Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2008-042 July 14, 2008

Knowledge Benchmarks in Adversarial Mechanism Design (Part I) and Implementation in Surviving Strategies (Part I)

Jing Chen and Silvio Micali

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

Knowledge Benchmarks in Adversarial Mechanism Design and Implementation in Surviving Strategies (Part I)

Jing Chen and Silvio Micali

July 13, 2008

Abstract

We put forward new benchmarks and solution concepts for *Adversarial Mechanism Design*, as defined by [MV07.a], and we exemplify them in the case of truly combinatorial auctions.

We benchmark the *combined performance* (the sum of the auction's efficiency and revenue) of a truly combinatorial auction against a very relevant but private knowledge of the players: essentially, *the maximum revenue that the best informed player could guarantee if he were the seller.* (I.e., by offering each other player a subset of the goods for a take-it-or-leave-it price.)

We achieve this natural benchmark within a factor of 2, by means of a new and probabilistic auction mechanism, in *knowingly* Σ^1/Σ^2 *surviving strategies.* That is, the above performance of our mechanism is guaranteed in any rational play, independent of any possible beliefs of the players. Indeed, our performance guarantee holds for *any possible choice of strategies,* so long as each player chooses a strategy among those surviving iterated elimination of knowingly dominated strategies.

Our mechanism is extremely robust. Namely, its performance guarantees hold even if all but one of the players collude (together or in separate groups) in any possible but reasonable way. Essentially, the only restriction for the collective utility function of a collusive subset S of the players is the following: the collective utility increases when one member of S is allocated a subset of the goods "individually better" for him and/or his "individual price" is smaller, while the allocations and prices of all other members of S stay the same.

Our results improve on the yet unpublished ones of [MV07.b]. The second part of this paper, dealing with a more aggressive benchmark (essentially, the maximum welfare privately known to the players) is forthcoming.

1 A New Approach to Adversarial Mechanism Design

Truly Combinatorial Auctions. In a combinatorial auction, a set of goods are going to be divided into subsets and sold to several players. Each player i has a valuation for these goods —a mapping from subsets of goods to non-negative reals— which is called his *true valuation*, and denoted by TV_i. The *profile* (i.e., a vector indexed by the players) TV is called the true valuation profile of the auction. Following the terms used in [MV07.a], in this paper, we focus on *truly combinatorial auctions*. That is, auctions which have completely no restriction on how they are formed. In particular, in truly combinatorial auctions, for any players i and any subsets of goods S, T such that $S \neq T$, knowing the value of $TV_i(S)$ gives out no information about the value of $TV_i(T)$. The truly combinatorial auction is the most general form of combinatorial auctions, and it is easy to see that if a statement holds with respect to such a general form, it must also hold with respect to any special form of combinatorial auctions, e.g., auctions where the true valuations are sub-modular or additive or single-minded,[1] or where the supply of goods is unlimited, etc.

Combining Efficiency and Revenue. In Auction Mechanism Design, the goal is to design a particular mechanism such that, when executed in a combinatorial auction with a set N of n players ($N = \{1, \ldots, n\}$) and a set G of m goods ($G = \{g_1, \ldots, g_m\}$), produces an outcome Ω —indicating how to sell the goods— which achieves some pre-specified property $\mathbb{P}$. The outcome consists of: (1) an *allocation* A, that is, a partition of G in to $n+1$ subsets, $A = (A_0, A_1, \ldots, A_n)$, and (2) a *price profile* P, that is, a profile of real numbers. We refer to A_0 as the set of unallocated goods, and for each player i, we refer to A_i as the set of goods allocated to i and to P_i as the price of i. The property $\mathbb{P}$ is typically either to maximize the *social welfare*, $SW(\Omega, TV)$, which is defined as $\sum_i TV_i(A_i)$, or to maximize the *revenue*, $REV(\Omega)$, which is defined as $\sum_i P_i$. Towards finding a uniform goal which is sited beneath these goals, we find it natural to consider that *the sum of the social welfare and the revenue* be taken care of when designing new mechanisms. Indeed, if a mechanism achieves a "large" value for the sum, then at least one of the social welfare and the revenue is also "large". Accordingly, in this paper, we consider maximizing $SW(\Omega, TV) + REV(\Omega)$ to be the goal of mechanism design, and

[1] *sub-modular* means that $TV_i(S \cup T) \leq TV_i(S) + TV_i(T)$ for any player i and any subsets S, T of the goods; *additive* means that $TV_i(S) = TV_i(g_1) + \cdots + TV_i(g_k)$ whenever $S = \{g_1, \ldots, g_k\}$; and *single-minded* means that for any player i, there exists a subset of goods S_i and a non-negative real v_i such that for any subset of goods T, $TV_i(T) = v_i$ whenever $T \supseteq S_i$, and $TV_i(T) = 0$ otherwise.

talk about the value of the sum as a *combined performance* of a mechanism.[2] It is easy to understand that when the mechanism and/or the behavior of the players are probabilistic, by the "sum" we mean the sum of the *expected* social welfare and the *expected* revenue. In particular, we aim at answering the following question:

What is the combined performance achievable in truly combinatorial auctions?

Adversarial Mechanism Design. We answer the above question not only in the classical auction setting where each player is assumed to be independent and rational, and acts only to maximize his own utility (i.e., his true valuation for the subset of goods sold to him minus his price), but we answer it in more general and adversarial settings: the players may collude arbitrarily to form one or more collusion sets, and members of a collusion set can secretly coordinate their actions in the auction to maximize a secret and universal collective utility for the whole set. Such an *Adversarial Mechanism Design* has been put forward by [MV07.a]. Here we further explore this direction, and develop a new approach to deal with this problem. Notice that it is the *adversarial setting* that makes our approach different and important, because otherwise the VCG mechanism already achieves the maximum social welfare.

Knowledge-Based Adversarial Mechanism Design. The results of [MV07.a] imply that, *even in the traditional setting,* where all players are rational and independent, guaranteeing more revenue than $\mathbb{MSW}_{-\star}/c_{n,m}$ in a truly combinatorial auction requires either

(1) Take advantage of some *special knowledge* about the players true valuations, or

(2) Adopt a solution concept *weaker than dominant strategies.*

In this paper we choose "both way outs." But we do so *within* adversarial mechanism design.

Assume that a government wants to sell 8 licences, respectively for the "upcoming" Web 11.0 through 18.0, to 8 wireless companies —AT&T′, T-Mobile′, Verizon′, Sprint′, Mannesmann′, China-Mobile′, Orange′, and Telecom-Italia′— so as to maximize either revenue (which is already hard when all players bid independently!), or the sum of revenue and efficiency (which is hard when some of the players may collude). How can this be done beating the cited upperbounds of [MV07.a]?

[2]It worths mentioning that although in many mechanisms we met before the social welfare is an upperbound for the revenue, this is by no means a requirement, and is particularly not necessarily true in our result. Therefore we do not want the readers to confuse the combined performance of a mechanism with twice of the revenue generated from the auction.

The easy way out to assume (kindly to ourselves!) that the government can retain the services of an individual well informed about highly-future technologies and in particular the values attributed by each of 8 wireless companies to each of the possible 256 subsets of the licences. (In this case: congratulations! Just call me at 1-800-YOU-FOOL. I have great rates!)

But in mechanism design in its purest form, all knowledge lies with the players themselves. And in the spirit of Axiom 1 of Adversarial Mechanism Design we should reject the convenient temptation of assuming that some magical knowledge is available to the designer. We instead assume that the *players do indeed have useful knowledge,* and that the problem consists of designing a mechanism capable of successfully capitalizing on this proprietary knowledge.

What should "useful knowledge" be in our setting? Because the government's wish is to generate revenue from selling the licenses to the wireless companies, useful knowledge naturally is the revenue that each of the wireless companies could generate if it were to sell the licenses to its peers in a personalized sale: that is, by offering a separate subset of the licenses to each of the other companies in a take-it-or-leave it price. We then take the position that

> *Any of the 8 wireless players could do a better job in selling the 8 licenses to its other 7 peers than the government could in selling the licenses himself to all 8 players.*

Accordingly, a reasonable benchmark for the government is the minimum revenue that one of the 8 wireless players could guarantee. A better benchmark, could be the revenue that a randomly selected company among the 8 ones can guarantee. A better yet benchmark is *the maximum revenue that one of 8 wireless companies could guarantee.* This is indeed the benchmark we adopt.

One of our results is that the sum of efficiency and revenue can achieve this benchmark within a factor of 2, without assuming any special knowledge whatsoever for the auction mechanism, even in a adversarial setting with a general collusion model.

As for revenue alone, our benchmark is the second highest revenue known to the 8 wireless players. We show that this benchmark too can be achieve within a factor of 2, in an adversarial setting with the same collusion model.

The only restriction to the utility function of a collusive set is that, keeping fixed the allocations and the prices of all but one member of the set, the set's collective utility increases if this other player is given a better (for him individually) subset of goods or a lower price.

Let us now sketch our solution concept. As anticipated, we do not achieve our benchmarks in dominant-strategies. But we do achieve them by means of a novel notion: *implementation in surviving* *strategies.* By this we mean that we guarantee our properties in any play of our mechanism in which every strategy (of an independent player or of a every collusive set of players) is chosen among those surviving the iterated elimination of dominated strategies.

We believe this to be a notion of independent interest. In a sense, our mechanism induces a game that is dominant-solvable. More precisely, we do not prove the iterated elimination of dominated strategies leaves a single strategy for each (generalized) player. But this lack of uniqueness does not matter, because we prove that *our desired properties are achieved no matter which surviving strategy each (generalized) player chooses for himself!*

2 Our Knowledge Benchmark

As already said, each combinatorial auction intrinsically has a true-valuation profile TV. Without loss of generality, we now let it intrinsically have also another "external-knowledge" profile K.

Definition 1. *The* external knowledge *of a combinatorial auction is a profile K such that for all i $TV_{-i} \in K_i \subseteq \mathbb{V}_{-i}$, where $\mathbb{V}$ denotes the set of all possible valuation profiles.*

In essence, while TV_i is i's true valuation, K_i represents i's knowledge about the other players' true valuations: that is, from i's point of view, K_i is the set of all possible candidates for TV_{-i}. Such knowledge is truthful in the sense that the actual sub-profile TV_{-i} is among such candidates, but i has no further information about it. Notice that endowing each combinatorial auction with the external knowledge profile is indeed without loss of generality, because "in the worst case" K_i could coincide with $\mathbb{V}_{-i}$, in which case i has no external knowledge at all.

We stress that *no information about K_i is known to the auction designer.* For simplicity, in this version of our paper we actually assume that K_i is totally private to i. Our results however continue to hold under a more general definition of K. (In particular, players may know also information about each other's knowledge; and certain true-valuation information can be common knowledge.)

