is ℓ for an integral solution. However, the fractional solution which sets $x_i(\overline{S_j})=1/t$ for all t is feasible for the configuration LP, and achieves welfare $2(\ell-1)$. ∎

Proposition VI.13. *The integrality gap of the configuration LP with two (non-subadditive) MPH-2 bidders is $\frac{1}{2}$.*

Proof: Consider 4 items a,b,c,d. Alice has value 1 for bundle $\{a,b\}$ and for bundle $\{c,d\}$ (as well as for any containing bundle); and Bob has value 1 for bundle $\{a,c\}$ and for bundle $\{b,d\}$ (as well as for any containing bundle); all other values are 0. These valuations are MPH-2 since they can be described as the maximum over 2 hypergraphs, each with a positive hyperedge of size 2 corresponding to one of the two desired bundles. The best fractional solution to the configuration LP is $x_1(\{a,b\})=x_1(\{c,d\})=x_2(\{a,c\})=x_2(\{b,d\})=1/2$; one can check that all constraints are satisfied and the objective is 2. The best integral solution however cannot achieve welfare better than 1, completing the proof. ∎

Proposition VI.14. *Let $k\geq 2$. There exists an absolute constant C such that the (randomized) communication required to achieve a $(1/2+C/\log k)$-approximation for* WELFARE-MAXIMIZATION *even when both Alice and Bob are subadditive MPH-k is $\Omega(e^{\sqrt{k}})$.*

Proof: This is a direct application of Theorem V.1, after observing that any subadditive valuation defined on items $K\subseteq M$ (with zero value for all other items) is MPH-$|K|$. ∎

Observe that the above proposition further means that our approximation guarantees for subadditive MPH-k are asymptotically tight among protocols with poly(m) communication whenever $k=\Omega(\log^3 m)$ (tighter calculations could get this down to $\log^{(1+\varepsilon)} m$ for any constant $\varepsilon>0$, if desired). It remains open whether there is an impossibility result $<3/4$ for two subadditive MPH-2 valuations ($3/4$ is implied already by the impossibility for MPH-1 = XOS [DNS10]).

ACKNOWLEDGMENTS

The authors wish to thank Noga Alon for referring us to the notion of ℓ-independent sets. The work of M. Feldman was partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement number 337122, the Israel Science Foundation (grant number 317/17), and the Blavatnik Family Foundation. E. Neyman and S.M. Weinberg are supported by NSF CCF-1717899. I. Talgam-Cohen is supported by the ISRAEL SCIENCE FOUNDATION (grant No. 336/18) and by the Taub Family Foundation.

APPENDIX

Proof of Lemma III.2: It is clear that $f^\ell_{\mathcal{S}}(X)$ is always defined at least once. The only way in which $f^\ell_{\mathcal{S}}(X)$ could

be defined multiple times is if $\sigma_{\mathcal{S}}(X) < \frac{\ell}{2}$ (in which case $f^\ell_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X)$) and $\sigma_{\mathcal{S}}(\overline{X}) < \frac{\ell}{2}$ (in which case $f^\ell_{\mathcal{S}}(X) = \ell - \sigma_{\mathcal{S}}(\overline{X})$). So assume for contradiction that both events hold, and let $X \subseteq \cup_{i=1}^{\ell/2-1} T_i$, and $\overline{X} \subseteq \cup_{i=1}^{\ell/2-1} U_i$, where each $T_i, U_i \in \mathcal{S}$. Note that $\sigma_{\mathcal{S}}(X), \sigma_{\mathcal{S}}(\overline{X}) < \ell/2$ implies that such T_i, U_i must exist. But now consider that we can write $M = X \cup \overline{X}$ as a union of $\leq \ell - 2$ elements of $\mathcal{S}$, contradicting that $\mathcal{S}$ is ℓ-sparse. ∎

