Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2011-009 February 16, 2011

Mechanism Design With Approximate Player Types

Alessandro Chiesa, Silvio Micali, and Zeyuan Allen Zhu

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

Mechanism Design with Approximate Player Types

Alessandro Chiesa[*] MIT CSAIL
Silvio Micali[†] MIT CSAIL
Zeyuan Allen Zhu[‡] MIT CSAIL

February 16, 2011

Abstract

We investigate mechanism design when the players *do not* exactly know their types, but have instead only partial information about them.

[*]Email: alexch@csail.mit.edu.
[†]Email: silvio@csail.mit.edu.
[‡]Email: zeyuan@csail.mit.edu.

1 Introduction

The players' *true types* ultimately determines their utilities, and the traditional assumption in mechanism design is that *each player i knows his own true type.* (The true type of i is consistently denoted by $\mathtt{t}_i$. The space of all possible types for i is assumed to be finite and known to every one, and is consistently denoted by Θ_i.)

In this paper, we investigate mechanism design when the players *do not* exactly know their types, but have instead only partial information about them.

How to model this "type uncertainty?" There are for course several possible ways. Perhaps the most obvious one is via *"individual Bayesians."* That is, one could assume that each player i knows the distribution $\mathcal{D}_i$ over Θ_i from which $\mathtt{t}_i$ has been actually selected. We wish instead to explore a more extreme model.

A General Set-Theoretic Model of Type Uncertainty Our model is *set theoretic* (and indeed inspired by the "external-knowlege" model of [**?**]). At the highest level, we assume that each player i knows that $\mathtt{t}_i$ belongs to a specific subset K_i of Θ_i, but has no further information. Note that this assumption is "safer" (i.e., weaker) than assuming an individual Bayesian $\mathcal{D}_i$: that is, the latter assumption implies the former (but not viceversa).[1] In fact, ours in not even an assumption: the fact that i knows that $\mathtt{t}_i \in K_i$ does not imply that i has a given level of knowledge about his own type. For instance, if $K_i = \Theta_i$, then i knows nothing about his own type; if $K_i = \{\mathtt{t}_i\}$, then i knows precisely his own type, and we are in the traditional model.

A Special Sub-Model: Type Intervals We use the above general model to provide some key definitions, but we shall use a more specific instantiation of it for designing and analyzing auction mechanisms. In an auction of (in general) multiple goods, a player type is a *valuation*, that is, a function v mapping each possible subset of the goods for sale to a non-negative number $v(S)$. In an auction setting, we specialize our set-theoretic model of type uncertainty as "δ approximation." That is, for simplicity, we assume that there is a fixed real $\delta \in (0,1)$, common knowledge to everyone, such that for every subset S of the goods, each player i knows that his own true value for S, $\mathtt{t}_i(S)$, lies in the interval $x - \delta x$ and $x + \delta x$. We refer to δ as the *uncertainty factor*. For instance, if $\delta = .1$, then every one knows that each player knows his own valuation with a 10% accuracy.

Note that, although the set in which each $\mathtt{t}_i(S)$ belongs has a very simple and appealing structure, i has no information that enables him to further narrow down $\mathtt{t}_i(S)$. When choosing how to bid he only knows δ and his own interval for each subset of the goods. For what he is concerned, an *adversary* selects $\mathtt{t}_i$ in the proper interval at the start of the auction, but her selection is *secret*: i will realize only later on (e.g., a day or a year after the auction is over) what his true utility really is).

Why? Our model aims at including some existing (if not frequent) aspects of reality. For instance, no one would be too surprised if a firm about to participate to an auction asks different employees to figure out the firm's value for the same subset of the goods, and different values are ultimately reported. Our model aims also at simplicity, at least at an initial stage of this research. (For instance, the one-fits-all nature of the uncertainty factor δ may be removed later on.) More importantly, by investigating a model of an adversarial nature we may be able to prove less, but our results will then be guaranteed to hold for a larger variety of more "realistic/benign" models. We personally prefer this to getting stronger results by making stronger assumptions about the type uncertainty of the players. But there is definitely room for other approaches!

Non-traditional Model, Non-traditional Results One may imagine that the presence of an uncertainty factor would not cause any major changes in mechanism design. Perhaps, the worst that might happen is that the performance of our established mechanisms degrades a bit as a function of δ. For instance, one might hypothesize that the VCG mechanism continues to work in dominant strategies, although guaranteeing only a fraction —say— $(1-\delta)$, or $(1-3\delta)$, or $(1-\delta)^2$ than the (true!) maximum social welfare.

We shall prove, however, that this is not the case. Set-theoretic type uncertainty will cause major changes both in design and the performance that one may be able to guarantee.

[1] Indeed, if $\mathtt{t}_i$ has been drawn from $\mathcal{D}_i$, then one can set K_i to be $\mathcal{D}_i$'s support.

2 Notions of Mechanism Design with Approximate Player Types

We introduce the basic definitions needed for a formal treatment of mechanism design in our model, pointing out how these definitions"collapse" to standard ones when all the players' knowledge is exact.

All sets are finite in this paper (and in real life as well!). If we use the word "finite", it is for emphasis only.

2.1 Contexts, Mechanisms and Games with Approximate Types

The most basic notion is that of a pre-context: it keeps track of the number of players, the set of possible types for each of the players, the set of all possible outcomes, the utility function for each of the players, and a promise on each player's "quality of approximation" about his own type:

Definition 2.1. *A* **pre-context** *is a tuple* $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ *such that:*

- N *is a finite set of* $n \stackrel{\text{def}}{=} |N|$ *players;*
- $\Theta = \Theta_1 \times \cdots \times \Theta_n$ *where each* Θ_i *is a finite set of types for player* i*;*
- Ω *is a finite set of outcomes;*
- $U = (U_1, \ldots, U_n)$ *where each* $U_i \colon \Theta_i \times \Omega$ *is the utility function of player* i*; and*
- $Q = Q_1 \times \cdots \times Q_n$ *where each* Q_i *is a subset of* 2^{Θ_i}*.*

The "new" component in our pre-context (as compared to the usual definition) is the *approximation quality* Q that players are promised to have: for each player i, Q_i is an explicit collection of subsets of Θ_i; each subset in Q_i a possible *approximate type* for player i. Our definition reduces to the usual definition of a pre-context whenever each Q_i is the collection of all singletons in 2^{Θ_i} (i.e., the only "allowed" approximation is the perfect approximation).

A pre-context becomes a context when it is augmented with a *specific* set of players, i.e., a specific profile of approximate types:

Definition 2.2. *A* **context** *is a tuple* $\mathsf{Cnt} = (N, \Theta, \Omega, U, Q, K)$ *such that:*

- $(N, \Theta, \Omega, U, Q)$ *is a pre-context and*
- $K = K_1 \times \cdots \times K_n$ *where each* $K_i \in Q_i$*; we refer to* Q_i *as the approximate type for player* i*.*

Note that our notion of approximate type is *set-theoretic*, that is, each K_i is interpreted as comprising all the possible types that player i believes to be possible (though he is not sure which type is the true one). In particular, we can model any Bayesian on a player's type by considering its support.

Because each K_i is promised to be one of the sets in Q_i, the collection of subsets Q_i gives an idea of how knowledgeable player i will be about his own types. For example, Q_i may consist of all subsets of 2^{Θ_i} of cardinality that is at most some constant, say, 10.

As usual, a mechanism specifies the players' strategies and how these strategies determine outcomes:

Definition 2.3. *A* **mechanism** *is a tuple* $M = (\Sigma, F)$ *such that:*

- $\Sigma = \Sigma_1 \times \cdots \times \Sigma_n$ *where each* Σ_i *is a finite set of strategies for player* i*; and*
- $F \colon \Sigma \to \Omega$ *is the outcome function.*

We will denote pure strategies with Latin letters (such as s and t), and possibly mixed strategies with Greek letters (such as σ and τ).

Definition 2.4. *A* **game** *is a pair* $\mathcal{G} = (\mathsf{Cnt}, M)$ *where* Cnt *is a context and* M *is a mechanism.*

As our results concern mechanism *design*, we need to understand the strategic behaviors of rational players. Therefore, in the next subsection we start to discuss notions of dominance in the case of approximate types.

2.2 Dominance with Approximate Types

Throughout this section, we fix a game $\mathcal{G} = (N, \Theta, \Omega, U, Q, K, \Sigma, F)$. First, we define the three usual notions of dominance in the approximate player type case:

Definition 2.5 (Dominance)**.** *Fix a player* $i \in N$ *and an approximate type* $K_i \in Q_i$*. Then,*

- *We say that a strategy* $\sigma_i \in \Delta(\Sigma_i)$ **very-weakly dominates** *a strategy* $\sigma'_i \in \Delta(\Sigma_i)$ *for player* i *with respect to the approximate type* K_i*, in symbols* $\sigma_i \underset{i,K_i}{\overset{\text{vw}}{\succ}} \sigma'_i$*, if (and only if):*

$$\forall\, \theta_i \in K_i\,,\, \forall\, \tau_{-i} \in \Delta(\Sigma_{-i})\,:\, \mathbb{E}U_i(\theta_i, \sigma_i \sqcup \tau_{-i}) \geq \mathbb{E}U_i(\theta_i, \sigma'_i \sqcup \tau_{-i}) \enspace .$$

- *We say that a strategy* $\sigma_i \in \Delta(\Sigma_i)$ **weakly dominates** *a strategy* $\sigma'_i \in \Delta(\Sigma_i)$ *for player* i *with respect to the approximate type* K_i*, in symbols* $\sigma_i \underset{i,K_i}{\overset{\text{w}}{\succ}} \sigma'_i$*, if (and only if):*

$$\begin{aligned}&\forall\, \theta_i \in K_i\,,\, \forall\, \tau_{-i} \in \Delta(\Sigma_{-i})\,:\, \mathbb{E}U_i(\theta_i, \sigma_i \sqcup \tau_{-i}) \geq \mathbb{E}U_i(\theta_i, \sigma'_i \sqcup \tau_{-i}) \enspace , \textit{ and}\\ &\exists\, \theta_i \in K_i\,,\, \exists\, \tau_{-i} \in \Delta(\Sigma_{-i})\,:\, \mathbb{E}U_i(\theta_i, \sigma_i \sqcup \tau_{-i}) > \mathbb{E}U_i(\theta_i, \sigma'_i \sqcup \tau_{-i}) \enspace .\end{aligned}$$