Having made external knowledge intrinsic to combinatorial auctions, we can now consider *knowledge benchmarks*, that is functions mapping every external knowledge (sub)profile to a non-negative number. Towards defining our knowledge benchmark, we first put forwards a simple auxiliary concept.

Definition 2. (Feasible and Canonical Outcomes) *We say that an outcome (A, P) of a combinatorial auction is* canonical for a subset of players S *if, for all players i: (1) P_i is an integer≥ 0, (2) $P_i = 0$ whenever $A_i = \emptyset$, and (3) $A_i = \emptyset$ whenever $i \notin S$.*

For any subset of players S and any valuation sub-profile V_S, we say that an outcome (A, P) is V_S-feasible *if it is canonical for S and $P_j \leq V_j(A_j)$ whenever $A_j \neq \emptyset$.*

Notice that, for any player i, an outcome (A, P) that is TV_{-i}-feasible essentially consists of a way of selling the goods to the players in $-i$ that generates revenue equal to $\sum_j P_j$: namely, offer to each player j to either (1) buy the subset of goods A_j for price P_j, or (2) receive nothing and pay nothing. Indeed, since he receives positive utility, each player j prefers to accept such an offer to buy.

Let us now define in three steps our knowledge benchmark $\mathbb{MEW}$.

Definition 3. *For a combinatorial auction and any player i, we define $F(K_i)$ to be the set of outcomes V-feasible for every $V \in K_i$; and we define the* maximum external welfare *known to i to be*

$$\mathbb{MEW}_i(K_i) = \max_{\Omega \in F(K_i)} REV(\Omega).$$

Finally, we define $\mathbb{MEW}$*, the* maximum external welfare *benchmark, as follows: for all external-knowledge sub-profiles K_S,*

$$\mathbb{MEW}(K_S) = \max_{i \in S} \mathbb{MEW}_i(K_i).$$

Notice that $\mathbb{MEW}_i(K_i)$ indeed represents the maximum revenue that i could guarantee if he were in charge to sell the goods to the other players by making each one of them a take-it-or-leave-it offer for some subset of the goods.

Recall that, as for all benchmarks in Adversarial Mechanism Design, in order to measure the performance of a combinatorial-auction mechanism, $\mathbb{MEW}$ will be actually evaluated only on the external-knowledge subprofile of the independent players.

Finally, notice that $\mathbb{MEW}$ is "player-monotone" —see [MV07.a]. This implies that the value of $\mathbb{MEW}$ cannot but increase with the number of independent players.

3 Our Collusion Model

The collusion model of this paper envisages a world in which collusion is "illegal." (Forthcoming papers will consider different worlds, with some benefits too.) Accordingly, our mechanisms do not make any concessions to collusive players, nor envisage special bids for collusive players. Thus there is no possibility of proving collusion within the mechanism: *we* distinguish between independent and collusive players in our analysis, but *our mechanisms* have no idea of which players are collusive or independent.

The illegality of collusion, of course, has never stopped players form colluding in auctions, or in any other game, and this trend will likely continue. Our mechanisms therefore must be "collusion-resilient." The illegality of collusion might however have some positive influence on capping the number of collusive players. In turn, this should induce a higher value of our benchmark, because $\mathbb{MEW}$ is evaluated on the external knowledge of the *independent* players and it is *player monotone.* But, while in practice collusive players may be few, in our analysis we do not count on the number of collusive players to be small: we achieve our benchmark no matter how many collusive players there are, but such achievement is meaningful only if at least one independent player exists.

Besides assuming no restrictions on the number of collusive players, we also do not assume any restrictions on the number of collusive sets, nor on the capability of the players (in the same collusive set) to coordinate themselves. In particular, they are free to enter binding agreements with each other on how to conduct themselves in an auction. But we do not rely on the perfect coordination inside on collusive set either! In essence, we aim at achieving our benchmark no matter how players collude. As we clarify in Section 3.3, however, we do make a technical restriction on collusive sets. Such restriction is in particular needed to prevent our collusive players from actually being *irrational.* (Were we dealing with irrational players, we would be facing the worst setting, as defined by [MV07.a], rather than the adversarial one. And guaranteeing any property in the worst setting, although occasionally possible —as indeed proved in [MV07.a]— it is much harder, and tipically impossible.)

3.1 The Static Model (and a Hybrid Presentation)

For simplicity, in this version of our paper we formalize a *static* collusion model. In essence, we envisage that every player starts with his individual true valuation and external knowledge and then before the mechanism starts, by the effect of some magical (but adversarial!) power, some of the players are *suddenly* partitioned into disjoint and fixed collusive sets.

As we shall see, formalizing this model and its matching solution concepts requires some work. This model is very adequate for normal-form mechanisms, where the players act once and "simultaneously." In virtue of the *revelation principle*, our mechanism of course has a normal-form version. However, it also has extensive-form versions that enjoy additional advantages. (Indeed, the *revelation principle* not withstanding, extensive-form mechanisms can enjoy properties denied to normal-form ones. One of these properties is perfect privacy —see [ILM05].) The version of our mechanism presented here is extensive-form, in particular, *public extensive-form.* In a truly combinatorial auction with n players and m goods, an auction mechanism M of public extensive-form is played in k stages, where k is an integer greater than 1. At each stage j, each player i publicly announces a string x_i^j simultaneously with the other players. Then M is evaluated on the profile x, where x_i is the sequence $x_i^1, \ldots, x_i^k$, so as to produce a final outcome. Although we do not have the time to elaborate on the advantages gained from such an extensive-form mechanism over its normal-form counterpart, we shall do so in a subsequent version of our paper.

In that same version we shall also argue that the usage of extensive-form mechanisms justifies considering a much more difficult *dynamic* collusion model. In such a model, collusive sets may grow or change during the running of the mechanism, based on the actions taken in the actual execution, and the knowledge which may be collectively gained when one more player joins a collusive set. This model is ideal for a rigorous and general study of "collusive players formation" provides the "adversary" with more power, whenever the players' strategies and the mechanism itself are probabilistic. Fortunately, our mechanism is robust in this model too.

The dynamic model however is much harder to define. For instance, since collusive players change over time, it becomes less clear what the independent players are, and thus how our benchmark should be evaluated. Accordingly, in this version of our paper we are satisfied with an hybrid presentation. Namely we present our mechanism in extended form and prove its properties under the static collusion model. (We shall "rectify this situation" soon.)

3.2 Collusion Systems

As preannounced, in a combinatorial auction each player i has an individual true valuation TV_i, an individual external knowledge K_i, and an individual utility function u_i, namely: for any outcome (A, P), $u_i((A, P), TV) = TV_i(A_i) - P_i$. Let us now explain the additions brought by a *collusion system* to this initial landscape.

To begin with, a collusion system specifies a partition $\mathbb{C}$ of the players into mutually disjoint sets. Relative to this partition, a player is said to be independent if he belongs to a subset in $\mathbb{C}$ containing only him. All other players are are termed collusive, and the subsets they belong to are termed *collusive sets.*

A collusion system also specifies a utility function $\mathbb{U}_C$ for each subset C in $\mathbb{C}$, and the members of

C act so as to maximize $\mathbb{U}_C$. The vector $\mathbb{U}$ must therefore satisfy the following constraint: if $\{i\} \in \mathbb{C}$, that is if i is an independent player, then $\mathbb{U}_{\{i\}}$ must coincide with i's individual utility function u_i.

If C is a collusive set, however, no restrictions are envisaged for the *collective utility function* $\mathbb{U}_C$: it can be any function of TV_C, —that is the true-valuation sub-profile of the members of C, of P_C —that is the profile of prices of the members of C, and of A_C, that is of the subset of goods allocated to the members of C. In particular, $\mathbb{U}_C$ may be the function that, for any outcome (A, P) returns

- the sum of the individual utilities of C's members.
- the sum of: the individual utility of C's first member, half of the utility of C's second member, a third of the individual utility of C's third member, and so on.
- A constant c.

A collusion system also need to specify the knowledge $\mathbb{K}_C$ about the true-valuation profile TV available to each set C in $\mathbb{C}$. This knowledge is sometimes crucial for C's maximization of its collective utility. The only restriction on the vector $\mathbb{K}$ is about independent players. Namely, if $\{i\} \in \mathbb{C}$, $\mathbb{K}_{\{i\}}$ must coincide with K_i, that is with i's individual external knowledge. Although it is natural to assume, and will indeed be required in forthcoming papers, that "the *collective knowledge* of a collusive set C in $\mathbb{C}$, $\mathbb{K}_C$, cannot be more accurate than the intersection of the external individual knowledge of its members", for the results of this present paper, no restriction is required for it. That is, our results in this present paper do not dependent on the members of C having collective knowledge of a particular form, and hold even when their collective knowledge is wrong. Because of this, in the definitions following, we do not consider $\mathbb{K}_C$ as part of the collusion system, implying that it is totally arbitrary.

Generality of the model. Before summarizing the above discussion into Definition 4, let us point out how general our static model is.

Our collusive sets are suddenly fixed, but chosen arbitrarily. As a result, although we do not enter into "coalition formation, our collusion systems validly represent the end point of enormously many possible collusion-formation processes. Sometimes players join a collusive set C in order to maximize their own *broadly construed* utilities. Notice that such "individual aims" are totally compatible with our requirement that C's members maximize $\mathbb{U}_C$. For instance, after maximizing the sum of their individual utilities, the members of C are free to divide this collective utility among themselves in any way they want. In particular, one half can go to a distinguished member, and the balance be partitioned equally among the remaining members. As for another example, $\mathbb{U}_C$ may coincide with the individual utility of just one of C's members, and the other players may get along with maximizing his utility only because they are blackmailed by him, or because he offers them a million dollar each and binds them by a contract to do his bidding.

Notice too that making the collective knowledge of a collusive set C totally arbitrary is actually crucial to capture possible ways for C to come into existence. For instance, the members of C may have decided to form a collusive set and how to divide their collective utilities after a laborious negotiation. But because the final outcome of this negotiation was not a forgone conclusion, they may not have right volunteered their total knowledge about the true-valuation profile TV, and after reaching a mutually satisfactory agreement, it might have become awkward or impossible for them to tell that they actually had even more precise knowledge. Indeed, the final knowledge on which a collusive set C relies upon, may be "less" than that of any of C's members, or even false. For instance, consider a player i offering one million dollars to any other player that he wishes to enter a collusive set whose collective utility coincides with u_i and whose bids are chosen by i. In this case, i may want to invite and pay only players whose knowledge will be particularly beneficial to him, but then it is also possible that some other player j lies to i about his own knowledge and is believed by i.