Proof of Lemma III.3: We show (1) similarly to [BR11]. First, $\sigma_{\mathcal{S}}(\cdot)$ is monotone because if X is a subset of $\cup_{i\in Y} S_i$, then $X' \subseteq X$ is also a subset of $\cup_{i\in Y} S_i$ (and therefore, any subcollection of $\mathcal{S}$ that covers X also covers X', and $\sigma_{\mathcal{S}}(X') \leq \sigma_{\mathcal{S}}(X)$). Otherwise if X is not covered by $\mathcal{S}$, then $\sigma_{\mathcal{S}}(X) = \max\{\ell, k\}$ and certainly $\sigma_{\mathcal{S}}(X') \leq \sigma_{\mathcal{S}}(X)$. Second, $\sigma_{\mathcal{S}}$ is subadditive: Note that at least one of X, W is not covered by $\mathcal{S}$ if and only if $X \cup W$ is not covered by $\mathcal{S}$, and in this case $\sigma_{\mathcal{S}}(X \cup W) = \max\{\ell, k\} \leq \sigma_{\mathcal{S}}(X) + \sigma_{\mathcal{S}}(W)$. Otherwise, consider the index sets Y, Z witnessing $\sigma_{\mathcal{S}}(X)$ and $\sigma_{\mathcal{S}}(W)$, respectively (that is, $X \subseteq \cup_{i\in Y} S_i$ and $\sigma_{\mathcal{S}}(X) = |Y|$, ditto for W and Z). Then $X \cup W \subseteq \cup_{i\in Y\cup Z} S_i$, and $\sigma_{\mathcal{S}}(X \cup W) \leq |Y| + |Z| = \sigma_{\mathcal{S}}(X) + \sigma_{\mathcal{S}}(W)$.

(2) follows directly from the definition of $f^\ell_{\mathcal{S}}$, and the fact that it is well-defined: (2) clearly holds for any set X for which $f^\ell_{\mathcal{S}}(X)$ is defined in (a), and also for any set X for which $f^\ell_{\mathcal{S}}(X)$ is defined in (b) (as both $f^\ell_{\mathcal{S}}(X)$ and $f^\ell_{\mathcal{S}}(\overline{X})$ are $\ell/2$).

(3) holds because if $\sigma_{\mathcal{S}}(X) < \frac{\ell}{2}$ then $f^\ell_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X)$ by definition. And if $f^\ell_{\mathcal{S}}(X) < \frac{\ell}{2}$, the only way this is possible (given the definition of $f^\ell_{\mathcal{S}}(\cdot)$) is if $f^\ell_{\mathcal{S}}(X)$ is defined in (a) to be $\sigma^\ell_{\mathcal{S}}(X)$.

(4) holds again because if $f^\ell_{\mathcal{S}}(X) > \frac{\ell}{2}$, the only way this is possible (given the definition of $f^\ell_{\mathcal{S}}(\cdot)$) is if $f^\ell_{\mathcal{S}}(X)$ is defined to be $\ell - \sigma_{\mathcal{S}}(\overline{X})$.

(5) holds because by definition of $f^\ell_{\mathcal{S}}(\cdot)$, one of three possible cases must occur: either $f^\ell_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X)$, or $f^\ell_{\mathcal{S}}(X) = \ell - \sigma_{\mathcal{S}}(\overline{X}) \leq \sigma_{\mathcal{S}}(M) - \sigma_{\mathcal{S}}(\overline{X}) \leq \sigma_{\mathcal{S}}(X)$ (using the ℓ-sparsity of $\mathcal{S}$ by which $\sigma_{\mathcal{S}}(M) \geq \ell$ (which holds even if M is not covered by $\mathcal{S}$) and subadditivity of $\sigma_{\mathcal{S}}$), or $f^\ell_{\mathcal{S}}(X) = \ell/2$ and $\sigma_{\mathcal{S}}(X) \geq \ell/2$. ∎

Proof of Corollary III.4: (Monotonicity) Let $X \subseteq T$ and suppose for contradiction that $f^\ell_{\mathcal{S}}(T) < f^\ell_{\mathcal{S}}(X)$. First suppose that $f^\ell_{\mathcal{S}}(T) < \frac{\ell}{2}$. By part (3) of Lemma III.3, we can conclude that $\sigma_{\mathcal{S}}(T) < \frac{\ell}{2}$, and therefore by part (1) (specifically, monotonicity of $\sigma_{\mathcal{S}}(\cdot)$), $\sigma_{\mathcal{S}}(X) < \frac{\ell}{2}$. Thus, by another application of (3), we get: $f^\ell_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X) \leq \sigma_{\mathcal{S}}(T) = f^\ell_{\mathcal{S}}(T)$, a contradiction.