- *We say that a strategy* $\sigma_i \in \Delta(\Sigma_i)$ **strictly dominates** *a strategy* $\sigma'_i \in \Delta(\Sigma_i)$ *for player* i *with respect to the approximate type* K_i*, denoted* $\sigma_i \underset{i,K_i}{\overset{\text{s}}{\succ}} \sigma'_i$*, if (and only if):*

$$\forall\, \theta_i \in K_i\,,\, \forall\, \tau_{-i} \in \Delta(\Sigma_{-i})\,:\, \mathbb{E}U_i(\theta_i, \sigma_i \sqcup \tau_{-i}) > \mathbb{E}U_i(\theta_i, \sigma'_i \sqcup \tau_{-i}) \enspace .$$

Remark 2.6. Let us observe the relationship between the three notions of dominance and the traditional ones, that is, when $K_i = \{\mathsf{t}_i\}$ for all players i. Denoting by "$\longrightarrow$" logical implication, we predictably have:

$$\left(\sigma_i \underset{i,K_i}{\overset{\text{vw}}{\succ}} \sigma'_i\right) \longleftrightarrow \left(\forall\, \theta_i \in K_i \,:\, \sigma_i \underset{i,\{\theta_i\}}{\overset{\text{vw}}{\succ}} \sigma'_i\right) \enspace ,$$

and

$$\left(\sigma_i \underset{i,K_i}{\overset{\text{s}}{\succ}} \sigma'_i\right) \longleftrightarrow \left(\forall\, \theta_i \in K_i \,:\, \sigma_i \underset{i,\{\theta_i\}}{\overset{\text{s}}{\succ}} \sigma'_i\right) \enspace .$$

However, perhaps less predictably, for the notion of weak dominance we have slightly different relation to the usual notion, namely:

$$\left(\sigma_i \underset{i,K_i}{\overset{\text{w}}{\succ}} \sigma'_i\right) \underset{\not\longrightarrow}{\longleftarrow} \left(\forall\, \theta_i \in K_i \,:\, \sigma_i \underset{i,\{\theta_i\}}{\overset{\text{w}}{\succ}} \sigma'_i\right) \enspace .$$

This is so because the existential quantifier of the second condition of weak dominance interferes with the universal quantification of θ_i within K_i.[2]

The three notions of Definition 2.5 give rise as usual to three corresponding sets of "undominated" strategies and three corresponding sets of "dominant" strategies, both also in the approximate player type case.

Definition 2.7 (Undominated Strategies)**.** *Fix a player* $i \in N$ *and an approximate type* $K_i \in Q_i$*. Then,*

- *The set of not-very-weakly-dominated strategies of* i *with respect to* K_i *is defined as*

$$\mathsf{NVWDed}_i(K_i) \overset{\text{def}}{=} \left\{\sigma_i \in \Delta(\Sigma_i) : \nexists\, \tau_i \in \Delta(\Sigma_i) \textit{ s.t. } \tau_i \underset{i,K_i}{\overset{\text{vw}}{\succ}} \sigma_i\right\} \enspace .$$

- *The set of not-weakly-dominated strategies of* i *with respect to* K_i *is defined as*

$$\mathsf{NWDed}_i(K_i) \overset{\text{def}}{=} \left\{\sigma_i \in \Delta(\Sigma_i) : \nexists\, \tau_i \in \Delta(\Sigma_i) \textit{ s.t. } \tau_i \underset{i,K_i}{\overset{\text{w}}{\succ}} \sigma_i\right\} \enspace .$$

- *The set of not-strictly-dominated strategies of* i *with respect to* K_i *is defined as*

$$\mathsf{NSDed}_i(K_i) \overset{\text{def}}{=} \left\{\sigma_i \in \Delta(\Sigma_i) : \nexists\, \tau_i \in \Delta(\Sigma_i) \textit{ s.t. } \tau_i \underset{i,K_i}{\overset{\text{s}}{\succ}} \sigma_i\right\} \enspace .$$

[2] In a sense, the notion of weak dominance intends to capture the "weakest" condition for which a strategy σ'_i should be discarded in favor of σ_i. Assuming that σ'_i is already weakly dominated by σ_i with respect to a given approximate type K_i, it suffices for there to *exist one* type $\theta_i \in K_i$ for which σ'_i is strictly worse than σ_i for some strategy sub-profile τ_{-i} in order to discard σ'_i in favor of σ_i. (Requiring it for *all* types $\theta_i \in K_i$ would strengthen the condition for dominance in a way that would make it impossible to eliminate strategies that should in fact be "rationally" discarded.)

Definition 2.8 (Dominant Strategies)**.** *Fix a player $i \in N$ and an approximate type $K_i \in Q_i$. Then,*

- *The set of very-weakly-dominant strategies of i with respect to K_i is defined as*

$$\mathsf{VWDnt}_i(K_i) \stackrel{\text{def}}{=} \left\{ \sigma_i \in \Delta(\Sigma_i) : \forall\, \tau_i \in \Delta(\Sigma_i)\,,\ \tau_i \overset{\text{vw}}{\underset{i,K_i}{\succ}} \sigma_i \right\} \ .$$

- *The set of weakly-dominant strategies of i with respect to K_i is defined as*

$$\mathsf{WDnt}_i(K_i) \stackrel{\text{def}}{=} \left\{ \sigma_i \in \Delta(\Sigma_i) : \forall\, \tau_i \in \Delta(\Sigma_i)\,,\ \tau_i \overset{\text{w}}{\underset{i,K_i}{\succ}} \sigma_i \right\} \ .$$

- *The set of strictly-dominant strategies of i with respect to K_i is defined as*

$$\mathsf{SDnt}_i(K_i) \stackrel{\text{def}}{=} \left\{ \sigma_i \in \Delta(\Sigma_i) : \forall\, \tau_i \in \Delta(\Sigma_i)\,,\ \tau_i \overset{\text{s}}{\underset{i,K_i}{\succ}} \sigma_i \right\} \ .$$

Note that all the above definitions reduce to the usual ones in the case where all the players' types are exact (i.e., K_i is a singleton).

Notation. Whenever we ignore the subscript we denote the Cartesian product of the set among players. For instance, $\mathsf{NWDed}(K) \stackrel{\text{def}}{=} \mathsf{NWDed}_1(K_1) \times \cdots \times \mathsf{NWDed}_n(K_n)$.

2.2.1 Relation of Dominance with Approximate Types

We prove an intuitive lemma that we will use several times in our proofs: whenever the approximate type of a player becomes larger, the sets of undominated strategies grow, while the sets of dominant strategies shrink. Formally:

Lemma 2.9. *Fix two games $\mathcal{G}$ and $\mathcal{G}'$ that differ only in the respective approximate type profiles K and K'. For every player $i \in N$, if $K_i \subseteq K_i'$, then the following set inclusions hold:*

$$\begin{array}{cc} \mathsf{NVWDed}_i(K_i) \subseteq \mathsf{NVWDed}_i(K_i') & \mathsf{VWDnt}_i(K_i) \supseteq \mathsf{VWDnt}_i(K_i') \\ \mathsf{NWDed}_i(K_i) \subseteq \mathsf{NWDed}_i(K_i') & \mathsf{WDnt}_i(K_i) \supseteq \mathsf{WDnt}_i(K_i') \\ \mathsf{NSDed}_i(K_i) \subseteq \mathsf{NSDed}_i(K_i') & \mathsf{SDnt}_i(K_i) \supseteq \mathsf{SDnt}_i(K_i') \end{array}$$

Proof. We proof the top set inclusion in each set. The proofs for the other four set inclusions are similar. Suppose that some strategy $\sigma_i \in \mathsf{NVWDed}_i(K_i)$, i.e., there does not exist a strategy $\tau_i \in \Delta(\Sigma_i)$ such that $\tau_i \overset{\text{vw}}{\underset{i,K_i}{\succ}} \sigma_i$. However, $K_i \subseteq K_i'$ implies the implication $\left(\tau_i \overset{\text{vw}}{\underset{i,K_i'}{\succ}} \sigma_i\right) \to \left(\tau_i \overset{\text{vw}}{\underset{i,K_i}{\succ}} \sigma_i\right)$, and thus we are forced to conclude that $\sigma_i \in \mathsf{NVWDed}_i(K_i')$ as well. Now suppose that some strategy $\sigma_i \in \mathsf{VWDnt}_i(K_i')$, i.e., for every strategy $\tau_i \in \Delta(\Sigma_i)$ it the case that $\sigma_i \overset{\text{vw}}{\underset{i,K_i'}{\succ}} \tau_i$. Again, as $K_i \subseteq K_i'$ implies the implication $\left(\sigma_i \overset{\text{vw}}{\underset{i,K_i'}{\succ}} \tau_i\right) \to \left(\sigma_i \overset{\text{vw}}{\underset{i,K_i}{\succ}} \tau_i\right)$, and thus we are forced to conclude that $\sigma_i \in \mathsf{VWDnt}_i(K_i)$ as well. $\square$

2.3 Implementation with Approximate Types

We can now state what it means for a mechanism to "implement" a social property relative to a certain class of contexts with approximate types. A social property is simply a predicate over the players' types and an outcome distribution:

Definition 2.10. *A* **social property** *is a function* $\Pi \colon \Theta \times \Delta(\Omega) \to \{0, 1\}$.