Definition 4. *In a combinatorial auction,* $(\mathbb{C}, I, TV, K, \mathbb{U})$ *is a* collusion system *if and only if*

- $\mathbb{C}$ *is a partition of the player set.*
- I *is the set of all players* i *such that* $\{i\} \in \mathbb{C}$*.* (Set I is explicitly specified for convenience only.)
- TV *is the true valuation profile, and* K *is the external knowledge profile of the auction.*
- $\mathbb{U}$ *is a vector of functions indexed by the subsets in* $\mathbb{C}$*. Specifically, for each subset of players* $C \in \mathbb{C}$, $\mathbb{U}_C$ *is the collusive utility function of* C*, and* $\mathbb{U}_{\{i\}}((A,P),TV) = TV_i(A_i) - P_i$ *whenever* $i \in I$*.*

We refer to a player in I *as* independent*, to a player not in* I *as* collusive*, and to a subset in* $\mathbb{C}$ *with cardinality* > 1 *as a* collusive set*. We denote the subset in* $\mathbb{C}$ *to which a player* i *belongs by* $\mathbb{C}_i$*. We refer to* $\mathbb{U}$ *as* the utility vector of $\mathbb{C}$*.*

A strategy vector σ *for* $(\mathbb{C}, I, TV, K, \mathbb{U})$ *is a vector of functions indexed by the subsets in* $\mathbb{C}$*:*

$$\sigma = \bigsqcup_{C \in \mathbb{C}} \sigma_C$$

where each σ_C *specifies the actions of all players in* C *during the auction. If* C *is a collusive set, we refer to* σ_C *as a collective strategy for* C*.*

Comments

- The fact that a player j is collusive does not imply that his individual utility function u_j ceases from "existing". Although j will now act so as to maximize the collective utility of his collusive set, we shall still reason about his individual utility function u_j.

- Here σ_C is a deterministic function —although the players in C may want to use a mixture of different strategies.

- We refer to a strategy vector σ as a *play* of a combinatorial auction, or of an auction mechanism. Since outcomes of auctions are ultimately determined by the players' strategies, it is convenient to consider a mechanism M as mapping plays σ to outcomes $M(\sigma)$, and to refer to its allocation and price component as M_a and M_p: that is, $M(\sigma) = (M_a(\sigma), M_p(\sigma))$ for all plays σ. When M is probabilistic, the *expected social welfare* and the *expected revenue* of a play σ, $\mathbb{E}[SW(\sigma, TV)]$ and $\mathbb{E}[REV(\sigma)]$, are respectively defined as $\mathbb{E}[SW(M(\sigma), TV)]$ and $\mathbb{E}[REV(M(\sigma))]$. More simply, we use $\mathbb{E}[SW]$ and $\mathbb{E}[REV]$ respectively, when σ is clear from context.

- Since again, outcomes of auctions are ultimately determined by the players' strategies, we may consider each collective utility function $\mathbb{U}_C$ and each individual utility function u_i to be the —possibly probabilistic— function that, for TV and σ, returns —respectively— the utility of the collusive set C and the individual utility of player i under σ and TV. Accordingly, the expected collective utility $\mathbb{E}[\mathbb{U}_C(\sigma, TV)]$ and the expected individual utility $\mathbb{E}[u_i(\sigma, TV)]$ are respectively defined as $\mathbb{E}[\mathbb{U}_C(M(\sigma), TV)]$ and $\mathbb{E}[u_i(M(\sigma), TV)]$. When the sequences of the coin tosses r used by the mechanism is fixed, we may also include r as $\mathbb{U}_C$ and u_i's input.

3.3 Individually Monotone Collective Utilities

As we have seen, in a collusion system $(\mathbb{C}, I, TV, K, \mathbb{U})$ each C in $\mathbb{C}$ tries to maximize its own collective utility u_C. But so far each u_C is totally unrestricted whenever C is a collusive set. Accordingly, the players in C are much more than collusive: they are "de facto" *irrational* —that is, as per the definition of [MV07.a], they are free to play arbitrary strategies. Accordingly, we need to restrict collective utilities somewhat.

To be reasonable, the collective utility of a collusive set C must ultimately be related to the individual utilities of its members. For example, the members of C may have agreed to maximizing the sum of their individual utilities, but then split the proceeds in some, not necessarily fair, way. As for another example, the collective utility of C may coincide with the individual utility of just one of his members, who may convince the other to go along, in a variety of ways, including blackmail, promise of future cooperation in totally different settings, a lump-sum payment, etc.

To maximize the meaningfulness of our mechanism, we want to assume collective utilities that are as general as possible. Informally, we assume that they are "individually monotone." By this we mean that, fixing the allocations and the prices of all players in C except for a player i, C's collective utility cannot but increase with i's individual utility. That is: (1) C's collective utility remains unchanged if the only change in the outcome consists of swapping the subset of goods A_i allocated to i with another set A'_i such that $TV_i(A_i) = TV_i(A'_i)$ and leaving i's price unchanged; and (2) C's collective utility increases if the only outcome change consists of either decreasing i's price keeping the subset of goods allocated to him the same, or keeping i's price the same but allocating him a subset of goods that he values more. Let us now be more precise.

Definition 5. *We say that the collective utility function u_C of a collusive set C is* individually monotone *if (1) for all players $i \in C$ and (2) for all outcomes (A, P) and (A', P') such that $(A_j, P_j) = (A'_j, P'_j)$ whenever $j \in C \setminus \{i\}$, we have:*

- *If $TV_i(A_i) - P_i = TV_i(A'_i) - P'_i$, then $u_C((A, P), TV) = u_C((A', P'), TV)$;*
- *If $TV_i(A_i) - P_i > TV_i(A'_i) - P'_i$, then $u_C((A, P), TV) > u_C((A', P'), TV)$.*

We say that a collusion system $(\mathbb{C}, I, TV, K, \mathbb{U})$ is individually monotone *if $\mathbb{U}_C$ is individually monotone for each collusive set $C \in \mathbb{C}$.*

Recall that C is a collusive set only if its cardinality is greater than 1, but note that $\mathbb{U}_{\{i\}}$ is individually monotone whenever $i \in I$. Note too that the first implication in the above definition does not imply that $(A_i, P_i) = (A'_i, P'_i)$.

4 Our Solution Concept

Mechanism design is typically satisfied with guaranteeing a given property $\mathbb{P}$ "at equilibrium." However, equilibria (besides being defined relative to the "deviation" of a single player) do not solely depend on the players' rationality, but also on their *beliefs*. Saying that a profile of strategies σ is an equilibrium only means that, for every player i, deviating from σ_i is an irrational thing to do (i.e., yields a lower utility for i) only if he believes that any other player j will stick to his strategy σ_j. When, as it is typically the case, there are multiple equilibria —σ, τ, ...— if some players believe that the equilibrium about to be played is σ while others believe it is τ, etc., the auction may not end up in any equilibrium at all. In fact "mixing and matching" the strategies of different equilibria needs not to result in an equilibrium! Accordingly, even if the desired property $\mathbb{P}$ were to be guaranteed at each of the possible equilibria, the problem of *equilibrium selection* prevents this to be a truly meaningful guarantee.

Aiming at robust guarantees, Adversarial Mechanism Design advocates the usage of solution concepts immune to the problems of equilibrium selection. One adequate solution concept is equilibrium in dominant strategies. Indeed, if σ is such an equilibrium then, for any player i, σ_i is i's best response to any possible strategies of the other players. (Adversarial Mechanism Design in dominant strategy is indeed possible, as demonstrated by the revenue mechanisms of [MV07.a] for truly combinatorial auctions.) In this paper we want to put forward other solution concepts that are adequate too. We believe these to be of independent interest. To bypass the problem of beliefs and equilibrium selection, our solution concepts are *equilibrium-less.*

4.1 The Bird-Eye View of Our Solution Concept

Let us present a series of new and equilibrium-less solution concepts, and then formalize just the last one, which is the one we actually use to analyze our mechanism. At the highest level all our solution concepts aim to be *"practically as good as dominant-solvability."*

Implementation in Surviving Strategies. There is no question that a game is adequately *solved* if, after iteratively removing all dominated strategies, a single strategy survives for each player.

Unfortunately, we cannot guarantee that our mechanism yields a dominant-solvable game. Indeed, after iteratively removing dominated strategies for all our *agents* (i.e., independent players and collusive sets), plenty of *surviving strategies* will exist for each agent. Accordingly, we cannot predict with certainty which strategy vector will be actually played. But predictability of the "actual" strategy vector is *a useful mean, not the goal.* Indeed, to guarantee that a mechanism satisfies its desired property it suffice to prove that *the desired property holds for any vector of surviving strategies.*

We call this notion *implementation in surviving strategy.* As we shall see, our mechanism essentially is an example of such an implementation.

Note that although we define this notion in a collusive setting, the notion a fortiori can be defined (and achieved) in the traditional setting, where all players are independent. For such traditional setting, our notion generalizes that of an implementation in undominated strategies of [BLP06].

Implementation in Knowingly Surviving Strategies. We believe and hope that implementation in surviving strategies will be useful in its own. But we predict that it will play a larger role in games of perfect information. When this is not the case, it may be hard for an agent to figure out whether one of its strategies is indeed dominated by another. Specifically, consider an independent player i in our setting. Such player i has no idea about which sets —if any— may be collusive nor about the collective utility function that each collusive set tries to maximize. Accordingly, the best i can do at each "iteration" is to eliminate those of his strategies that, according to his knowledge K_i, he is sure that are dominated. As a result, upon completion of the elimination process, the set of surviving strategies for each agent will be even "larger." Indeed, *knowingly surviving strategies* are a superset of surviving strategies. Nonetheless, if we could guarantee that the desired property hold for any vector of knowingly surviving strategies, we would be equally happy.

We call this (sketched) solution concept *implementation in knowingly surviving strategies.* Our mechanism actually satisfies a stronger version of this solution concept.

Implementation in Knowingly Σ^1/Σ^2 Surviving Strategies. Experience seems to indicated that there are different "levels of rationality." By this we mean that, in practice, many players are capable of completing the first few iterations of elimination of dominated strategies, but fail to go "all the way."

Accordingly, the above solution concept would be "stronger" if the desired property were guaranteed for any vector of strategies surviving just the first few iterations. This is exactly the case for our property and our mechanism. Specifically, we achieve our benchmark whenever

- Each collusive set C plays any knowingly undominated strategy σ_C (i.e., one surviving the first iteration); and
- Each independent player i plays any strategy σ_i surviving the first two iterations.

We call such a solution concept *Implementation in Knowingly Σ^1/Σ^2 Surviving Strategies.*

The rest of this Section 4 is devoted formalize this latter notion. To do so, we must first formalize Σ^1 for both independent and collusive agents, and Σ^2 for independent ones.

Notice that in the following subsections, the utility of an agent is "expected" only because the mechanism may be probabilistic. Indeed, we are not in the Bayesian setting, and the strategies to be compared are deterministic. Therefore when the sequence r of the coin tosses used by the mechanism is fixed, the utility function of an agent is also deterministic. In such a case, we further consider the utility function of an agent to be the function that for TV, r and any strategy vector σ, returns the utility of that agent under σ, TV and r.