Next suppose that $f^\ell_{\mathcal{S}}(T) \geq \frac{\ell}{2}$. By assumption, this means that $f^\ell_{\mathcal{S}}(X) > \frac{\ell}{2}$, so by part (4) we have $\sigma_{\mathcal{S}}(\overline{X}) = \ell - f^\ell_{\mathcal{S}}(X)$. Since $\overline{T} \subseteq \overline{X}$, by (1) we have $\sigma_{\mathcal{S}}(\overline{T}) \leq \sigma_{\mathcal{S}}(\overline{X}) = \ell - f^\ell_{\mathcal{S}}(X) < \frac{\ell}{2}$. So by applying parts (2) and (3) we have $\ell - f^\ell_{\mathcal{S}}(T) = f^\ell_{\mathcal{S}}(\overline{T}) = \sigma_{\mathcal{S}}(\overline{T}) \leq \ell - f^\ell_{\mathcal{S}}(X)$, implying that $f^\ell_{\mathcal{S}}(X) \leq f^\ell_{\mathcal{S}}(T)$, a contradiction. Therefore, we have a contradiction in both cases, and we must have $f^\ell_{\mathcal{S}}(X) \leq f^\ell_{\mathcal{S}}(T)$.

(Subadditivity) Suppose for contradiction that $f^\ell_{\mathcal{S}}(X \cup T) > f^\ell_{\mathcal{S}}(X) + f^\ell_{\mathcal{S}}(T)$. We first show that $f^\ell_{\mathcal{S}}(X \cup T) > \frac{\ell}{2}$. Indeed, suppose that $f^\ell_{\mathcal{S}}(X \cup T) \leq \frac{\ell}{2}$. Then $f^\ell_{\mathcal{S}}(X)$ and $f^\ell_{\mathcal{S}}(T)$ are both $< \frac{\ell}{2}$, so $f^\ell_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X)$ and $f^\ell_{\mathcal{S}}(T) = \sigma_{\mathcal{S}}(T)$ by (3). So by subadditivity of $\sigma_{\mathcal{S}}(\cdot)$, we have $\sigma_{\mathcal{S}}(X \cup T) \leq f^\ell_{\mathcal{S}}(X) + f^\ell_{\mathcal{S}}(T)$. Note finally that $f^\ell_{\mathcal{S}}(X \cup T) \leq \sigma_{\mathcal{S}}(X \cup T)$ by part (5). Thus, $f^\ell_{\mathcal{S}}(X \cup T) \leq f^\ell_{\mathcal{S}}(X) + f^\ell_{\mathcal{S}}(T)$, a contradiction.

Now assume for contradiction that $f^\ell_{\mathcal{S}}(X \cup T) > \frac{\ell}{2}$. This means that $f^\ell_{\mathcal{S}}(X \cup T) = \ell - \sigma_{\mathcal{S}}(\overline{X \cup T})$ by (4). Observe also that $f^\ell_{\mathcal{S}}(X \cup T) \leq \ell$. Since by assumption, $f^\ell_{\mathcal{S}}(X \cup T) > f^\ell_{\mathcal{S}}(X) + f^\ell_{\mathcal{S}}(T)$, at least one of $f^\ell_{\mathcal{S}}(X)$ and $f^\ell_{\mathcal{S}}(T)$ is $< \frac{\ell}{2}$; without loss of generality, assume that $f^\ell_{\mathcal{S}}(X) < \frac{\ell}{2}$, so $f^\ell_{\mathcal{S}}(X) = \sigma_{\mathcal{S}}(X)$ by (3). Using what we've concluded so far, we may rewrite $f^\ell_{\mathcal{S}}(X \cup T) > f^\ell_{\mathcal{S}}(X) + f^\ell_{\mathcal{S}}(T)$ as $\ell - \sigma_{\mathcal{S}}(\overline{X \cup T}) > \sigma_{\mathcal{S}}(X) + f^\ell_{\mathcal{S}}(T)$, i.e.,