Notation. To avoid repetition, we denote by $\mathbb{S}$ any of the following "solution concepts": NVWDed, NWDed, NSDed, VWDnt, WDnt, and SDnt. We use $\mathbb{S}_i$ to denote the corresponding component, e.g., NWDed_i for the the i-th component of NWDed.

Definition 2.11 (Implementation)**.** *Let M be a mechanism, $\mathcal{C}$ a class of contexts, Π a social property, and $\mathbb{S}$ one of the above solution concepts. We say that:*

- M **partially implements Π in $\mathbb{S}$ strategies w.r.t.** $\mathcal{C}$ *if, for every context* $\mathsf{Cnt} = (N, \Theta, \Omega, U, Q, K) \in \mathcal{C}$,

$$\exists\, \sigma \in \mathbb{S}(K) \;\; \forall\, \theta \in K \; : \; \Pi(\theta, F(\sigma)) = 1 \; .$$

- M **fully implements Π in $\mathbb{S}$ strategies w.r.t.** $\mathcal{C}$ *if, for every context* $\mathsf{Cnt} = (N, \Theta, \Omega, U, Q, K) \in \mathcal{C}$,

$$\forall\, \sigma \in \mathbb{S}(K) \;\; \forall\, \theta \in K \; : \; \Pi(\theta, F(\sigma)) = 1 \; .$$

Remark 2.12. Note that, in extending the classical notion of partial and full implementation to the case of approximate types, we actually choose a *worst-case perspective*. That is, rather than requiring the property to hold at only one possible type profile in K, we believe it crucial to require that the property hold at *every combination of possible types*.

Clearly, partial implementation is weaker than full implementation, even with approximate types. Accordingly, we shall use partial implementation only to prove impossibility results.

Example 2.13. Let us spell out two of the solution concepts about which we prove our theorems:

- When we set $\mathbb{S}_i(K_i) \stackrel{\text{def}}{=} \mathsf{NWDed}_i(K_i)$, we obtain the notions of **partial and full implementation in not-weakly-dominated strategies**, respectively.
- When we set $\mathbb{S}_i(K_i) \stackrel{\text{def}}{=} \mathsf{VWDnt}_i(K_i)$, we obtain the notions of **partial and full implementation in very-weakly-dominant strategies**, respectively.

2.4 A Revelation Principle with Approximate Types

We prove an analogue of the Revelation Principle in the setting of approximate player types. A version of the classical Revelation Principle states that, when considering partial implementations in very-weakly-dominant strategies, it suffices to consider (very-weakly DST) mechanisms where strategy sets are identical to type spaces Θ, and reporting the true type is a very-weakly-dominant strategy.

Lemma 2.14 (Revelation Principle)**.** *Fix any pre-context* $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$, *mechanism* $M = (\Sigma, F)$, *and social choice function* $W\colon Q \to \Delta(\Omega)$ *such that for every* $K \in Q$,

$$\exists\, \sigma \in \mathsf{VWDnt}(K) \text{ s.t. } F(\sigma) = W(K) \; .$$

Then there exists a direct *mechanism* $M' = (\Sigma', F')$ *for which* $\Sigma' = Q$ *and for every* $K \in Q$,

$$K \in \mathsf{VWDnt}'(K) \text{ and } F'(K) = W(K) \; .$$

In other words, the direct mechanism M' is a very-weakly dominant strategy truthful (DST) *one that guarantees reporting the truth K is very-weakly-dominant, and yields the same outcome. Here VWDnt' is the very-weakly-dominant strategies of the game with M'.*

Proof. There exists a function $f\colon Q \to \Delta(\Sigma)$ such that $f(K)$ is the lexicographically first strategy $\sigma \in \mathsf{VWDnt}(K)$ for which $M(\sigma) = W(K)$. Consider the mechanism $M' = (\Sigma', F')$ for which $\Sigma' \stackrel{\text{def}}{=} Q$ and $F'(K) \stackrel{\text{def}}{=} F(f(K))$ for every $K \in Q$.

Fix any $K \in Q$ and suppose that $\exists \sigma \in \mathsf{VWDnt}(K)$ for which $M(\sigma) = W(K)$, and let it be the lexicographically first one. By construction, $F'(K) = F(f(K)) = F(\sigma) = W(K)$. We are left to prove $K \in \mathsf{VWDnt}'(K)$. Indeed, consider any alternative strategy $K_i' \in Q$ for player i. Then, for every $\theta_i \in K_i$ and every $K_{-i} \in Q_{-i}$,

$$\begin{aligned} \mathbb{E}U\big(\theta_i, F'(K_i \sqcup K_{-i})\big) &= \mathbb{E}U\big(\theta_i, F(f(K_i \sqcup K_{-i}))\big) \\ \geq \mathbb{E}U\big(\theta_i, F(f(K_i' \sqcup K_{-i}))\big) &= \mathbb{E}U\big(\theta_i, F'(K_i' \sqcup K_{-i})\big) \; , \end{aligned}$$

where the inequality comes from the fact that $f(K_i \sqcup K_{-i}) \in \mathsf{VWDnt}(K)$. This completes the proof. $\square$

Corollary 2.15. *Fix any pre-context* $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ *and a mechanism* $M = (\Sigma, F)$ *that partially implements some social property* $\Pi\colon \Theta \times \Delta(\Omega) \to \{0, 1\}$ *in very-weakly-dominant strategies. We can construct a direct (very weakly DST) mechanism* M' *such that reporting the the true* K *is very-weakly-dominant and implements* Π.

Remark 2.16. An analogue of Lemma 2.14 also holds for implementation in ex-post Nash, and therefore all of our impossibility results in this subsection can be extended to implementation in ex-post Nash equilibria. (See, for example, Corollary 9.26 in Algorithmic Game Theory book for the idea.)

3 Approximate-Type Combinatorial Auctions and Social Welfare Benchmarks

The deinitions of this section are quite straightforward extensions of the classical ones, and can be easily guessed by people familiar with classical notions.

We begin our study of mechanism design with approximate player types by focusing on a specific class of pre-contexts and a specific social property: we focus on combinatorial auctions and consider the goal of achieving a given "ϵ-fraction" of the maximum social welfare.

Combinatorial auctions are very well studied in mechanism design and, in particular, enjoy very strong positive results in terms of achieving social efficiency: for example, the VCG mechanism guarantees *maximum* social welfare for any number of players and any number of goods. Intuitively, our results say that certain widely-used "classical" game-theoretic solution concepts (such as the implementation in very-weakly-dominant strategies of the VCG mechanism) are *not "robust" with respect to approximate players types* when attempting to maximize social welfare, even when the approximation of the types is as good as *any* positive constant; we stress that our negative results are completely generic, and not particular to VCG. In particular, our results cast some doubt on the meaningfulness of, for example, the VCG mechanism. After all, how could the players be required to know their types exactly? Assuming that it is indeed the case appears to be a very strong assumption to us.

Let us be more precise by introducing the appropriate notions. First, we define a general class of widely-studied pre-contexts called quasi-linear pre-contexts:

Definition 3.1. *Fix* $n, B \in \mathbb{N}$. *A* **quasi-linear pre-context** *with* n *players and maximum valuation* B *is a tuple* $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ *such that:*

- $(N, \Theta, \Omega, U, Q)$ *is a pre-context;*
- $n = |N|$*;*
- $\Omega = \mathcal{A} \times \mathbb{R}^n$ *for some finite set* $\mathcal{A}$*;*
- *for each* $i \in N$, $\Theta_i \stackrel{\text{def}}{=} \{0, 1, \dots, B\}^{\mathcal{A}}$*; and*
- *for each* $i \in N$, $U_i\big(\theta_i, (A, P)\big) \stackrel{\text{def}}{=} \theta_i(A) - P$.

In other words, in a quasi-linear pre-context, the outcome set has a specific structure: it consists of pairs (A, P) where A is called a *social alternative* and $P = (P_1, \dots, P_n)$ is a *price profile*. Then, a player's types are assumed to be integer valuations between 0 and some bound B over the *alternatives set* $\mathcal{A}$, and his utility are simply the difference between the valuation on the given alternative minus the price.

The technical results in this paper concern a specific kind of quasi-linear pre-context, which is called a combinatorial auction pre-context:

Definition 3.2. *Fix* $n, m, B \in \mathbb{N}$. *A* **combinatorial auction pre-context** *with* n *players,* m *goods, and maximum valuation* B *is a pre-context* $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ *such that, letting* $G \stackrel{\text{def}}{=} \{1, \dots, m\}$, *the following properties hold:*

- $(N, \Theta, \Omega, U, Q)$ *is a quasi-linear pre-context with* n *players and maximum valuation* B*;*
- *the alternatives set* $\mathcal{A}$ *is equal to the set of all partitions of* G *in* $n+1$ *sets* $A_0, \dots, A_n$*; and*
- *each* $U_i(\theta_i, \omega) = \theta_i(A) - P_i = \theta_i(A_i) - P_i$*; and*
- *each* $\theta_i(\emptyset) = 0$.

Recall that, in a quasi-linear pre-context the set of outcomes is $\Omega = \mathcal{A} \times \mathbb{R}^n$ for some finite set of alternatives $\mathcal{A}$. In a combinatorial auction, the finite set of alternatives has a specific form: it is the the set of all *allocations* of the m goods in G among $n+1$ players, where the "0-th player" represents the goods that are not allocated. Then, each player's utility function is assumed to depend *only* on his own allocation. These additional constraints on the quasi-linear pre-context are responsible for the "combinatorial flavor" of the auction.