For any collusion system $(\mathbb{C}, I, TV, K\mathbb{U})$ and any $C \in \mathbb{C}$, we denote by $\Sigma^0_{C,(\mathbb{C},I,TV,K,\mathbb{U})}$ the set of all possible deterministic-strategies of C, and by $\Sigma^0_{(\mathbb{C},I,TV,K,\mathbb{U})} = \prod_{C\in\mathbb{C}} \Sigma^0_{C,(\mathbb{C},I,TV,K,\mathbb{U})}$ the set of all possible deterministic-strategy vectors for $(\mathbb{C}, I, TV, K, \mathbb{U})$.

4.2 Knowingly Σ^1 Surviving Strategies for Independent Players in a Collusion System

Defining knowingly dominated and undominated strategies is easier for independent players, because it is clear what their knowledge is. Indeed, our definition parallels the traditional one. The main difference is that when comparing a strategy σ_i with another strategy σ'_i, i needs to consider his utility only for those true-valuations profiles V *compatible* with his knowledge K_i; that those V such that (1) V_i coincides with TV_i, which player i knows with certainty, while (2) V_{-i} coincides with any one of the candidates known to him, that is with one of the sub-profiles in K_i.

Definition 6. (Knowingly Dominated Strategies of Independent players in a Collusion System) *Let*

- $(\mathbb{C}, I, TV, K, \mathbb{U})$ *be a collusion system;*
- Σ' *a set of deterministic-strategy vectors for* $(\mathbb{C}, I, TV, K, \mathbb{U})$, $\Sigma' = \prod_{C\in\mathbb{C}} \Sigma'_C$*; and*
- i *an independent player, and* σ_i *and* σ'_i *two deterministic strategies of* i *in* $\Sigma'_{\{i\}}$.

We say that σ_i *is* knowingly dominated by σ'_i over Σ' and $(\mathbb{C}, I, TV, K, \mathbb{U})$, *if* $\forall$ *valuation profiles* V *such that* $V_i = TV_i$ *and* $V_{-i} \in K_i$, *assuming that the true valuation profile is* V, *we have:*

1. $\forall \tau_{-i} \in \Sigma'_{-i}$, $E[u_i(\sigma_i \sqcup \tau_{-i}, V)] \le E[u_i(\sigma'_i \sqcup \tau_{-i}, V)]$.
2. $\exists \tau_{-i} \in \Sigma'_{-i}$ *such that* $E[u_i(\sigma_i \sqcup \tau_{-i}, V)] < E[u_i(\sigma'_i \sqcup \tau_{-i}, V)]$.

We say that σ_i *is* knowingly undominated over Σ' and $(\mathbb{C}, I, TV, K, \mathbb{U})$ *if it is not knowingly dominated by any* σ'_i *over* Σ' *and* $(\mathbb{C}, I, TV, K, \mathbb{U})$.

We denote by $\Sigma^1_{i,(\mathbb{C},I,TV,K,\mathbb{U})}$ *(or equivalently* $\Sigma^1_{\{i\},(\mathbb{C},I,TV,K,\mathbb{U})}$*) the set of deterministic strategies of* i *that are knowingly undominated over* $\Sigma^0_{(\mathbb{C},I,TV,K,\mathbb{U})}$*, and refer to it as* the set of knowingly Σ^1 surviving strategies of i over $(\mathbb{C}, I, TV, K, \mathbb{U})$.

4.3 Knowingly Σ^1 Surviving Strategies for Collusive Sets

Defining knowingly dominated and undominated strategies for collusive sets is much harder, because their collective knowledge is far from being clearly determined. Of course, this problem can be solved by *assuming* that the collective knowledge of a collusive set C, $\mathbb{K}_C$, has a specific form. For instance, one can assume that $\mathbb{K}_C = \cap_{i\in C}(TV_i \sqcup K_i)$; that is one can assume that the collective knowledge of a collusive set C coincides with the true valuations of its members and the intersection of their external knowledge. This would indeed be the most accurate collective knowledge derivable from their individual knowledge.

However, this assumption cannot be made with great confidence. Although we focus on the "end result", collusive sets must form in some specific fashion. And the way they get formed ultimately dictate their collective knowledge (as well as their collective utility). Let us consider two examples. In the first example, the players in some subset C enter a binding agreement to coordinate their bids and then split in predetermined, and possible different, proportions the sum of their individual utilities, using *side payments* as necessary. Thus it is in everybody's best interest to maximize the utility function u_C consisting of the sum of their individual utilities. First, however, they must precisely determine u_C. To do this, they must share with each other their true valuations and their external knowledge. But then, the temptation arise for some player to to report a different true valuation, so as to have to make a smaller side payment in the end. Indeed, by receiving the agreed upon portion of a smaller "total utility" but making a "smaller side payment" he may be better off than being truthful. Accordingly, unless C can avail itself of the help of an angel that truthfully announces to them their collective utility in the final outcome, $\mathbb{K}_C$ may be vastly different than $\cap_{i\in C}(TV_i \sqcup K_i)$.

As for a second example, consider a collusive set C arising from the efforts of a "leader" j, who having 5 million dollars to spend, he is ready to pay one million dollars to a player i in exchange of learning TV_i and K_i, and of obtaining that i bids the way j wants. Clearly, j would rather corrupt players who know a lot, and thus again incentives are created for players to boast a knowledge more accurate of what they have in order to be chosen to enter the collusive set and being paid the million dollar. Hopefully, if they are lucky they will not be discovered and will not pay any consequences.

Again therefore, no matter what the desired collective knowledge for C may be, the true $\mathbb{K}_C$ may be vastly different from it.

Accordingly, any one is free to assume anything they want, but to assume $\cap_{i\in C}(TV_i \sqcup K_i)$, or other precise forms for the collective knowledge of a collusive set, requires more and more unrealistic settings, such as those in which the total knowledge of a player is "provable".

Finding all such assumptions unappealing and unsafe, we are happy to rely on the safest possible way of determining a knowingly dominated collective strategy for a collusive set C. Indeed, since we are already assuming that all utility functions are individually monotone, *we shall determine that a collective strategy of C is knowingly dominated based only on the individual monotonicity of C's collective utility function.* Of course, this will enable us to eliminated "much fewer" collective strategies, thus making Σ^1 "much bigger", but then this does not ultimately matter since we shall prove any way that we can achieve our benchmark in Σ^1/Σ^2 surviving strategies!

Let us now see how individual monotonicity enables us to knowingly eliminate a collective strategy σ_C for C. As we said, this elimination is quite minimal, indeed it can only apply whenever σ_C and σ'_C are "individually-monotone comparable", IM-comparable for short. We thus need to define this notion first.

Essentially, we say that two different, deterministic, and collective strategies σ_C and σ'_C are "IM-comparable" if (1) there exists a set of decision nodes S such that σ_C and σ'_C are equal at every node outside S and different at every node in S and (2) for every two deterministic strategy vectors τ and τ', equal except that $\tau_C = \sigma_C$ and $\tau'_C = \sigma'_C$, whenever a node $v \in S$ is reached in both an execution of τ and an execution of τ' in which the coin tosses of the mechanism are the same, then, if v is a decision node of i (necessarily in C) the final respective outcomes of these executions, (A, P) and (A', P') are identical for every member of C except for the allocation and the price of i.

Now, letting σ_C and σ'_C are IM-comparable, we say that the first is dominated by the second if, in any play, the latter strategy gives player i an individual utility that always is at least as good as that given by σ_C, and sometimes strictly better.

So far, we have *de facto* argued that eliminating σ_C in favor of σ'_C is beter for i i*individually.* Let us now point out that, whenever the utility of C is individually monotone, i's individual incentives and C's collective incentives are locally aligned at every decision node of i in S. Accordingly, what is "dominated for i is also dominated for C." Moreover, whoever may coordinate the actions of the players in C so as to maximize C's collective utility, can recognize the set S, since S's definition does not depend on the individual knowledge of the members of C! Moreover, for each node v of S, he can also determine the player i such that v is i's decision node. Accordingly, C's coordinator always has the option, whenever node v is reached of delegating the choice of action on behalf of i to i himself. Knowing full well that, since their interests are properly aligned, he can get the best response on behalf of C, without having to trust any "revealed knowledge" to be genuine!

Let us now be more precise.

Definition 7. (IM-comparable Collective Strategies for Collusive Sets in a Collusion System) *Let*

- $(\mathbb{C}, I, TV, K, \mathbb{U})$ *be a collusion system;*
- $\Sigma' = \prod_{C \in \mathbb{C}} \Sigma'_C$ *a set of strategy vectors for* $(\mathbb{C}, I, TV, K, \mathbb{U})$ *where* Σ'_C *is a set of strategies of* C*;*
- C *a collusive set in* $\mathbb{C}$*,* σ_C *and* σ'_C *two of its different, deterministic and collective strategies in* Σ'_C*, and* S *a set of decision nodes such that* σ_C *and* σ'_C *differ and only differ at every node in* S*.*

We say that σ_C *and* σ'_C *are* IM-comparable collective strategies of C over Σ' and $(\mathbb{C}, I, TV, K, \mathbb{U})$*, if* $\forall \tau_{-C} \in \Sigma'_{-C}$ *and* $\forall$ *sequences* r *of the coin tosses used by the mechanism, we have:*

1. *either the executions of* $\sigma_C \sqcup \tau_{-C}$ *and* $\sigma'_C \sqcup \tau_{-C}$ *with the same* r *reach the same sequence of decision nodes (therefore yield the same outcome); or*
2. *the executions of* $\sigma_C \sqcup \tau_{-C}$ *and* $\sigma'_C \sqcup \tau_{-C}$ *with the same* r *both reach a node* $v \in S$ *belonging to some player* i *(in* C*), and respectively yield two outcomes* (A, P) *and* (A', P') *such that* $(A_{C\setminus\{i\}}, P_{C\setminus\{i\}}) = (A'_{C\setminus\{i\}}, P'_{C\setminus\{i\}})$*.*

Definition 8. (Knowingly Dominated Strategies for Collusive Sets in a Collusion System) *Let*

- $(\mathbb{C}, I, TV, K, \mathbb{U})$ *be an individually monotone collusion system;*
- $\Sigma' = \prod_{C \in \mathbb{C}} \Sigma'_C$ *a set of strategy vectors for* $(\mathbb{C}, I, TV, K, \mathbb{U})$ *where* Σ'_C *is a set of strategies of* C*;*
- C *a collusive set in* $\mathbb{C}$*, and* σ_C *and* σ'_C *two of its deterministic and collective strategies in* Σ'_C*.*