$$\sigma_{\mathcal{S}}(\overline{X \cup T}) + \sigma_{\mathcal{S}}(X) + f^\ell_{\mathcal{S}}(T) < \ell. \tag{11}$$

We have that $\overline{T} \subseteq \overline{X \cup T} \cup X$ (De Morgan), and so $\sigma_{\mathcal{S}}(\overline{X \cup T}) + \sigma_{\mathcal{S}}(X) \geq \sigma_{\mathcal{S}}(\overline{T})$ by subadditivity of $\sigma_{\mathcal{S}}(\cdot)$. Plugging this observation into Equation (11), we get $\sigma_{\mathcal{S}}(\overline{T}) + f^\ell_{\mathcal{S}}(T) < \ell$. But by parts (1) and (5) of Lemma III.3, $f^\ell_{\mathcal{S}}(X) = \ell - f^\ell_{\mathcal{S}}(\overline{X}) \geq \ell - \sigma_{\mathcal{S}}(\overline{X})$ for all X, a contradiction. We therefore conclude that $f^\ell_{\mathcal{S}}(\cdot)$ must be subadditive, as we have derived contradictions whether $f^\ell_{\mathcal{S}}(X \cup T) \leq \frac{\ell}{2}$ (first paragraph) or $f^\ell_{\mathcal{S}}(X \cup T) > \frac{\ell}{2}$ (just now). ∎

Proof of Lemma IV.3: Consider the following randomized construction of $\mathcal{S}$: For each $i \in [k], j \in M$ independently, flip a fair coin. If heads, put $j \in S_i$. Otherwise, don't. We wish to show that the probability that $\mathcal{S}$ is ℓ-independent is non-zero. We'll again use $S^1_i := S_i$, and $S^0_i := \overline{S_i}$.

So fix any set Y of ℓ indices, and any vector $y \in \{0,1\}^\ell$ (for simplicity of notation, index these ℓ bits using the indices of Y). We wish to consider the probability that $\cup_{i\in Y} S^{y_i}_i = M$. If there exists a single Y, y such that $\cup_{i\in Y} S^{y_i}_i = M$, then $\mathcal{S}$ is *not* ℓ-independent. But if for all Y, y $\cup_{i\in Y} S^{y_i}_i \neq M$, then $\mathcal{S}$ *is* ℓ-independent. So we simply wish to analyze the probability that this occurs for a fixed Y, y and take a union bound.

For a fixed Y, y, observe that each element $j \in M$ is in each $S^{y_i}_i$ independently with probability $1/2$. So the probability that j is in *some* $S^{y_i}_i$ is just $1 - 1/2^\ell$. Moreover, these events are independent across items j, so $\cup_{i\in Y} S^{y_i}_i = M$ with probability $(1 - 1/2^\ell)^m$. Now we wish to take a union bound over all 2^ℓ choices of y times $\binom{k}{\ell}$ choices of Y to get that the probability that $\mathcal{S}$ is *not* ℓ-independent is upper bounded by (the final two steps use our choice of $\ell = \log_2(m) - \log_2(x)$ and $k = e^{x/\ell}$):

$$\begin{aligned} 2^\ell \binom{k}{\ell} \left(1 - \frac{1}{2^\ell}\right)^m &< \frac{2^\ell \cdot k^\ell}{\ell!} \left(1 - \frac{1}{2^\ell}\right)^m \\ &< k^\ell \exp\left(\frac{-m}{2^\ell}\right) \\ &= \exp\left(\ell \cdot \ln(k) - x\right) = 1. \end{aligned}$$