Next, we define the notion of social welfare for a quasi-linear pre-context, which represents the total amount of "happiness" among the players, specifically the sum of each player's individual happiness, when given a certain outcome:

Definition 3.3. *Let* $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ *be a quasi-linear pre-context. For every type profile* $\theta \in \Theta$ *and outcome* $\omega = (A, P) \in \Omega$*, the* **social welfare** *of* ω *with respect to* θ *is*

$$\mathrm{SW}(\theta, \omega) \stackrel{\text{def}}{=} \sum_{i \in N} \theta_i(A) \ .$$

Having defined the social welfare function, we can now state precisely the corresponding social property that attempts to maximize it. For each $\epsilon \in [0, 1]$, we consider the social property ϵ-MSW that attempts to achieve at least an ϵ fraction of the maximum social welfare (in expectation, because the output of a mechanism could be a probability distribution over outcomes):

Definition 3.4. *For any* $\epsilon \in [0, 1]$*, the* ϵ-MSW *is the social property so defined: for all* $\theta \in \Theta$ *and all* $W \in \Delta(\Omega)$*,*

$$\epsilon\text{-MSW}(\theta, W) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } \mathbb{E}\big[\mathrm{SW}(\theta, W)\big] \geq \epsilon \cdot \max_{\omega \in \Omega} \mathrm{SW}(\theta, \omega) \\ 0 & \text{otherwise} \end{cases} .$$

Accordingly, 1*-MSW is the* **maximum social welfare** *property.*

The results that we will prove will be characterizations of which $\epsilon \in (0, 1)$ allow or do not allow for implementation of ϵ-MSW in a given solution concept (e.g., very-weakly-dominant strategies or not-weakly-dominated strategies) within combinatorial auctions.

As we are interested in exploring the relation between the "quality" of the players' approximate types and the difficulty of implementing certain social properties, we introduce an explicit "approximation parameter" that we call δ. The parameter δ is to be interpreted as an upper bound on the coarseness of the players' approximations to their types; for example, we can think of $\delta = 0.1$ to mean that all the players have approximation of their true valuations up to a 0.1 multiplicative approximation (but, of course, players are allowed to have more knowledge than that, e.g., perfect approximation of their types). We are therefore interested in considering classes of quasi-linear pre-contexts with an explicit approximation parameter, therefore motivating the following definition:

Definition 3.5. *Fix* $n, B \in \mathbb{N}$ *and* $\delta \in (0, 1)$*. A* **δ-approximate quasi-linear pre-context** *with* n *players and maximum valuation* B *is a pre-context* $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ *such that:*

- $(N, \Theta, \Omega, U, Q)$ *is a quasi-linear pre-context with* n *players and maximum valuation* B*; and*
- *for each* $K_i \in Q_i$*, we must have for all* $A \in \mathcal{A}$*, there is some* $x_{i,A} \in \mathbb{R}$ *for which* $K_i(A) \subseteq \big[(1-\delta)x_{i,A}, (1+\delta)x_{i,A}\big] \cap \mathbb{Z}$.

We denote by $\mathsf{QL}(n, B, \delta)$ *the pre-context class consisting of all* δ*-approximate quasi-linear pre-contexts with* n *players and maximum valuation* B.

We are now finally able to define the main object of study of this paper:

Definition 3.6. *Fix* $n, B \in \mathbb{N}$ *and* $\delta \in (0, 1)$*. A* **δ-approximate combinatorial auction** *with* n *players,* m *goods, and maximum valuation* B *is a context* $\mathsf{Cnt} = (N, \Theta, \Omega, U, Q, K)$ *such that:*

- $(N, \Theta, \Omega, U, Q)$ *is a* δ*-approximate quasi-linear pre-context with* n *players and maximum valuation* B*;*
- $(N, \Theta, \Omega, U, Q)$ *is a combinatorial auction pre-context with* n *players,* m *goods, and maximum valuation* B.

We denote by $\mathsf{CA}(n, m, B, \boldsymbol{\delta})$ *the pre-context class consisting of all* δ*-approximate combinatorial auction with* n *players,* m *goods, and maximum valuation* B.

Note that, as expected, a 0-approximate quasi-linear pre-context is a quasi-linear pre-context according to the usual definition, and a 0-approximate combinatorial auction is a combinatorial auction according to the usual definition

Remark 3.7. A definition similar to Definition 3.5 can be meaningfully made for *any* context, by simply requiring that the cardinality of each K_i is at most, say, 10 (and possibly by further requiring that the types in K_i are "clustered", according to some notion of distance, to avoid the realistically unlikely event that a player has in mind several "very different" types as possibilities). We do not use such a definition as the technical results in this paper solely concern combinatorial auctions.

4 Formal Statements of Our Results

For approximate types, we prove both positive and negative results about two meaningful and popular notions of implementation, namely in dominant strategies and in undominated strategies. (Recall that type-approximation factor δ is known to everybody, including the designer, which makes our negative results only stronger.)

4.1 Results about Implementation in Dominant Strategies

Let us begin by recalling that, in the case of $\delta = 0$, the well-known VCG mechanism guarantees *maximum* social welfare when all the players report their true types, and reporting the true type very-weakly dominates all other strategies. In our notation,

Fact 4.1. *Fix any $n, m, B \in \mathbb{N}$. The* VCG *mechanism partially implements* 1-MSW *in very-weakly-dominant strategies with respect to* $\mathsf{CA}(n, m, B, \mathbf{0})$.

Perhaps one might hope that, when the players have δ-approximate knowledge about their own types, the VCG (or a variant thereof) would continue to work in dominant strategies and return a fraction of the optimal welfare that is solely dependent on δ (e.g., $(1-\delta)$ or $(1-\delta)^2$). We show however that, even if we demand partial implementation in very-weakly-dominant the fraction of the maximum social welfare that one can guarantee cannot be bounded solely in terms of δ. That is true even in the simple case of a single-good auction, only a fraction of $\frac{1}{n}$ of the social welfare can be guaranteed. Formally,

Theorem 4.2. *Fix any $n, B \in \mathbb{N}$ and $\delta \in (0, 1)$ with $B > \frac{3-\delta}{2\delta}$. If M partially implements ϵ-*MSW *in very-weakly-dominant strategies with respect to* $\mathsf{CA}(n, 1, B, \boldsymbol{\delta})$*, then*

$$\epsilon \leq \frac{1}{n} + \frac{\lfloor \frac{3-\delta}{2\delta} \rfloor + 1}{B} .$$

We give the proof in the case of $m = 1$ in Section 5.1.1 and $m > 1$ in Section 5.1.2.

Notice that a fraction of $\frac{1}{n}$ of the social welfare can be trivially achieved, even when the types are completely unknown, by just assigning the good to a randomly chosen player! The situation becomes substantially worse in combinatorial auctions where there are multiple goods for sale. Indeed, we prove that any very-weakly-dominant strategy mechanism must lose an *additional* fraction of the social welfare that is exponential in the number of goods for sale. Formally,

Theorem 4.3. *Fix any $n, m, B \in \mathbb{N}$ and $\delta \in (0, 1)$ with $B > \frac{3-\delta}{2\delta}$. If M partially implements ϵ-*MSW *in very-weakly-dominant strategies with respect to* $\mathsf{CA}(n, m, B, \boldsymbol{\delta})$*, then*

$$\epsilon \leq \frac{1}{n2^{m-1}} + \frac{\lfloor \frac{3-\delta}{2\delta} \rfloor + 1}{B} .$$

Both of the impossibility results above trivially extend to stronger notions of implementation in dominant strategies. In particular,

Corollary 4.4. *Analogous negative results as Theorem 4.2 and Theorem 4.3 hold for full implementation in* VWDnt *strategies, partial/full implementation in* WDnt *strategies, and partial/full implementation in* SDnt *strategies.*

On a "mildly positive" side, we show that our results are not only tight (up to small constant factors), but can be matched via the strongest implementation in dominant strategies. Namely,

Theorem 4.5 (Trivial)**.** *Fix any $n, m, B \in \mathbb{N}$. There exists a mechanism M such that, for any $\delta \in (0,1)$, M fully implements ϵ-MSW in strictly-dominant strategies with respect to $\mathsf{CA}(n,m,B,\boldsymbol{\delta})$ for*

$$\epsilon \stackrel{\text{def}}{=} \frac{1}{n(2^m - 1)} \ .$$

Interestingly, the mechanism of "Theorem 4.5" is completely stupid! It simply picks a player at random, and awards him a random subset of the goods — ignoring all the input strategies.

One interpretation of the results above is the following: if one insists on implementation in very-weakly-dominant strategies, and the players' types are only approximate (and likely they are!), then the mechanism will fail in learning *any information* about the players from their dominant strategies.

4.2 Results about Implementation in Undominated Strategies

As we have seen the news is bad for implementation in dominant strategies when the players have approximate types. We shall now see that the news for implementation in undominated strategies is better, at least for auctions of a single good. Indeed, the second-price mechanism (i.e., the VCG mechanism for $m = 1$), performs quite well (although not in dominant strategies!). That is, it returns a fraction of the maximum social welfare that depends solely on δ.

Theorem 4.6. *Fix $n, m, B \in \mathbb{N}$. There exists a mechanism M (i.e., the VCG mechanism) such that, for any $\delta \in (0,1)$, M fully implements ϵ-MSW in not-weakly-dominated strategies with respect to $\mathsf{CA}(n,m,B,\boldsymbol{\delta})$ for*

$$\epsilon \stackrel{\text{def}}{=} \frac{(1-\delta)^2}{(1+\delta)^2} - \frac{2}{(1+\delta)^2 \text{MSW}} \ .$$

The proof of Theorem 4.6 is given in Section 6.1.1.

On the other hand, if our players can use only a finite number of random bits in order to generate their mixed strategies (which probably is the case!), then the performance of the second-price mechanism cannot be improved much via any deterministic mechanism. Moreover, flipping coins can at best marginally improve the performance guarantee. Namely,

Theorem 4.7. *Fix any $n, m, B \in \mathbb{N}$ and $\delta \in (0,1)$ with $B \geq \frac{1}{\delta}$.*

- *If a deterministic mechanism M fully implements ϵ-MSW in not-weakly-dominated strategies with respect to $\mathsf{CA}(n,1,B,\boldsymbol{\delta})$, then*
$$\epsilon \leq \frac{(1-\delta)^2}{(1+\delta)^2} + \frac{4}{B} \ .$$

- *If a randomized mechanism M fully implements ϵ-MSW in not-weakly-dominated strategies with respect to $\mathsf{CA}(n,1,B,\boldsymbol{\delta})$, then*
$$\epsilon \leq \frac{(1-\delta)^2 + \frac{4\delta}{n}}{(1+\delta)^2} + \frac{4}{B} \ .$$

The proof of Theorem 4.7 is given in Section 6.1.2.