We say that σ_C *is* knowingly dominated by σ'_C over Σ' and $(\mathbb{C}, I, TV, K, \mathbb{U})$ *if*

1. σ_C *and* σ'_C *are IM-comparable collective strategies over* Σ' *and* $(\mathbb{C}, I, TV, K, \mathbb{U})$*.*
2. *Letting* S *be the set of decision nodes at and only at which* σ_C *and* σ'_C *differ.*
   $\forall i \in C$ *such that there exists* $\tau_{-C} \in \Sigma'_{-C}$ *and a sequence* r *of coin tosses of the mechanism for which, the executions of* $\sigma_C \sqcup \tau_{-C}$ *and* $\sigma'_C \sqcup \tau_{-C}$ *with the same* r *both reach some* $v \in S$ *belonging*

to i*; and* $\forall$ *valuation profiles* V *such that* $V_i = TV_i$ *and* $V_{-i} \in K_i$*, assuming that the true valuation profile is* V*, we have:*

a. $\forall \tau_{-C} \in \Sigma'_{-C}$ *and sequences* r *of coin tosses of the mechanism such that the executions of* $\sigma_C \sqcup \tau_{-C}$ *and* $\sigma'_C \sqcup \tau_{-C}$ *with the same* r *both reach some* $v \in S$ *belonging to* i*:* $u_i(\sigma_C \sqcup \tau_{-C}, r, V) \leq u_i(\sigma'_C \sqcup \tau_{-C}, r, V)$ *(therefore* $\mathbb{U}_C(\sigma_C \sqcup \tau_{-C}, r, V) \leq \mathbb{U}_C(\sigma'_C \sqcup \tau_{-C}, r, V)$*).*

b. $\exists \tau_{-C} \in \Sigma'_{-C}$ *and a sequence* r *of coin tosses of the mechanism such that: the executions of* $\sigma_C \sqcup \tau_{-C}$ *and* $\sigma'_C \sqcup \tau_{-C}$ *with the same* r *both reach some* $v \in S$ *belonging to* i*, and* $u_i(\sigma_C \sqcup \tau_{-C}, r, V) < u_i(\sigma'_C \sqcup \tau_{-C}, r, V)$ *(therefore* $\mathbb{U}_C(\sigma_C \sqcup \tau_{-C}, r, V) < \mathbb{U}_C(\sigma'_C \sqcup \tau_{-C}, r, V)$*).*

We say that σ_C *is* knowingly undominated over Σ' and $(\mathbb{C}, I, TV, K, \mathbb{U})$ *if it is not knowingly dominated by any* σ'_C *over* Σ' *and* $(\mathbb{C}, I, TV, K, \mathbb{U})$*.*

We denote by $\Sigma^1_{C,(\mathbb{C},I,TV,K,\mathbb{U})}$ *the set of deterministic collective strategies of* C *that are knowingly undominated over* $\Sigma^0_{(\mathbb{C},I,TV,K,\mathbb{U})}$*, and refer to it as* the set of knowingly Σ^1 surviving strategies of C over $(\mathbb{C}, I, TV, K, \mathbb{U})$.

Definition 9. (Knowingly Σ^1 Surviving Strategies for Collusive Sets) *Let*

- C *be a collusive set in some (implicit) collusion system with individually monotone utility function* u_C*, true valuation sub-profile* tv_C *and external knowledge sub-profile* k_C*; and*
- $\mathtt{IM}(C, u_C, tv_C, k_C)$ *be the set of all individually monotone collusion systems* $\mathbb{S} = (\mathbb{C}, I, TV, K, \mathbb{U})$ *such that* $C \in \mathbb{C}$*,* $TV_C = tv_C$*,* $K_C = k_C$ *and* $\mathbb{U}_C = u_C$*.*

We define

$$\Sigma_C^{1,\mathtt{IM}(C,u_C,tv_C,k_C)} = \bigcup_{\mathbb{S} \in \mathtt{IM}(C,u_C,tv_C,k_C)} \Sigma^1_{C,\mathbb{S}}$$

and refer to it as the set of knowingly Σ^1 surviving strategies of C over $\mathtt{IM}$.

4.4 Knowingly Σ^2 Surviving Strategies for Independent Players

Definition 10. (Knowingly Σ^2 Surviving Strategies for Independent Players in a Collusion System) *Let* $\mathbb{S} = (\mathbb{C}, I, TV, K, \mathbb{U})$ *be an individually monotone collusion system and* i *be an independent player. We define* $\Sigma^1_{\mathbb{S}} = \prod_{C \in \mathbb{C}} \Sigma^1_{C,\mathbb{S}}$*, denote by* $\Sigma^2_{i,\mathbb{S}}$ *the set of deterministic strategies of* i *that are knowingly undominated over* $\Sigma^1_{\mathbb{S}}$ *and* $\mathbb{S}$*, and refer to it as* the set of knowingly Σ^2 surviving strategies of i over $\mathbb{S}$.

Definition 11. (Knowingly Σ^2 Surviving Strategies for Independent Players) *Let i be an independent player in some (implicit) collusion system with true valuation tv_i and external knowledge k_i, and $\mathtt{IM}(i, tv_i, k_i)$ be the set of all individually monotone collusion systems $\mathbb{S} = (\mathbb{C}, I, TV, K, \mathbb{U})$ such that $i \in I$, $TV_i = tv_i$ and $K_i = k_i$. We define*

$$\Sigma_i^{2,\mathtt{IM}(i,tv_i,k_i)} = \bigcup_{\mathbb{S} \in \mathtt{IM}(i,tv_i,k_i)} \Sigma_{i,\mathbb{S}}^2$$

and refer to it as the set of knowingly Σ^2 surviving strategies of i over $\mathtt{IM}$.

Definition 12. *Let $(\mathbb{C}, I, TV, K, \mathbb{U})$ be an individually monotone collusion system, and σ a strategy vector for $(\mathbb{C}, I, TV, K, \mathbb{U})$. We say that σ is a* knowingly Σ^1/Σ^2 surviving play for $(\mathbb{C}, I, TV, K, \mathbb{U})$ *over* $\mathtt{IM}$ *if*

$$\sigma \in \prod_{i \in I} \Sigma_i^{2,\mathtt{IM}(i,TV_i,K_i)} \times \prod_{C \in \mathbb{C}, |C|>1} \Sigma_C^{1,\mathtt{IM}(C,\mathbb{U}_C,TV_C,K_C)}.$$

If $(\mathbb{C}, I, TV, K, \mathbb{U})$ are clear from context, we more simply call σ a knowingly surviving play.

Definition 13. (Implementation in knowingly Σ^1/Σ^2 Surviving Strategies) *Let $\mathbb{P}$ be a property over auction outcomes, and M an auction mechanism. We say that M* implements $\mathbb{P}$ in knowingly Σ^1/Σ^2 surviving strategies *if, for all truly combinatorial auctions with n players and m goods, all individually monotone collusion systems $(\mathbb{C}, I, TV, K, \mathbb{U})$, and all knowingly surviving play σ:*

$\mathbb{P}$ is expected to hold over outcomes distributed according to $M(\sigma)$.

5 The Statement of Our Theorem

Recall that in this paper we aim at answer the following question in Adversarial Mechanism Design:

What is the combined performance achievable in truly combinatorial auctions?

Our main theorem below provides a very general and robust answer to this question. In the language of our solution concept, it says that, letting the property $\mathbb{P}$ be such that "the sum of the expected social welfare and the expected revenue reaches at least half of the $\mathbb{MEW}$ benchmark evaluated on the external knowledge sub-profile of independent players", then there exists an auction mechanism that implements $\mathbb{P}$ in knowingly Σ^1/Σ^2 surviving strategies.

Theorem 1. *There exists an auction mechanism M such that, for all integers n and m, all truly combinatorial auctions with n players and m goods, all individually monotone collusion systems*

$(\mathbb{C}, I, TV, K, \mathbb{U})$, *and all knowingly surviving play* σ*, the sum of the expected social welfare and the expected revenue of* $M(\sigma)$ *is at least*

$$\frac{\mathbb{MEW}(K_I)}{2}.$$

We prove Theorem 1 by explicitly constructing a simple and probabilistic auction mechanism.

6 Our Mechanism

As promised, our mechanism is of public extensive-form, and actually consists of three stages: two player stages followed by a final mechanism stage, where the mechanism produces the final outcome (A, P).

In the first stage, each player i publicly (and simultaneously with the others) announces (1) a canonical outcome Ω_i for the players in $-i$; and (2) a subset of goods S_i. (Allegedly, Ω_i is actually feasible, and indeed represents the "best way known to i to sell the goods to the other players." Allegedly too, S_i is i's favorite subset of goods, that is the one i values the most.)

After the first stage, everyone can compute (a) the revenue R_i of Ω_i for each player i, (b) the highest and second highest of such revenues, respectively denoted by $R_\star$ and R', and (c) the player whose announced outcome has the highest revenue —the lexicographically first player in case of "ties". Such player is called the "star player" and is denoted by "$\star$". (Thus, $\star \in N$.)

In the second stage, each player i, envisioned to receive a non-empty set of goods (for a positive price) in $\Omega_\star$, publicly (and simultaneously with the other such players) answers yes or no to the following implicit question: "are you willing to pay your envisioned price for your envisioned goods?" (The players not receiving any goods according to $\Omega_\star$ announce the empty string.)

After the second stage, for each asked player i who answers no, the star player is punished with a fine equal to the price he envisioned for i.

In the third and final stage, the mechanism flips a fair coin. If Heads, $S_\star$ is given to the star player at no *additional* charge (and thus player $\star$ pays nothing altogether if no player says no in the second stage). If Tails, (1) the goods are sold according to $\Omega_\star$ to the players who answered yes in the second stage, (2) all the revenue generated by this sale is given to the star player, and (3) the star player additionally pays R' to the seller/auctioneer. (Thus, the star player pays only R' if he has not been fined.) A more precise description of our mechanism is given below. In it, for convenience, we also include three "variable-update stages" and mark them by the symbol "•". In such stages the contents of some public variables are updated based on the strings announced so far.

Mechanism $\mathcal{M}$

- Set $A_i = \emptyset$ and $P_i = 0$ for each player i.

1. Each player i simultaneously and publicly announces (1) a canonical outcome for $-i$, $\Omega_i = (\alpha^i, \pi^i)$, and (2) a subset S_i of the goods.

- Set: $R_i = REV(\Omega_i)$ for each player i, $\star = \arg\max_i R_i$, and $R' = \max_{i \neq \star} R_i$.

  (We shall refer to player $\star$ as the "star player", and to R' as the "second highest revenue".)

2. Each player i such that $\alpha_i^\star \neq \emptyset$ simultaneously and publicly announces YES or NO.

- For each player i who announces NO, $P_\star = P_\star + \pi_i^\star$.

3. Publicly flip a fair coin.

   – If Heads, reset $A_\star = S_\star$.