As the probability that $\mathcal{S}$ is not ℓ-independent is < 1, we are guaranteed the existence of some $\mathcal{S}$ that is ℓ-independent

by the probabilistic method. ∎

Proposition A.1. *Every monotone and subadditive function over a set of m items is MPH-$\lceil \frac{m}{2} \rceil$.*

Proof: Let f be a monotone subadditive function. For every set $S \subseteq [m]$ we construct a positive hypergraph G_S with weights w_S, as follows: Let S' be an arbitrary subset of S of size $\min(|S|, \lceil \frac{m}{2} \rceil)$. Set $w_S(S') = f(S')$ and $w_S(S \setminus S') = f(S) - f(S')$. All other hyperedges have weight 0. Observe that in our construction, every hyperedge with a non-zero weight has size at most $\lceil \frac{m}{2} \rceil$. For every set T, let $f_S(T)$ denote the value of T in G_S; so $f_S(\cdot)$ is PH-$\lceil \frac{m}{2} \rceil$. We argue that $f(T) = \max_S \{f_S(T)\}$ and so $f(\cdot)$ is MPH-$\lceil \frac{m}{2} \rceil$. It is sufficient to show that for every set S the following two properties hold: (1) $f_S(S) = f(S)$; (2) $f_S(T) \leq f(T)$ for every set T.

The first property holds since $f_S(S) = w_S(S') + w_S(S \setminus S') = f(S') + f(S) - f(S') = f(S)$. The proof of the second property is divided into four cases:

1) If $S \subseteq T$ then $f_S(T) = w(S') + w(S \setminus S') = f(S) \leq f(T)$.
2) If $S \not\subseteq T$ and $S' \subseteq T$ then $S \setminus S' \not\subseteq T$, therefore $f_S(T) = w(S') = f(S') \leq f(T)$.
3) If $S \not\subseteq T$ and $S \setminus S' \subseteq T$ then $S' \not\subseteq T$, therefore $f_S(T) = w(S \setminus S') = f(S) - f(S') \leq f(S \setminus S') \leq f(T)$, where the first inequality is due to the subadditivity of f.
4) If $S \setminus S', S' \not\subseteq T$ then $f_S(T) = 0 \leq f(T)$.

∎

Proof of Lemma V.2: Suppose that for some i, $|X_i \cap Y_i| = 2$. Consider the allocation that awards Y_i to Alice and $\overline{Y_i}$ to Bob. Then Bob clearly has value $\ell - 1$, as $f^\ell_{\mathcal{Y}}(\overline{Y_i}) = \ell - 1$. Also, as $|Y_i| = |X_i| = m/2 + 1$ and $|X_i \cap Y_i| = 2$, we necessarily have $X_i \cup Y_i = M$. This implies that $Y_i \supseteq \overline{X_i}$, and therefore $v^\ell_{\mathcal{X}}(Y_i) \geq v_{\mathcal{X}}(\overline{X_i}) = \ell - 1$. So welfare $2\ell - 2$ is achievable (and again, optimal, as no bidder can achieve value ℓ without receiving all of M).

Now suppose that for all i, $|X_i \triangle Y_i| = 2$, and further suppose for contradiction that total welfare $> \ell + 1$ is achievable, by giving S to Alice and $\overline{S}$ to Bob. Then one of the players (without loss of generality, say it is Bob) has value $> \frac{\ell}{2}$, so it must have been the case that $f^\ell_{\mathcal{Y}}(\overline{S})$ was defined to be $\ell - \sigma_{\mathcal{Y}}(S)$, and $\sigma_{\mathcal{Y}}(S) < \frac{\ell}{2}$.