The situation seems to change dramatically for auctions of multiple goods. For now we have been able to prove this explicitly for the VCG mechanism. Namely,

Theorem 4.8. *Fix any $n, m, B \in \mathbb{N}$ and $\delta \in (0,1)$. The VCG mechanism (breaking ties at random) cannot fully implement ϵ-MSW with respect to $\mathsf{CA}(n,m,B,\boldsymbol{\delta})$ for*

$$\epsilon \geq \Omega\left(\left(\frac{1-\delta}{1+\delta}\right)^{2^m-2}\right) \ .$$

The proof of this theorem will be given in the full version of this paper.

5 Proof Details: Implementation in Dominant Strategies

We study the relation between the approximation factor δ and implementations of ϵ-fraction of maximum social welfare in the three types of dominant strategies (i.e., those from Definition 2.8). In this section, we present the following results:

Section 5.1: An impossibility result for partial implementation in very-weakly-dominant strategies showing that, essentially, one cannot do better than $\frac{1}{n2^{m-1}}$-fraction of the maximum social welfare. Recall that this impossibility result implies the same for the stronger notion of full implementation in VWDnt, and the stronger notions of partial/full implementations in WDnt and SDnt.

Section 5.2: A mechanism that implements a $\frac{1}{n(2^m-1)}$-fraction of the maximum social welfare, under *any* solution concept — this is the mechanism that, ignoring the players' input strategies, chooses a player at random, an assigns him a random subset of the goods.

At high level, our tight results demonstrate that even partial implementations in dominant strategies completely fail at learning anything about the players.

5.1 Impossibility Result

We seek a negative result for maximizing social welfare in combinatorial auctions when considering partial implementation in very-weakly-dominant strategies. We consider separately single-good auctions (in Section 5.1.1) and multiple-good auctions (in Section 5.1.2). While the proof technique used in the two cases is similar, the intuition is clearer in the single-good case and thus we present them as separate results.

Here is where we will make use of the Revelation Principle for approximate player types from Corollary 2.15: it will suffice to rule out all mechanisms for which the strategy set is identified with the approximation quality and reporting the true approximate type is a very-weakly-dominant strategy — and thus we will restate both theorems accordingly.

5.1.1 Single-Good Auctions

We present our negative result for partial implementation of ϵ-MSW in very-weakly-dominant strategies within *single*-good auctions. Fix $n, B \in \mathbb{N}$ and consider any δ-approximate combinatorial auction pre-context $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ for n players, 1 good, and maximum valuation B.

Theorem (Part 1 of Theorem **??**, restated)**.** *Consider any (possibly randomized) mechanism $M = (\Sigma, F)$ such that $\Sigma = Q$ and, for every $K \in Q$, K is a very-weakly-dominant strategy of the game with pre-context $\mathsf{Pre\,Cnt}$, players with approximate types K, and mechanism M. If $B > \frac{3-\delta}{2\delta}$, then there exists an approximate type $K \in Q$ and a type profile $\mathsf{tv} \in K$ such that*

$$\mathbb{E}\big[\mathrm{SW}\big(\mathsf{tv}, F(K)\big)\big] \le \left(\frac{1}{n} + \frac{\lfloor (3-\delta)/2\delta \rfloor + 1}{B}\right) \mathrm{MSW}(\mathsf{tv}) \ .$$

The main idea of the proof is to give a characterization of a player's winning probabilities and expected prices, by exploiting the fact that reporting the true approximate type is a very-weakly-dominant strategy. Essentially, a player's winning probability "saturates" to a constant value whenever his reported approximate type is sufficiently high; this we prove as a separate lemma. After that we consider a world where all the player's approximate types are just above the saturation point, and identify a player that does not win too often (it must exist); invoking our characterization of the winning probability allows us to deduce that even if the player deviates to reporting a very high approximate type, he still wins with the same low probability. Finally, "switching worlds" to one where the player's true approximate type is indeed very high, we show that the mechanism could not have been one that guarantees a significant fraction of the maximum social welfare. Details follow.

Proof. For each $K \in Q$ and player $i \in N$, define $M_i^A(K)$ to be the probability that player i wins the good and $M_i^{\mathbb{E}P}(K)$ to be his expected price, whenever the reported strategy profile is K. We first prove the characterization on a player's winning probabilities and expected prices:

Lemma 5.1. *Fix player $i \in N$, $x \in \{0,1,\dots,B\}$, and $K_{-i} \in Q_{-i}$. If $x > \frac{3-\delta}{2\delta}$, then*

$$M_i^A(\mathsf{int}_\delta(x) \sqcup K_{-i}) = M_i^A(\mathsf{int}_\delta(x+1) \sqcup K_{-i}) \text{ and}$$
$$M_i^{\mathbb{E}P}(\mathsf{int}_\delta(x) \sqcup K_{-i}) = M_i^{\mathbb{E}P}(\mathsf{int}_\delta(x+1) \sqcup K_{-i}) \ .$$

Proof. Define $K_i \stackrel{\text{def}}{=} \mathsf{int}_\delta(x)$ and $K_i' \stackrel{\text{def}}{=} \mathsf{int}_\delta(x+1)$. Then:

- If player i has approximate type K_i then reporting K_i very-weakly dominates reporting K_i':

$$\begin{aligned}\forall\, \mathsf{tv}_i \in K_i : \quad & M_i^A(K_i \sqcup K_{-i}) \cdot \mathsf{tv}_i - M_i^{\mathbb{E}P}(K_i \sqcup K_{-i}) \\ & \geq M_i^A(K_i' \sqcup K_{-i}) \cdot \mathsf{tv}_i - M_i^{\mathbb{E}P}(K_i' \sqcup K_{-i}) \ .\end{aligned}$$

- If player i has approximate type K_i' then reporting K_i' very-weakly dominates reporting K_i:

$$\begin{aligned}\forall\, \mathsf{tv}_i' \in K_i' : \quad & M_i^A(K_i' \sqcup K_{-i}) \cdot \mathsf{tv}_i' - M_i^{\mathbb{E}P}(K_i' \sqcup K_{-i}) \\ & \geq M_i^A(K_i \sqcup K_{-i}) \cdot \mathsf{tv}_i' - M_i^{\mathbb{E}P}(K_i \sqcup K_{-i}) \ .\end{aligned}$$

On one hand, we can choose $\mathsf{tv}_i = x$ and $\mathsf{tv}_i' = x+1$, and sum the two corresponding inequalities. The $M^{\mathbb{E}P}$ price terms cancel, and we get:

$$M_i^A(K_i' \sqcup K_{-i}) \geq M_i^A(K_i \sqcup K_{-i}) \ .$$

On the other hand, we can choose $\mathsf{tv}_i = \lfloor x(1+\delta) \rfloor$ and $\mathsf{tv}_i' = \lceil (x+1)(1-\delta) \rceil$,[3] and sum the two corresponding inequalities to get:

$$\big(M_i^A(K_i \sqcup K_{-i}) - M_i^A(K_i' \sqcup K_{-i})\big) \cdot \big(\lfloor x(1+\delta) \rfloor - \lceil (x+1)(1-\delta) \rceil\big) \geq 0 \ .$$

Therefore, whenever $x > \frac{3-\delta}{2\delta}$, we always have $\lfloor x(1+\delta) \rfloor - \lceil (x+1)(1-\delta) \rceil > 0$ and thus $M_i^A(K_i \sqcup K_{-i}) = M_i^A(K_i' \sqcup K_{-i})$. Finally, going back to the two inequalities for very-weak dominance, we also deduce that $M_i^{\mathbb{E}P}(K_i \sqcup K_{-i}) = M_i^{\mathbb{E}P}(K_i' \sqcup K_{-i})$, as desired. □

We can now go back to the proof of the theorem. Define $c \stackrel{\text{def}}{=} \lfloor \frac{3-\delta}{2\delta} \rfloor + 1$ and $K \stackrel{\text{def}}{=} (\mathsf{int}_\delta(c), \mathsf{int}_\delta(c), \dots, \mathsf{int}_\delta(c))$. There must exist some player, say player 1, such that $M_1^A(K) \leq 1/n$, because the summation of $M_i^A(K)$ over $i \in N$ cannot be greater than 1. Invoking Lemma 5.1 multiple times with player 1, K_{-i} of this proof, and x going from c to B, we obtain that

$$M_1^A(\mathsf{int}_\delta(B), \mathsf{int}_\delta(c), \dots, \mathsf{int}_\delta(c)) = M_1^A(K) \leq \frac{1}{n} \ .$$

Now suppose that the true approximate type profile of the players is $K' \stackrel{\text{def}}{=} (\mathsf{int}_\delta(B), \mathsf{int}_\delta(c), \dots, \mathsf{int}_\delta(c))$. Then, for the choice of true type profile $\mathsf{tv} = (B, c, \dots, c) \in K'$, we get the following social welfare:

$$\mathbb{E}\big[\mathrm{SW}\big(\mathsf{tv}, F(K')\big)\big] \leq \frac{1}{n} B + \frac{n-1}{n} c \leq \left(\frac{1}{n} + \frac{c}{B}\right) B = \left(\frac{1}{n} + \frac{c}{B}\right) \cdot \mathrm{MSW}(\mathsf{tv}) \ .$$

□

5.1.2 Multiple-Good Auctions

We present our negative result for partial implementation of ϵ-MSW in very-weakly-dominant strategies within *multiple*-good auctions. Fix $n, m, B \in \mathbb{N}$ and consider any δ-approximate combinatorial auction pre-context $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ for n players, m goods, and maximum valuation B.