   – If Tails: (1) reset $P_\star = P_\star + R'$; and (2) for each player i who announced YES in Stage 2, reset: $A_i = \alpha_i^\star$, $P_i = \pi_i^\star$, and $P_\star = P_\star - P_i$.

Comment. The outcome (A, P) may not be canonical, as the price of the star player may be non-zero even though he may receive nothing.

7 Analysis of Our Mechanism

Theorem 2. $\forall$ *integers* n *and* m*;* $\forall$ *individually monotone collusion systems* $(\mathbb{C}, I, TV, K, \mathbb{U})$ *for* n *players and* m *goods; and* $\forall$ *knowingly surviving plays* σ *of* $\mathcal{M}$*:*

the sum of the expected social welfare and the expected revenue of $\mathcal{M}(\sigma)$ *is at least*

$$\frac{\mathbb{MEW}(K_I)}{2}.$$

Proof. We base our proof on that of five simpler claims about the actions of the independent players i in any strategy in $\Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $\Sigma^2_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, and the actions of the collusive sets C in any collective strategy in $\Sigma^1_{C,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, where $(\mathcal{C}, \mathcal{I}, \mathcal{TV}, \mathcal{K}, \mathcal{U})$ is any individually monotone collusion system for n players and m goods, and n and m are as above.

Claim 1. *$\forall$ individually monotone collusion systems $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, $\forall i\in\mathcal{I}$, and $\forall\sigma_i\in\Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, if $i\neq\star$, then in Stage 2, the following two implications hold for σ_i:*

1. *i answers YES whenever $\mathcal{TV}_i(\alpha_i^\star)>\pi_i^\star$, and*
2. *i answers NO whenever $\mathcal{TV}_i(\alpha_i^\star)<\pi_i^\star$.*

Proof of Claim 1. We restrict ourselves to just prove the first implication. (The proof of the second implication is indeed totally symmetric.) We proceed by contradiction. Assume that $\exists$ individually monotone collusion systems $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, $\exists i\in\mathcal{I}$, and $\exists\sigma_i\in\Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, such that it is possible that when $i\neq\star$ and $\mathcal{TV}_i(\alpha_i^\star)>\pi_i^\star$, in Stage 2, i answers NO. Consider the following alternative strategy for player i:

Strategy σ_i':

- *Stage 1.* Run σ_i (with stage input "1" and private inputs $\mathcal{TV}_i$ and $\mathcal{K}_i$) and announce Ω_i and S_i as σ_i does.
- *Stage 2.* If $i=\star$ or $\mathcal{TV}_i(\alpha_i^\star)\leq\pi_i^\star$, run σ_i and answer whatever σ_i does.[3]
  Else (i.e., $i\neq\star$ and $\mathcal{TV}_i(\alpha_i^\star)>\pi_i^\star$), answer YES.

We derive a contradiction by proving that σ_i is knowingly dominated by σ_i' over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$. Therefore $\sigma_i\notin\Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, which contradicts the assumption that $\sigma_i\in\Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$.

Towards proving that σ_i is knowingly dominated by σ_i' over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, for all valuation profiles $\mathcal{V}$ such that $\mathcal{V}_i=\mathcal{TV}_i$ and $\mathcal{V}_{-i}\in\mathcal{K}_i$, we only consider all strategy sub-vectors $\tau_{-i}\in\Sigma^0_{-i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ such that in play $\sigma_i\sqcup\tau_{-i}$ $i\neq\star$, $\mathcal{V}_i(\alpha_i^\star)>\pi_i^\star$ and i answers NO in Stage 2 (notice that such a strategy indeed exists). Because for any other $\tau_{-i}\in\Sigma^0_{-i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, the plays $\sigma_i\sqcup\tau_{-i}$ and $\sigma_i'\sqcup\tau_{-i}$ coincide, therefore for any coin toss of $\mathcal{M}$, the outcomes in these two plays also coincide, and thus $\mathbb{E}[u_i(\sigma_i\sqcup\tau_{-i},\mathcal{V})]=\mathbb{E}[u_i(\sigma_i'\sqcup\tau_{-i},\mathcal{V})]$.

For all such τ_{-i}, observe that, since σ_i' coincides with σ_i in Stage 1, the outcome profile Ω is the same in the plays $\sigma_i\sqcup\tau_{-i}$ and $\sigma_i'\sqcup\tau_{-i}$. Accordingly, the star player too is the same in both plays. Since (by hypothesis) $i\neq\star$ in play $\sigma_i\sqcup\tau_{-i}$, $i\neq\star$ also in play $\sigma_i'\sqcup\tau_{-i}$. Finally, since $\alpha_i^\star$ (as well as $\pi_i^\star$) is the same in both plays and (by hypothesis) $\mathcal{V}_i(\alpha_i^\star)>\pi_i^\star$ in play $\sigma_i\sqcup\tau_{-i}$, $\mathcal{V}_i(\alpha_i^\star)>\pi_i^\star$ also in play $\sigma_i'\sqcup\tau_{-i}$.

We now distinguish two cases, each occurring with probability 1/2.

(1) $\mathcal{M}$'s coin toss c comes up Heads.

[3] The first implication of Claim 1 specifies that $i\neq\star$ and $\mathcal{TV}_i(\alpha_i^\star)>\pi_i^\star$. However, a strategy must be specified in all cases, and thus σ_i' must be specified also when $i=\star$ or $\mathcal{TV}_i(\alpha_i^\star)\leq\pi_i^\star$.

In this case,

$$u_i(\sigma_i \sqcup \tau_{-i}, c, \mathcal{V}) = u_i(\sigma_i' \sqcup \tau_{-i}, c, \mathcal{V}) = 0,$$

that is the individual utility of player i is 0, since only the star player receives goods.

(2) $\mathcal{M}$'s coin toss c comes up Tails.

In this case, by hypothesis $\mathcal{V}_i(\alpha_i^\star) > \pi_i^\star$, player i answers NO in play $\sigma_i \sqcup \tau_{-i}$ and answers YES in play $\sigma_i' \sqcup \tau_{-i}$. Thus the individual utility of i is different in the two plays: specifically

$$u_i(\sigma_i \sqcup \tau_{-i}, c, \mathcal{V}) = 0 \quad \text{and} \quad u_i(\sigma_i' \sqcup \tau_{-i}, c, \mathcal{V}) = \mathcal{V}_i(\alpha_i^\star) - \pi_i^\star = \mathcal{TV}_i(\alpha_i^\star) - \pi_i^\star > 0.$$

Combining the two cases yields

$$\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] < \mathbb{E}[u_i(\sigma_i' \sqcup \tau_{-i}, \mathcal{V})].$$

Therefore σ_i is knowingly dominated by σ_i' over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$. ■

Without loss of generality, we assume that $\forall$ individually monotone collusion systems $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, $\forall i \in \mathcal{I}$, and $\forall \sigma_i \in \Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, if $i \neq \star$, then in Stage 2, i answers YES if $\mathcal{TV}_i(\alpha_i^\star) = \pi_i^\star$.[4]

Claim 2. *$\forall$ individually monotone collusion systems $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, $\forall$ collusive sets $C \in \mathcal{C}$, and $\forall \sigma_C \in \Sigma^1_{C,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, if $\star \notin C$, then in Stage 2, the following two implications hold for σ_C and all players $i \in C$:*

1. *i answers YES whenever $\mathcal{TV}_i(\alpha_i^\star) > \pi_i^\star$, and*
2. *i answers NO whenever $\mathcal{TV}_i(\alpha_i^\star) < \pi_i^\star$.*

Proof of Claim 2. We again restrict ourselves to just prove the first implication, and proceed by contradiction. Assume that $\exists$ individually monotone collusion systems $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, $\exists$ collusion sets $C \in \mathcal{C}$, $\exists \sigma_C \in \Sigma^1_{C,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, and $\exists i \in C$, such that it is possible that when $\star \notin C$ and $\mathcal{TV}_i(\alpha_i^\star) > \pi_i^\star$, in Stage 2 i answers NO. Consider the following alternative collective strategy for C.

Strategy σ_C':

- *Stage 1.* Run σ_C and announce Ω_i and S_i as σ_i does.
- *Stage 2.* If $\star \in C$ or $\mathcal{TV}_i(\alpha_i^\star) \leq \pi_i^\star$, run σ_C and answer whatever σ_C does for all $j \in C$.

[4] Else, we can easily modify $\mathcal{M}$ such that when Stage 3 gets Tails, reset $P_i = \pi_i^\star - \epsilon$ where ϵ is an arbitrarily small positive number. Doing this only changes our benchmark in Theorem 2 from $\frac{\text{MEW}(K_I)}{2}$ to $\frac{\text{MEW}(K_I)}{2} - \gamma$, where γ is another arbitrarily small number. Therefore we ignore this point in the analysis later.

Else (i.e. $\star \notin C$ and $\mathcal{TV}_i(\alpha_i^\star) > \pi_i^\star$), run σ_C, answer whatever σ_C does for all $j \in C \setminus \{i\}$, and answer YES for i.

We derive a contradiction by proving that σ_C is knowingly dominated by σ'_C over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$. Therefore $\sigma_C \notin \Sigma^1_{C,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, which contradicts the fact that $\sigma_C \in \Sigma^1_{C,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$.

Towards proving that σ_C is knowingly dominated by σ'_C over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, we focus ourselves on proving that σ_C and σ'_C are IM-comparable over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$.

First observe that σ_C and σ'_C only differ at player i's decision nodes where i should answer YES or NO. Therefore we only consider all strategy sub-vectors $\tau_{-C} \in \Sigma^0_{-C,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ such that in play $\sigma_C \sqcup \tau_{-C}$ $\star \notin C$, $\mathcal{TV}_i(\alpha_i^\star) > \pi_i^\star$ and i answers NO in Stage 2 (notice that such a strategy indeed exists). Because for any other $\tau_{-C} \in \Sigma^0_{-C,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, for any coin toss of $\mathcal{M}$, the plays $\sigma_C \sqcup \tau_{-C}$ and $\sigma'_C \sqcup \tau_{-C}$ coincide everywhere (therefore the outcomes in these two plays are the same).

For all such τ_{-C}, for any coin toss used by $\mathcal{M}$, $\sigma_C \sqcup \tau_{-C}$ and $\sigma'_C \sqcup \tau_{-C}$ both reach i's decision node (in Stage 2) where i answers NO in the play $\sigma_C \sqcup \tau_{-C}$ and answers YES in the play $\sigma'_C \sqcup \tau_{-C}$. Observe that the star player is the same in both plays, and thus so is $\alpha^\star_C$, as well as $\pi^\star_C$. For any player $j \in C \setminus \{i\}$, the answer announced in Stage 2 is also the same in both plays. Therefore no matter what $\mathcal{M}$'s coin toss comes up,

$$(\mathcal{M}_a(\sigma_C \sqcup \tau_{-C})_{C\setminus\{i\}}, \mathcal{M}_p(\sigma_C \sqcup \tau_{-C})_{C\setminus\{i\}}) = (\mathcal{M}_a(\sigma'_C \sqcup \tau_{-C})_{C\setminus\{i\}}, \mathcal{M}_p(\sigma'_C \sqcup \tau_{-C})_{C\setminus\{i\}}).$$

Thus σ_C and σ'_C are IM-comparable over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$.