Now, observe that because $|X_i \triangle Y_i| = 2$ for all i that $\sigma_{\mathcal{X}}(S) \leq \sigma_{\mathcal{Y}}(S) + 1$. Indeed, let $S = Y_{i_1} \cup \cdots \cup Y_{i_{\sigma_{\mathcal{Y}}(S)}}$. Then there is exactly one element in Y_{i_j} that is not also in X_{i_j}, and we therefore conclude that $|S \setminus (X_{i_1} \cup \cdots \cup X_{i_{\sigma_{\mathcal{Y}}(S)}})| \leq \sigma_{\mathcal{Y}}(S) < \ell/2$. By criterion (4) of ℓ-compatibility, there is some X_i that contains all of these elements, and so $S \subseteq X_i \cup X_{i_1} \cup \cdots \cup X_{i_{\sigma_B(S)}}$, witnessing that $\sigma_{\mathcal{X}}(S) \leq \sigma_{\mathcal{Y}}(S) + 1$.

Finally, if $\sigma_{\mathcal{X}}(S) < \frac{\ell}{2}$ then $f^\ell_{\mathcal{X}}(S) = \sigma_{\mathcal{X}}(S) \leq \sigma_{\mathcal{Y}}(S) + 1$, so the total welfare is at most $\ell + 1$, a contradiction. Otherwise, $\sigma_{\mathcal{X}}(S) = \frac{\ell}{2}$, which we claim implies that $f^\ell_{\mathcal{X}}(S) \leq \ell/2$. Indeed, if $f^\ell_{\mathcal{X}}(S) > \ell/2$, then it is because $\sigma_{\mathcal{X}}(\overline{S}) < \ell/2$. But then we would have $\sigma_{\mathcal{X}}(S) + \sigma_{\mathcal{X}}(\overline{S}) < \ell$, implying a cover of M with $< \ell$ and contradicting ℓ-sparsity of $\mathcal{X}$. Observe that in both cases we may conclude that $f^\ell_{\mathcal{X}}(S) \leq \sigma^\ell_{\mathcal{X}}(S)$.

Now we may conclude that the total welfare is $f^\ell_{\mathcal{X}}(S) + f^\ell_{\mathcal{Y}}(\overline{S}) \leq \sigma_{\mathcal{X}}(S) + \ell - \sigma_{\mathcal{Y}}(S) \leq \ell + 1$, again a contradiction. We have now reached a contradiction from all branches starting from the assumption that welfare $> \ell + 1$ is achievable, and may now conclude that the maximum possible welfare is indeed $\leq \ell + 1$, as desired. ∎

REFERENCES

[ABDR12] Ittai Abraham, Moshe Babaioff, Shaddin Dughmi, and Tim Roughgarden. Combinatorial auctions with restricted complements. In *13th ACM Conference on Electronic Commerce (EC)*, 2012.

[ABKS17] Itai Ashlagi, Mark Braverman, Yash Kanoria, and Peng Shi. Communication requirements and informative signaling in matching markets. In *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*, page 263, 2017.

[Alo86] Noga Alon. Explicit construction of exponential sized families of k-independent sets. *Discrete Mathematics*, 58:191–193, 1986.

[ANRW15] Noga Alon, Noam Nisan, Ran Raz, and Omri Weinstein. Welfare maximization with limited interaction. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 1499–1512, 2015.

[Ass17] Sepehr Assadi. Combinatorial auctions do need modest interaction. In *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*, pages 145–162, 2017.

[BBK+16] Joshua Brody, Harry Buhrman, Michal Koucký, Bruno Loff, Florian Speelman, and Nikolai K. Vereshchagin. Towards a reverse Newman's theorem in interactive information complexity. *Algorithmica*, 76(3):749–781, 2016.

[BDF+10] David Buchfuhrer, Shaddin Dughmi, Hu Fu, Robert Kleinberg, Elchanan Mossel, Christos H. Papadimitriou, Michael Schapira, Yaron Singer, and Christopher Umans. Inapproximability for VCG-Based Combinatorial Auctions. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2010.

[BDF+12] Ashwinkumar Badanidiyuru, Shahar Dobzinski, Hu Fu, Robert Kleinberg, Noam Nisan, and Tim Roughgarden. Sketching valuation functions. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 1025–1035, 2012.

[BDN19] Yakov Babichenko, Shahar Dobzinski, and Noam Nisan. The communication complexity of local search. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019*, 2019.