[3] As long as $x > \frac{1}{2\delta}$ it is guaranteed that, for these choices, $\mathsf{tv}_i \in K_i$ and $\mathsf{tv}_i' \in K_i'$. But later we will choose $x > \frac{3-\delta}{2\delta}$, so we are safe.

Theorem (Part 2 of Theorem **??**, restated)**.** *Consider any (possibly randomized) mechanism $M = (\Sigma, F)$ such that $\Sigma = Q$ and, for every $K \in Q$, K is a very-weakly-dominant strategy of the game with pre-context* $\mathsf{Pre\,Cnt}$*, players with approximate types K, and mechanism M. If $B > \frac{3-\delta}{2\delta}$, then there exists an approximate type $K \in Q$ and a type profile $\mathsf{tv} \in K$ such that*

$$\mathbb{E}\big[\mathrm{SW}\big(\mathsf{tv}, F(K)\big)\big] \leq \left(\frac{1}{n2^{m-1}} + \frac{\lfloor (3-\delta)/2\delta \rfloor + 1}{B}\right) \cdot \mathrm{MSW}(\mathsf{tv}) \ .$$

The main idea of the proof is similar to the one that we used in the case of single-good auctions in Section 5.1.1. Again, we give a characterization of a player's probability of winning a particular subset of the goods. We then find a player that wins that subset of the goods very rarely, and then use the same "switching worlds" trick to move to a world where that player has a high valuation on that subset, and yet still wins very rarely, thus forcing the mechanism to have a poor social welfare guarantee. What lets us achieve a much stronger negative result in the multiple-good case is the fact that now we have many more valuations (namely, $n(2^m - 1)$ in total) among which the probability mass is spread. The lemma giving the characterization is somewhat more complicated, as the type space is a set of function as opposed to a set of integers.

Proof. For every $K \in Q$ and $\omega \in \Omega$, define $M^A_\omega(K)$ to be the probability that $F(K)$ chooses the outcome ω. For every player $i \in N$, subset $S \subseteq G$, and outcome $\omega \in \Omega$, we say that the pair (i, S) is *consistent* with ω, denoted $(i, S) \sim \omega$, if $\omega = (A, P)$ and $A_i = S$. Next, for every $K \in Q$, player $i \in N$, and subset $S \subset G$, we define $M^A_{i,S}(K)$ to be the probability that player i receives subset S, i.e.,

$$M^A_{i,S}(K) \stackrel{\text{def}}{=} \sum_{\substack{\omega \in \Omega \\ (i,S) \sim \omega}} M^A_\omega(K) \ .$$

This last definition is motivated by the fact that, in a combinatorial auction, player i is only interested in his own allocation A_i, and is indifferent to A_{-i}. We first prove the characterization on a player's winning probability of S and expected prices:

Lemma 5.2. *Fix player $i \in N$, $K \in Q$, $K'_i \in Q_i$, and non-empty $S \subset G$. If $x > \frac{3-\delta}{2\delta}$, $K_i(S) = \mathsf{int}_\delta(x)$, $K'_i(S) = \mathsf{int}_\delta(x+1)$, and $K_i(T) = K'_i(T)$ for all $T \neq S$, then*

$$M^A_{i,S}(K_i \sqcup K_{-i}) = M^A_{i,S}(K'_i \sqcup K_{-i}) \text{ and}$$
$$M^{\mathbb{E}P}_i(K_i \sqcup K_{-i}) = M^{\mathbb{E}P}_i(K'_i \sqcup K_{-i}) \ .$$

Proof. We consider two cases:

- If player i has approximate type K_i then reporting K_i very-weakly dominates reporting K'_i:

$$\begin{aligned}\forall\, \mathsf{tv}_i \in K_i : \quad & \sum_{T \subseteq G} \big(M^A_{i,T}(K_i \sqcup K_{-i}) \cdot \mathsf{tv}_i(T)\big) - M^{\mathbb{E}P}_i(K_i \sqcup K_{-i}) \\ \geq & \sum_{T \subseteq G} \big(M^A_{i,T}(K'_i \sqcup K_{-i}) \cdot \mathsf{tv}_i(T)\big) - M^{\mathbb{E}P}_i(K'_i \sqcup K_{-i}) \ .\end{aligned}$$

- If player i has approximate type K'_i then reporting K'_i very-weakly dominates reporting K_i:

$$\begin{aligned}\forall\, \mathsf{tv}'_i \in K'_i : \quad & \sum_{T \subseteq G} \big(M^A_{i,T}(K'_i \sqcup K_{-i}) \cdot \mathsf{tv}'_i(T)\big) - M^{\mathbb{E}P}_i(K'_i \sqcup K_{-i}) \\ \geq & \sum_{T \subseteq G} \big(M^A_i(K_i \sqcup K_{-i}) \cdot \mathsf{tv}'_i(T)\big) - M^{\mathbb{E}P}_i(K_i \sqcup K_{-i}) \ .\end{aligned}$$

On one hand, we can choose $\mathsf{tv}_i(S) = x$ and $\mathsf{tv}'_i(S) = x + 1$, and for all $T \neq S$ choose $\mathsf{tv}_i(T) = \mathsf{tv}'_i(T)$ to be some arbitrary point in $K_i(T) = K'_i(T)$. Summing up the two inequalities, we get:

$$M^A_{i,S}(K'_i \sqcup K_{-i}) \geq M^A_{i,S}(K_i \sqcup K_{-i}) \ .$$

On the other hand, we can choose $\mathsf{tv}_i(S) = \lfloor x(1+\delta) \rfloor \in K_i(S)$ and $\mathsf{tv}'_i(S) = \lceil (x+1)(1-\delta) \rceil \in K'_i(S)$, and for all $T \neq S$ choose $\mathsf{tv}_i(T) = \mathsf{tv}'_i(T)$ to be some arbitrary point in $K_i(T) = K'_i(T)$.[4] Again summing the two inequalities, most of the terms cancel, and we are left with the following:

$$\big(M^A_{i,S}(K_i \sqcup K_{-i}) - M^A_{i,S}(K'_i \sqcup K_{-i})\big) \cdot \big(\lfloor x(1+\delta) \rfloor - \lceil (x+1)(1-\delta) \rceil\big) \geq 0 \ .$$

Therefore, whenever $x > \frac{3-\delta}{2\delta}$, we always have $\lfloor x(1+\delta) \rfloor - \lceil (x+1)(1-\delta) \rceil > 0$ and thus $M^A_i(K_{i,S} \sqcup K_{-i}) = M^A_{i,S}(K'_i \sqcup K_{-i})$. Finally, going back to the two inequalities for very-weak dominance, we can also deduce that $M^{\mathbb{E}P}_i(K_i \sqcup K_{-i}) = M^{\mathbb{E}P}_i(K'_i \sqcup K_{-i})$, as desired. □

We can now go back to the proof of the theorem. Define $c \stackrel{\text{def}}{=} \lfloor \frac{3-\delta}{2\delta} \rfloor + 1$ and $K_i(T) \stackrel{\text{def}}{=} \mathsf{int}_\delta(c)$ for all nonempty $T \subseteq G$ and players $i \in N$. Because the mechanism assigns disjoint outcomes with a maximum total probability of 1, we have:

$$\sum_{i \in N} \sum_{\substack{T \subset G \\ 1 \in T}} M^A_{i,T}(K) \leq 1 \ .$$

Again, the events in the summation are disjoint, because there is only one good $1 \in G$ and it can be assigned to only one of the n players. Also because $|\{T \subset G : 1 \in T\}| = 2^{m-1}$, at least one of the probabilities, say $M^A_{1,S}(K)$, is at most $\frac{1}{n2^{m-1}}$.

Now define $K'_1(S) \stackrel{\text{def}}{=} \mathsf{int}_\delta(B)$ and $K'_1(T) \stackrel{\text{def}}{=} \mathsf{int}_\delta(c)$ for all nonempty $T \neq S$, and $K'_i(T) = \mathsf{int}_\delta(c)$ for all players $i \neq 1$ and nonempty $T \subseteq G$. Invoking Lemma 5.2 multiple times with player 1 and the subset S, K'_{-i} of this proof, and x going from c to B, we obtain that

$$M^A_{1,S}(K') = M^A_{1,S}(K) \leq \frac{1}{m2^{m-1}} \ .$$

Now suppose that the true approximate type profile of the players is K'. Then, for the choice of the true type profile $\mathsf{tv} = (\mathsf{tv}_1, \ldots, \mathsf{tv}_n)$ with $\mathsf{tv}_1(S) = B$ and $\mathsf{tv}_1(T) = c$ for all nonempty $T \neq S$, and $\mathsf{tv}_i(T) = c$ for all players $i \neq 1$ and nonempty $T \subseteq G$, we get the following social welfare:

$$\mathbb{E}\big[\mathrm{SW}(\mathsf{tv}, F(K'))\big] \leq \frac{B}{n2^{m-1}} + \left(1 - \frac{1}{n2^{m-1}}\right) \cdot c < \left(\frac{1}{n2^{m-1}} + \frac{c}{B}\right) \cdot \mathrm{MSW}(\mathsf{tv}) \ .$$

□

5.2 Possibility Result

We now provide a mechanism that asymptotically matches the impossibility theorems provided in the previous section. Fix $n, m, B \in \mathbb{N}$ and consider any combinatorial auction pre-context $\mathsf{Pre\,Cnt} = (N, \Theta, \Omega, U, Q)$ for n players, m goods, and maximum valuation B. Consider the mechanism M for $\mathsf{Pre\,Cnt}$ such that each player's strategy set simply consists of the singleton $\{\perp\}$ and works as follows:

$M(\perp^n) \stackrel{\text{def}}{=}$

1. Choose a player $i \in N$ uniformly at random.
2. Choose a non-empty subset of goods $A \subseteq G$ uniformly at random.
3. Assign A to player i with price set to zero.