The remaining part to prove that σ_C is knowingly dominated by σ'_C over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$ is quite similar to that in Claim 1, and is therefore ignored. ■

Claim 3. *Let* $(\mathbb{C}, I, TV, K, \mathbb{U})$ *be the individually monotone collusion system in Theorem 2.* $\forall C \in \mathbb{C}$ *and* $\forall \sigma_C \in \Sigma_C^{1,\mathtt{IM}(C,\mathbb{U}_C,TV_C,K_C)}$, *if* $\star \neq C$, *then in Stage 2, the following two implications hold for* σ_C *and all players* $i \in C$:

1. *i answers YES whenever* $TV_i(\alpha_i^\star) > \pi_i^\star$, *and*
2. *i answers NO whenever* $TV_i(\alpha_i^\star) < \pi_i^\star$.

Proof of Claim 3. This claim follows directly from Definition 9 and Claim 2. ■

Claim 4. $\forall$ *individually monotone collusion systems* $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, $\forall i \in \mathcal{I}$, *and* $\forall \sigma_i \in \Sigma^2_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, *in Stage 1, the following two implications hold for* σ_i:

1. *i chooses S_i to be his favorite subset of goods, that is $S_i = \arg\max_{S \subseteq G} \mathcal{TV}_i(S)$.*

2. *i does not "under-bid", that is he announces Ω_i such that $REV(\Omega_i) \geq \mathbb{MEW}_i(\mathcal{K}_i)$.*

In Stage 2, the following two implications hold for σ_i:

1. *i answers YES whenever $\mathcal{TV}_i(\alpha_i^\star) \geq \pi_i^\star$, and*

2. *i answers NO whenever $\mathcal{TV}_i(\alpha_i^\star) < \pi_i^\star$.*

Proof of Claim 4. The implications for Stage 2 follow directly from Claim 1 and the fact that $\Sigma^2_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})} \subseteq \Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$.

We prove the implications for Stage 1 by contradiction. Assume that $\exists$ individually monotone collusion system $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, $\exists i \in \mathcal{I}$, and $\exists \sigma_i \in \Sigma^2_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ such that in Stage 1, at least one of the following statements for σ_i is true:

1. i announces S_i such that $\mathcal{TV}_i(S_i) < \mathcal{TV}_i(T)$ for some subset T of the goods.

2. i announces Ω_i such that $REV(\Omega_i) < \mathbb{MEW}_i(\mathcal{K}_i)$.

Now consider the following alternative strategy for player i.

Strategy σ_i':

- *Stage 1.* Run σ_i so as to compute the outcome Ω_i and the "desired" subset of goods S_i;
  If $REV(\Omega_i) < \mathbb{MEW}_i(\mathcal{K}_i)$, announce $\Omega_i' = (\alpha'^i, \pi'^i)$ such that
  $\Omega_i' = \arg\max_{\omega \in F(\mathcal{K}_i)} REV(\omega)$; else announce $\Omega_i' = (\alpha'^i, \pi'^i)$ such that $\Omega_i' = \Omega_i$.
  If $\mathcal{TV}_i(S_i) < \mathcal{TV}_i(T)$ for some subset T of the goods, announce S_i' such that
  $S_i' = \arg\max_{S \subseteq G} \mathcal{TV}_i(S)$; else announce S_i' such that $S_i' = S_i$.

- *Stage 2.* If $\star = i$ or $\alpha_i^\star = \emptyset$, announce the empty string.
  Else, announce YES if $\mathcal{TV}_i(\alpha_i^\star) \geq \pi_i^\star$, and announce NO if $\mathcal{TV}_i(\alpha_i^\star) < \pi_i^\star$.

We derive a contradiction by proving that σ_i is knowingly dominated by σ_i' over $\Sigma^1_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$, contradicting the fact that $\sigma_i \in \Sigma^2_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$.

Towards proving this, first we show that $\sigma_i' \in \Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$. Observe that in σ_i', the action of i in Stage 2 is consistent with Claim 1, therefore σ_i' can not be knowingly dominated by any σ_i'' over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$ if the only difference between the two strategies is in Stage 2. While in $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ there is no restriction about the players' answers in Stage 2. It is easy to verify that due to such arbitrary answers of the players, for any σ_i'' which differs from σ_i' in Stage 1, σ_i' can not be knowingly dominated by σ_i'' over $\Sigma^0_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$. Therefore $\sigma_i' \in \Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$.

Moreover, since (by hypothesis) $\sigma_i \in \Sigma^1_{i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, in Stage 2 in σ_i, when $i \neq \star$, i answers YES if $\mathcal{TV}_i(\alpha_i^\star) \geq \pi_i^\star$ and answers NO otherwise.

For all valuation profiles $\mathcal{V}$ such that $\mathcal{V}_i = \mathcal{TV}_i$ and $\mathcal{V}_{-i} \in \mathcal{K}_i$, for all strategy sub-vectors $\tau_{-i} \in \Sigma^1_{-i,(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$, observe that Ω_{-i} is the same in plays $\sigma_i \sqcup \tau_{-i}$ and $\sigma'_i \sqcup \tau_{-i}$. Therefore if $i \neq \star$ in both plays, then the star player is the same in both plays, so is $\alpha_i^\star$, as well as $\pi_i^\star$; if $i = \star$ in both plays, then the second highest revenue R' is the same in both plays. Also observe that by Claim 1 and Claim 2, in any play, for all $j \neq \star$, in Stage 2, j answers YES if $\mathcal{TV}_j(\alpha_j^\star) \geq \pi_j^\star$ and answers NO otherwise. Finally, if i announces $\Omega'_i = \arg\max_{\omega \in F(\mathcal{K}_i)} REV(\omega)$ and $i = \star$, then all players $j \neq i$ such that $\alpha_j^i \neq \emptyset$ answer YES in Stage 2, because Ω'_i is $\mathcal{V}_{-i}$-feasible for $-i$.

Let us now compare the expected individual utility of i in the two plays, using the notation "$\sum_{j:YES}$" (respectively, "$\sum_{j:NO}$") for the sum taken over every player j who answers YES (respectively, NO) in Stage 2 in the play $\sigma_i \sqcup \tau_{-i}$, and "$\sum_{j:YES'}$" (respectively, "$\sum_{j:NO'}$") for the sum taken over every player j who answers YES (respectively, NO) in Stage 2 in the play $\sigma'_i \sqcup \tau_{-i}$. We distinguish three cases.

(1) *$i \neq \star$ in play $\sigma_i \sqcup \tau_{-i}$ and play $\sigma'_i \sqcup \tau_{-i}$.*

There are three sub-cases.

(a) $\alpha_i^\star = \emptyset$. In this case we have $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] = \mathbb{E}[u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V})] = 0$.

(b) *$\alpha_i^\star \neq \emptyset$ and i answers NO.* Also in this case we have $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] = \mathbb{E}[u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V})] = 0$.

(c) *$\alpha_i^\star \neq \emptyset$ and i answers YES.* In this case, with probability $\frac{1}{2}$, $u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V}) = u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V}) = 0$, and with probability $\frac{1}{2}$, $u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V}) = u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V}) = \mathcal{V}_i(\alpha_i^\star) - \pi_i^\star = \mathcal{TV}_i(\alpha_i^\star) - \pi_i^\star$. Overall therefore, also in case (c) we have $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] = \mathbb{E}[u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V})]$.

(2) *$i = \star$ in play $\sigma_i \sqcup \tau_{-i}$.*

In this case, i's expected utility in play $\sigma_i \sqcup \tau_{-i}$ is the weighted sum of his utility when $\mathcal{M}$'s coin toss is Heads and his utility when $\mathcal{M}$'s coin toss is Tails.[5] Therefore we have

$$\begin{aligned}\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] &= \frac{\mathcal{V}_i(S_i) - \sum_{j:NO} \pi_j^i}{2} + \frac{\sum_{j:YES} \pi_j^i - \sum_{j:NO} \pi_j^i - R'}{2} \\ &= \frac{\mathcal{TV}_i(S_i) - \sum_{j:NO} \pi_j^i}{2} + \frac{\sum_{j:YES} \pi_j^i - \sum_{j:NO} \pi_j^i - R'}{2}.\end{aligned}$$

Let us now compare this expected utility with $\mathbb{E}[u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V})]$. Towards computing the latter utility in Case 2, notice that the present case implies that $i = \star$ also in play $\sigma'_i \sqcup \tau_{-i}$. In fact, we

[5]Both individual utilities are expected, if the strategies of the other players are probabilistic.

have already argued that the sub-profile Ω_{-i} is the same in both plays, and $REV(\Omega_i') \geq REV(\Omega_i)$ by construction. This fact also implies that the second-highest revenue R' is the same in both plays. Accordingly, we have

$$\begin{aligned}\mathbb{E}[u_i(\sigma_i' \sqcup \tau_{-i}, \mathcal{V})] &= \frac{\mathcal{V}_i(S_i') - \sum_{j:NO'} \pi_j'^i}{2} + \frac{\sum_{j:YES'} \pi_j'^i - \sum_{j:NO'} \pi_j'^i - R'}{2} \\ &= \frac{\mathcal{TV}_i(S_i') - \sum_{j:NO'} \pi_j'^i}{2} + \frac{\sum_{j:YES'} \pi_j'^i - \sum_{j:NO'} \pi_j'^i - R'}{2}.\end{aligned}$$

There are two sub-cases.

(a) $REV(\Omega_i) < \mathbb{MEW}_i(\mathcal{K}_i)$.

In this case we have $\Omega_i' = \arg\max_{\omega \in F(\mathcal{K}_i)} REV(\omega)$ by construction. Because Ω_i' is a "feasible way of selling the goods to the players in $-i$," according to Claim 1 and 2, every player $j \neq i$ such that $\alpha_j^i \neq \emptyset$ answers YES in Stage 2: in our notation $\sum_{j:YES'} = \sum_j$. Accordingly, we have

$$\begin{aligned}\mathbb{E}[u_i(\sigma_i' \sqcup \tau_{-i}, \mathcal{V})] &= \frac{\mathcal{TV}_i(S_i')}{2} + \frac{\sum_j \pi_j'^i - R'}{2} = \frac{\mathcal{TV}_i(S_i') + REV(\Omega_i') - R'}{2} \\ &= \frac{\mathcal{TV}_i(S_i') + \mathbb{MEW}_i(\mathcal{K}_i) - R'}{2} \\ &> \frac{\mathcal{TV}_i(S_i) + REV(\Omega_i) - R'}{2} = \frac{\mathcal{TV}_i(S_i)}{2} + \frac{\sum_j \pi_j^i - R'}{2} \\ &\geq \frac{\mathcal{TV}_i(S_i) - \sum_{j:NO} \pi_j^i}{2} + \frac{\sum_{j:YES} \pi_j^i - \sum_{j:NO} \pi_j^i - R'}{2} \\ &= \mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})].\end{aligned}$$

(The strict inequality follows from the hypothesis of Case (a) and our construction of σ_C', that is $\mathcal{TV}_i(S_i') \geq \mathcal{TV}_i(S_i)$.)