[BG14] Mark Braverman and Ankit Garg. Public vs private coin in bounded-round information. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I*, pages 502–513, 2014.

[BGPW13] Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. From information to exact communication. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC)*, pages 151–160, 2013.

[BJKS04] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *J. Comput. Syst. Sci.*, 68(4):702–732, 2004.

[BMW18] Mark Braverman, Jieming Mao, and S. Matthew Weinberg. On simultaneous two-player combinatorial auctions. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 2256–2273, 2018.

[BN07] Liad Blumrosen and Noam Nisan. Combinatorial auctions. In Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, editors, *Algorithmic Game Theory*, chapter 11. Cambridge University Press, 2007.

[BN19] Simina Brânzei and Noam Nisan. Communication complexity of cake cutting. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019.*, page 525, 2019.

[BNS07] Liad Blumrosen, Noam Nisan, and Ilya Segal. Auctions with severely bounded communication. *J. Artif. Intell. Res.*, 28:233–266, 2007.

[BR11] Kshipra Bhawalkar and Tim Roughgarden. Welfare guarantees for combinatorial auctions with item bidding. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 700–709, 2011.

[BR14] Mark Braverman and Anup Rao. Information equals amortized communication. *IEEE Trans. Information Theory*, 60(10):6058–6069, 2014.

[BR17] Yakov Babichenko and Aviad Rubinstein. Communication complexity of approximate Nash equilibria. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 878–889, 2017.

[Bra12] Mark Braverman. Interactive information complexity. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 505–524, 2012.

[BSS10] Dave Buchfuhrer, Michael Schapira, and Yaron Singer. Computation and incentives in combinatorial public projects. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, EC '10, pages 33–42, New York, NY, USA, 2010. ACM.

[CKST16] George Christodoulou, Annamária Kovács, Alkmini Sgouritsa, and Bo Tang. Tight bounds for the price of anarchy of simultaneous first-price auctions. *ACM Trans. Economics and Comput.*, 4(2):9:1–9:33, 2016.

[CM16] Shuchi Chawla and J. Benjamin Miller. Mechanism design for subadditive agents via an ex ante relaxation. In *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, Maastricht, The Netherlands, July 24-28, 2016*, pages 579–596, 2016.

[CS05] Vincent Conitzer and Tuomas Sandholm. Communication complexity of common voting rules. In *Proceedings 6th ACM Conference on Electronic Commerce (EC-2005), Vancouver, BC, Canada, June 5-8, 2005*, pages 78–87, 2005.

[CZ17] Yang Cai and Mingfei Zhao. Simple mechanisms for subadditive buyers via duality. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 170–183, 2017.

[DMSW15] Nikhil R. Devanur, Jamie Morgenstern, Vasilis Syrgkanis, and S. Matthew Weinberg. Simple auctions with simple strategies. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, pages 305–322, 2015.

[DN15] Shahar Dobzinski and Noam Nisan. Multi-unit auctions: Beyond Roberts. *J. Economic Theory*, 156:14–44, 2015.

[DNS10] Shahar Dobzinski, Noam Nisan, and Michael Schapira. Approximation algorithms for combinatorial auctions with complement-free bidders. *Math. Oper. Res.*, 35(1):1–13, 2010. Preliminary version in STOC 2005.

[Dob07] Shahar Dobzinski. Two randomized mechanisms for combinatorial auctions. In *Proceedings of the 10th International Workshop on Approximation and the 11th International Workshop on Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 89–103, 2007.

[Dob11] Shahar Dobzinski. An impossibility result for truthful combinatorial auctions with submodular valuations. In *Proceedings of the 43rd ACM Symposium on Theory of Computing (STOC)*, 2011.

[Dob16a] Shahar Dobzinski. Breaking the logarithmic barrier for truthful combinatorial auctions with submodular bidders. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2016, pages 940–948, New York, NY, USA, 2016. ACM.

[Dob16b] Shahar Dobzinski. Computational efficiency requires simple taxation. In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)*, pages 209–218, 2016.

[DSS15] Amit Daniely, Michael Schapira, and Gal Shahaf. Inapproximability of truthful mechanisms via generalizations of the VC dimension. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 401–408, 2015.