Note that M trivially enjoys a strictly-dominant strategy, namely the only strategy $\perp$ for each player. We now state and prove the social welfare guarantees of M:

Theorem (Theorem 4.5, restated). *For any approximate type profile $K = (K_1, \ldots, K_n)$ and type profile $\mathsf{tv} \in K$,*

$$\mathbb{E}\big[\mathrm{SW}(\mathsf{tv}, M(\perp^n))\big] \geq \frac{\mathrm{MSW}(\mathsf{tv})}{n(2^m - 1)} \ .$$

[4] Again, as long as $x > \frac{1}{2\delta}$, it is guaranteed that, for these choices, $\mathsf{tv}_i \in K_i$ and $\mathsf{tv}'_i \in K'_i$. But later we will choose $x > \frac{3-\delta}{2\delta}$, so we are safe.

Proof. There exists an optimal allocation profile A^*, i.e. , for which $\mathrm{MSW}(\mathsf{tv}) = \sum_{i=1}^n \mathsf{tv}(A_i^*)$. The mechanism M will assign to each player i the subset A_i^* with probability $\frac{1}{n(2^m-1)}$. Hence the social welfare is at least $\sum_{i=1}^n \frac{\mathsf{tv}(A_i^*)}{n(2^m-1)}$ in expectation, as desired. □

Note that the theorem above holds regardless of the approximation quality Q in the pre-context. (Thus it holds, in particular, for any $\delta \in (0,1)$ and any δ-approximate combinatorial auction pre-context.)

6 Proof Details: Implementation in Undominated Strategies

We study the relation between the approximation factor δ and implementations of ϵ-fraction of maximum social welfare in *not-weakly-dominated strategies*. In this section, we present the following results:

> Section 6.1: We focus on auctions of a single good. First, we show that insisting VCG one can still guarantee a reasonable $\frac{(1-\delta)^2}{(1+\delta)^2}$-fraction of the maximum social welfare. Second, we provide a negative result stating that VCG is optimal as a deterministic mechanism, and asymptotically optimal even if randomized mechanism is allowed.
>
> Section 6.2: We provide a negative result that the VCG mechanism with a random tie-breaking rule fails exponentially with respect to the number of goods m. We also put some remarks in this sub-section about some other results not listed and the ongoing work.

We remark that the result from Section **??** ensures that our negative results for implementation in not-weakly-dominated strategies in the case of approximate player types are indeed due to the approximate nature of the types, and not because, say, not much social welfare can be guaranteed in not-weakly-dominated strategies even in the case of exact player types.

6.1 Single-Good Auctions

6.1.1 Possibility Result

Recall that the VCG mechanism M, whenever there is only one good on sale, is defined as follows: on input a bid profile $v \in \Theta = \{0,1,\dots,B\}^n$,

$M(v) \stackrel{\text{def}}{=}$

1. Assign the good to the highest bidder: $i^* \stackrel{\text{def}}{=} \arg\max_i v_i$.
2. Charge the highest bidder the second price: $P_{i^*} \stackrel{\text{def}}{=} \max_{i \neq i^*} v_i$.
   (And everyone else pays nothing.)

We show that the VCG mechanism, at least in the case of auctions of a single good, is still able to guaranteed a modest fraction of the maximum social welfare:

Theorem (Theorem 4.6, restated)**.** *Fix any $\delta \in (0,1)$. No matter how M chooses the tie breaking rule for its first step, for every δ-approximate knowledge profile $K = (K_1,\dots,K_n)$, every $\mathsf{tv} \in K$, and every $\sigma \in \mathsf{NWDed}(\mathsf{tv})$,*

$$\mathrm{SW}\big(\mathsf{tv}, M(\sigma)\big) \geq \frac{(1-\delta)^2}{(1+\delta)^2}\mathrm{MSW}(\mathsf{tv}) - \frac{2}{(1+\delta)^2} \quad . \tag{1}$$

Proof. It suffices to prove the theorem assuming $K_i = \mathsf{int}_\delta(\mathsf{mid}_i)$ for every player i because, by Lemma 2.9, whenever $K_i \subsetneq \mathsf{int}_\delta(\mathsf{mid}_i)$, the set of not-weakly-dominated strategies $\mathsf{NWDed}_i(K_i)$ only shrinks in size thus leaving fewer possible choices for $\mathsf{tv} \in K$; therefore, $K_i = \mathsf{int}_\delta(\mathsf{mid}_i)$ for all $i \in N$ is the "worst case" to analyze.

So consider any player with approximate type $K_i = \mathsf{int}_\delta(x)$. We claim that strategy $v_i = \lfloor x(1-\delta) \rfloor$ weakly dominates every strategy $s_i' < s_i$. We first prove that the very weakly dominance holds, and then find a witness such that playing s_i is strictly better than s_i'.

For any other v_{-i}, if in both $M(v_i \sqcup v_{-i})$ and $M(s'_i \sqcup s_{-i})$ player i gets the good, then the utilities are the same because they have the same $\mathsf{tv}_i \in \mathsf{int}_\delta(x)$ and the price depends on the second highest bid. Similarly, if in both cases player i loses, the utilities are zero. The only interesting case is if s_i wins but s'_i not. For such v_{-i}:

$$\forall \mathsf{tv}_i \in \mathsf{int}_\delta(x), \quad M_i^A(v_i \sqcup s_{-i})\mathsf{tv}_i - M_i^{\mathbb{E}P}(s_i \sqcup s_{-i}) \geq x(1-\delta) - \max_{j\neq i} s_j \geq s_i - \max_{j \neq i} s_j \geq 0$$
$$= M_i^A(v'_i \sqcup v_{-i})\mathsf{tv}_i - M_i^{\mathbb{E}P}(v'_i \sqcup v_{-i}) \ . \quad (2)$$

Therefore, we conclude that for all v_{-i} and $\mathsf{tv}_i \in \mathsf{int}_\delta(x)$, $u(\mathsf{tv}_i, M(v_i \sqcup v_{-i})) \geq u(\mathsf{tv}_i, M(v'_i \sqcup v_{-i}))$; in other words v_i very weakly dominates v'_i.

On the other hand, no matter which tie breaking rule M chooses, there must exist some v_{-i} such that playing v_i wins but v'_i not because they differ by at least one. In that specific case, let $\mathsf{tv}_i = \lfloor x(1+\delta) \rfloor$[5] and Eq. 2 will become strict. For similar reason, one can prove that strategy $v_i = \lceil x(1+\delta) \rceil$, it will (weakly) dominate every strategy $v'_i > v_i$. This proves that for any $x \in [0, B]$, the set of undominated strategies $\mathsf{NWDed}_i(x) \subset \Delta([\lfloor x(1-\delta) \rfloor, \lceil x(1+\delta) \rceil] \cap \mathbb{Z})$.

Now we prove the lower bound of the social welfare. For any knowledge profile $K_i = \mathsf{int}_\delta(\mathsf{mid}_i)$, let x be the largest of mid_i's and $y \leq x$ be the second largest. If $\lfloor x(1-\delta) \rfloor > \lceil y(1+\delta) \rceil$, then the two sets of undominated strategies are not intersecting, so x will win the auction and we have perfect social welfare. The second highest player may win the auction only if $\lfloor x(1-\delta) \rfloor \leq \lceil y(1+\delta) \rceil$, which implies that $x(1-\delta) \leq y(1+\delta) + 2$ and in this case $\frac{\mathrm{SW}}{\mathrm{MSW}} = \frac{y}{x}$. Therefore, the worst possible social efficiency is lower bounded bounded by:

$$\forall \sigma \in \mathsf{NWDed}(\mathsf{mid}), \quad \mathrm{SW}(\mathsf{mid}, M(\sigma) \geq \frac{x(1-\delta)}{1+\delta} - \frac{2}{1+\delta} \geq \frac{1-\delta}{1+\delta}\mathrm{MSW}(\mathsf{mid}) - \frac{2}{1+\delta} \ .$$

We now go to the last step of the proof:

$$\forall \mathsf{tv} \in K, \forall \sigma \in \mathsf{NWDed}(K), \quad \mathrm{SW}(\mathsf{tv}, M(\sigma)) \geq \frac{1}{1+\delta}\mathrm{SW}(\mathsf{mid}, M(\sigma) \geq \frac{1}{1+\delta}\frac{1-\delta}{1+\delta}\mathrm{MSW}(\mathsf{mid}) - \frac{2}{(1+\delta)^2}$$
$$\geq \frac{1}{1+\delta}\frac{1-\delta}{1+\delta}\frac{1-\delta}{\mathrm{MSW}}(\mathsf{tv}) - \frac{2}{(1+\delta)^2} = \frac{(1-\delta)^2}{(1+\delta)^2}\mathrm{MSW}(\mathsf{tv}) - \frac{2}{(1+\delta)^2} \ .$$

□

Remark 6.1. We note that:

- In terms of the positive result of deterministic mechanisms, one can choose any deterministic tie breaking rule for VCG, and the above theorem guarantees a $\frac{(1-\delta)^2}{(1+\delta)^2}$ fraction of the social welfare;
- One can check that the above theorem is also true for arbitrarily randomized tie breaking rules;
- One can check that all integer points in the interval $[\lfloor \mathsf{mid}_i(1-\delta) \rfloor + 1, \lceil \mathsf{mid}_i(1+\delta) \rceil - 1]$ are undominated for arbitrary (randomized or deterministic) tie breaking rules, so this bound cannot be improved.

6.1.2 Impossibility Result

Next, we prove another intuitive, this time more technical, lemma that we will also use several times in our proofs: whenever two approximate types K_i and K'_i intersect in at least two points, their corresponding not-weakly-dominated strategy sets also intersect. Note that the strategy space Σ_i may be arbitrary (and, in particular, not necessarily equal to 2^{Θ_i}).