(b) $REV(\Omega_i) \geq \mathbb{MEW}_i(\mathcal{K}_i)$.

Notice that by hypothesis, this case implies that $\mathcal{TV}_i(S_i) < \max_{S \subseteq G} \mathcal{TV}_i(S)$. By construction we have $\Omega_i' = \Omega_i$ and $\mathcal{TV}_i(S_i') = \max_{S \subseteq G} \mathcal{TV}_i(S)$. According to Claim 1 and 2, each player $j \neq i$ announces the same answer in both plays. Therefore

$$\begin{aligned}\mathbb{E}[u_i(\sigma_i' \sqcup \tau_{-i}, \mathcal{V})] &> \frac{\mathcal{TV}_i(S_i) - \sum_{j:NO} \pi_j^i}{2} + \frac{\sum_{j:YES} \pi_j^i - \sum_{j:NO} \pi_j^i - R'}{2} \\ &= \mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})].\end{aligned}$$

(3) $i \neq \star$ *in play* $\sigma_i \sqcup \tau_{-i}$ and $i = \star$ *in play* $\sigma'_i \sqcup \tau_{-i}$.

In this case, let us prove that $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] \leq \frac{TV_i(S'_i)}{2} \leq \mathbb{E}[u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V})]$.

To upperbound $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})]$ we consider three sub-cases for play $\sigma_i \sqcup \tau_{-i}$.

(a) $\alpha_i^\star = \emptyset$. In this case, $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] = 0$.

(b) $\alpha_i^\star \neq \emptyset$ *and* $\mathcal{TV}_i(\alpha_i^\star) < \pi_i^\star$. In this case, according to Claim 1, player i answers NO in Stage 2, and thus $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] = 0$.

(c) $\alpha_i^\star \neq \emptyset$ *and* $\mathcal{TV}_i(\alpha_i^\star) \geq \pi_i^\star$. In this case, according to Claim 1, player i answers YES in Stage 2, and thus $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] = \frac{\mathcal{V}_i(\alpha_i^\star) - \pi_i^\star}{2} = \frac{\mathcal{TV}_i(\alpha_i^\star) - \pi_i^\star}{2} \leq \frac{\mathcal{TV}_i(\alpha_i^\star)}{2} \leq \frac{\mathcal{TV}_i(S'_i)}{2}$, because by construction, no matter what S_i is, $\mathcal{TV}_i(S'_i) = \max_{S \subseteq G} \mathcal{TV}_i(S)$.

Therefore, in Case 3, $\mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i}, \mathcal{V})] \leq \frac{\mathcal{TV}_i(S'_i)}{2}$ as stated above.

Let us now lowerbound $\mathbb{E}[u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V})]$ in Case 3. First of all, notice that the present case implies that $REV(\Omega_i) \neq REV(\Omega'_i)$. By construction of σ'_i, this further implies that $REV(\Omega_i) < \mathbb{MEW}_i(\mathcal{K}_i)$ and $\Omega'_i = \arg\max_{\omega \in F(\mathcal{K}_i)} REV(\omega)$. Therefore as in Case 2(a), we have that $\mathbb{E}[u_i(\sigma'_i \sqcup \tau_{-i}, \mathcal{V})] = \frac{\mathcal{TV}_i(S'_i) + REV(\Omega'_i) - R'}{2}$. But now, since $REV(\Omega'_i) \geq R'$, we also have $\frac{\mathcal{TV}_i(S'_i)}{2} \leq \frac{\mathcal{TV}_i(S'_i) + REV(\Omega'_i) - R'}{2}$, as stated above.

Combining these three cases, we see that σ_i is knowingly dominated by σ'_i over $\Sigma^1_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$. ■

Claim 5. *Let* $(\mathbb{C}, I, TV, K, \mathbb{U})$ *be the individually monotone collusion system in Theorem 2.* $\forall i \in I$ *and* $\forall \sigma_i \in \Sigma_i^{2,\mathtt{IM}(i,TV_i,K_i)}$*, in Stage 1, the following two implications hold for* σ_i*:*

1. *i chooses S_i to be his favorite subset of goods, that is* $S_i = \arg\max_{S \subseteq G} TV_i(S)$.
2. *i does not "under-bid", that is he announces Ω_i such that* $REV(\Omega_i) \geq \mathbb{MEW}_i(K_i)$.

In Stage 2, the following two implications hold for σ_i*:*

1. *i answers YES whenever* $TV_i(\alpha_i^\star) \geq \pi_i^\star$*, and*
2. *i answers NO whenever* $TV_i(\alpha_i^\star) < \pi_i^\star$.

Proof of Claim 5. This claim follows directly from Definition 11 and Claim 4. ■

Comment. Note that we say nothing about whether the players will over-bid or not. In fact, besides $\mathcal{K}_i$, if player i has some Bayesian information about $\mathcal{TV}_j$, he may be able to compute for a subset of goods S that the probability that $\mathcal{TV}_j(S)$ is larger than a particular value, and announce a price for player j on subset S larger than $\min_{v \in \mathcal{K}_i} v_j(S)$, so that taking into account the probability that this price is rejected by j and the probability that it is accepted, the expected utility of i is higher than announcing $\min_{v \in \mathcal{K}_i} v_j(S)$. Therefore over-bid may not be a knowingly dominated strategy for player i over $\Sigma^1_{(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})}$ and $(\mathcal{C},\mathcal{I},\mathcal{TV},\mathcal{K},\mathcal{U})$. But as shown in the following proof, if a player over-bids, our result still holds, and thus we do not care whether over-bidding is knowingly dominated or not. In Part II of this paper, by envisioning an infinite penalty to the star player when one of his proposal is rejected, we can make sure that the players will not over-bid.

We are finally ready to prove Theorem 2, that is: for our mechanism $\mathcal{M}$, $\forall$ integers n and m, $\forall$ individually monotone collusion systems $(\mathbb{C}, I, TV, K, \mathbb{U})$ for n players and m goods, and $\forall$ knowingly surviving plays σ of $\mathcal{M}$:

$$\mathbb{E}[REV] + \mathbb{E}[SW] \geq \frac{\mathbb{MEW}(K_I)}{2}.$$

To this end, denote by $*$ the independent player "realizing" our benchmark: that is,

$$* = \arg\max_{i \in I} \mathbb{MEW}_i(K_i).$$

(Notice that the players $*$ and $\star$ need not to coincide.)

According to Claim 4, in any surviving play, $*$ announces an outcome Ω_* such that $REV(\Omega_*) \geq \mathbb{MEW}_*(K_*)$. Now, since by definition the star player is the one who announces an outcome with the largest revenue, we have $R_\star \geq REV(\Omega_*)$, and thus $R_\star \geq \mathbb{MEW}(K_I) = \mathbb{MEW}_*(K_*)$. To prove Theorem 1 we distinguish two cases.

(1) $\star = *$.

In this case, as player $*$ is independent, so is player $\star$, and thus $\star \notin C$ for all collusive sets $C \in \mathbb{C}$. Therefore Claim 3 and 5 guarantees that every $i \neq \star$ answers YES in Stage 2 if and only if $TV_i(\alpha_i^\star) \geq \pi_i^\star$. Accordingly, the following inequality holds for $\mathcal{M}$'s expected social welfare:

$$\mathbb{E}[SW] = \frac{TV_\star(S_\star)}{2} + \frac{\sum_{i:YES} TV_i(\alpha_i^\star)}{2} \geq \frac{\sum_{i:YES} TV_i(\alpha_i^\star)}{2} \geq \frac{\sum_{i:YES} \pi_i^\star}{2}.$$

At the same time,

$$\mathbb{E}[REV] = \frac{\sum_{i:NO} \pi_i^\star}{2} + \frac{R' + \sum_{i:NO} \pi_i^\star}{2} \geq \frac{\sum_{i:NO} \pi_i^\star}{2}.$$

Thus

$$\mathbb{E}[SW] + \mathbb{E}[REV] \geq \frac{\sum_{i:YES} \pi_i^\star + \sum_{i:NO} \pi_i^\star}{2} = \frac{R_\star}{2} \geq \frac{\mathbb{MEW}(K_I)}{2}.$$

(2) $\star \neq *$.

In this case, $* \in -\star$, thus —since player $*$ is independent— $R' \geq REV(\Omega_*) \geq \mathbb{MEW}(K_I)$. Therefore $\mathcal{M}$'s expected revenue is

$$\mathbb{E}[REV] = \frac{\sum_{i:NO} \pi_i^\star}{2} + \frac{R' + \sum_{i:NO} \pi_i^\star}{2} \geq \frac{R'}{2} \geq \frac{\mathbb{MEW}(K_I)}{2}.$$

Because

$$\mathbb{E}[SW] = \frac{TV_\star(S_\star)}{2} + \frac{\sum_{i:YES} TV_i(\alpha_i^\star)}{2} \geq 0,$$

we have

$$\mathbb{E}[SW] + \mathbb{E}[REV] \geq \frac{\mathbb{MEW}(K_I)}{2}.$$

Combining these two cases, Theorem 2 follows.

Q.E.D.

References

[MV07.a] S. Micali and P. Valiant, Revenue in Truly Combinatorial Auctions and Adversarial Mechanism Design, Technical Report, MIT-CSAIL- TR-2008-039, 2008.

(Previous versions include: MIT-CSAIL-TR-2007-052, 2007; submission to FOCS 2007, submission to SODA 2008; and the versions deposited with the Library of Congress in 2007)

[MV07.b] S. Micali and P. Valiant, Leveraging Collusion in Combinatorial Auctions, Unpublished Manuscript. 2007

[OR94] M. J. Osborne and A. Rubinstein, A Course in Game Theory, MIT Press, 1994.

[BLP06] M. Babaioff, R. Lavi and E. Pavlov, Single-Value Combinatorial Auctions and Implementation in Undominated Strategies, SODA'06, pages 1054-1063, 2006.

[ILM05] M. Lepinski, S. Izmalkov and S. Micali, Rational secure computation and ideal mechanism design. FOCS'05, pages 585-595, 2005.