[DV12a] Shahar Dobzinski and Jan Vondrák. The computational complexity of truthfulness in combinatorial auctions. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, 2012.

[DV12b] Shahar Dobzinski and Jan Vondrák. From query complexity to computational complexity. In *Proceedings of the 44th Symposium on Theory of Computing (STOC)*, 2012.

[DV13] Shahar Dobzinski and Jan Vondrák. Communication complexity of combinatorial auctions with submodular valuations. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1205–1215, 2013.

[Fei06] Uriel Feige. On maximizing welfare when utility functions are subadditive. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 41–50, 2006.

[Fei09] Uriel Feige. On maximizing welfare when utility functions are subadditive. *SIAM J. Comput.*, 39(1):122–142, 2009. Preliminary version in STOC 2006.

[FFGL13] Michal Feldman, Hu Fu, Nick Gravin, and Brendan Lucier. Simultaneous auctions are (almost) efficient. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 201–210, 2013.

[FFI+15] Uriel Feige, Michal Feldman, Nicole Immorlica, Rani Izsak, Brendan Lucier, and Vasilis Syrgkanis. A unifying hierarchy of valuations with complements and substitutes. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, pages 872–878, 2015.

[FFT16] Uriel Feige, Michal Feldman, and Inbal Talgam-Cohen. Oblivious rounding and the integrality gap. In *Proceedings of the 19th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, pages 8:1–8:23, 2016.

[FV10] Uriel Feige and Jan Vondrák. The submodular welfare problem with demand queries. *Theory of Computing*, 6(1):247–290, 2010.

[GNOR15] Yannai A. Gonczarowski, Noam Nisan, Rafail Ostrovsky, and Will Rosenbaum. A stable marriage requires communication. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 1003–1017, 2015.

[GR18] Mika Göös and Aviad Rubinstein. Near-optimal communication lower bounds for approximate nash equilibria. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 397–403, 2018.

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

[KS73] Daniel J. Kleitman and Joel Spencer. Families of k-independent sets. *Discrete Mathematics*, 6(3):255–262, 1973.

[LS11] Ron Lavi and Chaitanya Swamy. Truthful and near-optimal mechanism design via linear programming. *J. ACM*, 58(6):25:1–25:24, 2011.

[MSV08] Vahab S. Mirrokni, Michael Schapira, and Jan Vondrák. Tight information-theoretic lower bounds for welfare maximization in combinatorial auctions. In *Proceedings 9th ACM Conference on Electronic Commerce (EC-2008), Chicago, IL, USA, June 8-12, 2008*, pages 70–77, 2008.

[New91] Ilan Newman. Private vs. common random bits in communication complexity. *Inf. Process. Lett.*, 39(2):67–71, 1991.

[Nis00] Noam Nisan. Bidding and allocation in combinatorial auctions. In *EC*, pages 1–12, 2000.

[NR01] Noam Nisan and Amir Ronen. Algorithmic mechanism design. *Games and Economic Behavior (GEB)*, 35:166–196, 2001.

[NS06] Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *Journal of Economic Theory (JET)*, 129:192–224, 2006.

[PSS08] Christos H. Papadimitriou, Michael Schapira, and Yaron Singer. On the hardness of being truthful. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2008.

[Rou14] Tim Roughgarden. Barriers to near-optimal equilibria. In *Proceedings of the 55th Symposium on Foundations of Computer Science (FOCS)*, pages 71–80, 2014.

[RS17] Aviad Rubinstein and Sahil Singla. Combinatorial prophet inequalities. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 1671–1687, 2017.

[RW15] Aviad Rubinstein and S. Matthew Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. In *Proceedings of the 16th ACM Conference on Electronic Commerce*, 2015.

[RW16] Tim Roughgarden and Omri Weinstein. On the communication complexity of approximate fixed points. In *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 229–238, 2016.

[Von08] Jan Vondrák. Optimal approximation for the submodular welfare problem in the value oracle model. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 67–74, 2008.