Lemma 6.2. *For every player $i \in N$, if K_i and K'_i intersect in at least two points, then $\mathsf{NWDed}_i(K_i) \cap \mathsf{NWDed}_i(K'_i) \neq \emptyset$.*

Proof. Suppose by way of contradiction that $\mathsf{NWDed}_i(K_i)$ and $\mathsf{NWDed}_i(K'_i)$ are disjoint. W.L.O.G., let us assume that $\{x, y\} \subset K_i \cap K'_i$ and $x > y$. Note that the set of not-weakly-dominated strategies in a game is always non-empty and thus so are both $\mathsf{NWDed}_i(K_i)$ and $\mathsf{NWDed}_i(K'_i)$. So let us choose an arbitrary strategy $\sigma \in \mathsf{NWDed}_i(K_i)$, which is not in $\mathsf{NWDed}_i(K'_i)$ as $\mathsf{NWDed}_i(K_i)$ and $\mathsf{NWDed}_i(K'_i)$ are disjoint by assumption.

[5] We must have $\mathsf{tv}_i \in K_i = \mathsf{int}_\delta(x)$ because $\lfloor x(1+\delta) \rfloor \geq x > v_i = \lfloor x(1-\delta) \rfloor$.

In the game $\mathcal{G}'$, invoking the boundedness of $\mathcal{G}'$, σ must be weakly dominated by some $\sigma' \in \mathsf{NWDed}_i(K_i')$. Similarly, in the game $\mathcal{G}$, invoking the boundedness of $\mathcal{G}$, σ must be weakly dominated by some $\tau \in \mathsf{NWDed}_i(K_i)$. Continuing in this fashion, we can "jump" from game $\mathcal{G}$ to $\mathcal{G}'$ and back, each time invoking the boundedness condition and the disjointness of the two sets, creating a "chain" of strategies. However, by the finiteness of the NWDed_i, this chain will end up in a cycle and W.L.O.G. let us assume

$$\sigma \overset{\mathrm{w}}{\underset{i,K_i'}{\succ}} \sigma' \overset{\mathrm{w}}{\underset{i,K_i}{\succ}} \tau \overset{\mathrm{w}}{\underset{i,K_i'}{\succ}} \tau' \overset{\mathrm{w}}{\underset{i,K_i}{\succ}} \sigma \ .$$

We are going to deduce a contradiction from that. Fix an arbitrary σ_{-i}. Let us for the notation simplicity define $M(\sigma) = M_i(\sigma \sqcup \sigma_{-i})$, and the latter is the probability that player i is to be assigned the good. The four dominance statements tell us the follows:

$$\begin{array}{lrcl}
\forall \mathsf{tv}' \in K_i', & M^A(\sigma)\mathsf{tv}' - M^{\mathbb{E}P}(\sigma) & \leq & M^A(\sigma')\mathsf{tv}' - M^{\mathbb{E}P}(\sigma') \\
\forall \mathsf{tv} \in K_i, & M^A(\sigma')\mathsf{tv} - M^{\mathbb{E}P}(\sigma') & \leq & M^A(\tau)\mathsf{tv} - M^{\mathbb{E}P}(\tau) \\
\forall \mathsf{tv}' \in K_i', & M^A(\tau)\mathsf{tv}' - M^{\mathbb{E}P}(\tau) & \leq & M^A(\tau')\mathsf{tv}' - M^{\mathbb{E}P}(\tau') \\
\forall \mathsf{tv} \in K_i, & M^A(\tau')\mathsf{tv} - M^{\mathbb{E}P}(\tau') & \leq & M^A(\sigma)\mathsf{tv} - M^{\mathbb{E}P}(\sigma)
\end{array} \tag{3}$$

The first step is to choose $\mathsf{tv} = x$ and $\mathsf{tv}' = y$ and sum the four inequalities in Eq. 3 up. The prices will cancel and we can deduce that

$$(M^A(\sigma') + M^A(\tau'))(x - y) \leq (M^A(\sigma) + M^A(\tau))(x - y) \ .$$

On the other hand, if we choose $\mathsf{tv} = y$ and $\mathsf{tv}' = x$ and sum them up we get:

$$(M^A(\sigma) + M^A(\tau))(x - y) \leq (M^A(\sigma') + M^A(\tau'))(x - y) \ .$$

Combining these two we get $M^A(\sigma')+M^A(\tau') = M^A(\sigma)+M^A(\tau)$. Therefore, for both sets of values of $(\mathsf{tv}, \mathsf{tv}')$, all of the four inequalities in Eq. 3 are tight. This suggests that each inequality in Eq. 3 is tight regardless of the choice of tv or tv'. Therefore, we must have $M^A(\sigma) = M^A(\sigma')$ and $M^{\mathbb{E}P}(\sigma) = M^{\mathbb{E}P}(\sigma')$.

At last, the above claim holds for all σ_{-i} and therefore σ and σ' are essentially two equivalent strategies no matter how other players behave. However, $\sigma \in \mathsf{NWDed}_i(K_i)$ but $\sigma' \in \mathsf{NWDed}_i(K_i')$ so results in a contradiction. This ends the proof. □

Now we begin to prove the impossibility result for single good auction.

Theorem (Theorem 4.7, restated)**.** *When $B \geq \frac{1}{\delta}$:*

- *for any deterministic mechanism M, there exists some knowledge profile K and $\sigma \in \overline{\mathsf{NWDed}}(K)$, such that*

$$\forall \mathsf{tv} \in K, \quad \mathrm{SW}(\mathsf{tv}, M(\sigma)) \leq \left(\frac{(1-\delta)^2}{(1+\delta)^2} + \frac{4}{B}\right) \mathrm{MSW}(\mathsf{tv}) \ ;$$

- *for any randomized mechanism M, there exists some knowledge profile K and $\sigma \in \overline{\mathsf{NWDed}}(K)$, such that*

$$\forall \mathsf{tv} \in K, \quad \mathrm{SW}(\mathsf{tv}, M(\sigma)) \leq \left(\frac{(1-\delta)^2 + \frac{4\delta}{n}}{(1+\delta)^2} + \frac{4}{B}\right) \mathrm{MSW}(\mathsf{tv}) \ .$$

Proof. Let $x = B$ and $y = \lfloor \frac{x(1-\delta)+2}{1+\delta} \rfloor$ be two integers in the type space. One can verify that $x \geq y$ by our choice of $B \geq \frac{1}{\delta}$, and $\mathsf{int}_\delta(x) \cap \mathsf{int}_\delta(y)$ contains at least two different integer points. Using Lemma 6.2, we can always to pick $\sigma_i \in \mathsf{NWDed}_i(x) \cap \mathsf{NWDed}_i(y)$ (for every player-i) because they intersect. Now consider the outcome for $M(\sigma_1, \ldots, \sigma_n)$.

If M is deterministic, there must exist some player, say player 1, such that $M_1^A(\sigma) = 0$, i.e., he does not receive the good. Consider the following $K = (\mathsf{int}_\delta(x), \mathsf{int}_\delta(y), \ldots, \mathsf{int}_\delta(y))$ and $\mathsf{tv} = (\lfloor(1+\delta)x\rfloor, \lceil(1-\delta)y\rceil, \ldots, \lceil(1-\delta)y\rceil)$:

$$\begin{aligned}
\mathrm{SW}(\mathsf{tv}, M(\sigma)) = \lceil(1-\delta)y\rceil) \leq (1-\delta)y + 1 \leq \frac{x(1-\delta)^2}{1+\delta} + 3 &\leq \frac{(1-\delta)^2}{(1+\delta)^2}\lfloor(1+\delta)x\rfloor + 4 \\
&\leq (\frac{(1-\delta)^2}{(1+\delta)^2} + \frac{4}{B})\lfloor(1+\delta)x\rfloor = \left(\frac{(1-\delta)^2}{(1+\delta)^2} + \frac{4}{B}\right) \mathrm{MSW}(\mathsf{tv}) \ .
\end{aligned}$$

If M is randomized, there must exist some player, say player 1, such that $M_1^A(\sigma) \leq \frac{1}{n}$, i.e., he receives the good with no more than $\frac{1}{n}$ probability. We define the same K and tv as above:

$$\begin{aligned}
\mathrm{SW}(\mathsf{tv}, M(\sigma)) &\leq \frac{n-1}{n}\lceil(1-\delta)y\rceil) + \frac{1}{n}\lfloor(1+\delta)x\rfloor \\
&< \frac{n-1}{n}(1-\delta)y + \frac{1}{n}\lfloor(1+\delta)x\rfloor + 1 \leq \frac{n-1}{n}\frac{(1-\delta)^2 x}{1+\delta} + \frac{1}{n}\lfloor(1+\delta)x\rfloor + 3 \\
&< \frac{n-1}{n}\frac{(1-\delta)^2}{(1+\delta)^2}\lfloor(1+\delta)x\rfloor + \frac{1}{n}\lfloor(1+\delta)x\rfloor + 4 < \left(\frac{n-1}{n}\frac{(1-\delta)^2}{(1+\delta)^2} + \frac{1}{n} + \frac{4}{B}\right)\lfloor(1+\delta)x\rfloor + 4 \\
&= \left(\frac{n-1}{n}\frac{(1-\delta)^2}{(1+\delta)^2} + \frac{1}{n} + \frac{4}{B}\right)\mathrm{MSW}(\mathsf{tv}) = \left(\frac{(1-\delta)^2 + \frac{4\delta}{n}}{(1+\delta)^2} + \frac{4}{B}\right)\mathrm{MSW}(\mathsf{tv}) \ .
\end{aligned}$$

□

6.2 Multiple-Good Auctions

Since VCG does relatively well in single good auction, a direct question is how well it generalizes to multiple good. We start with a negative answer. Let us consider the following VCG mechanism M for combinatorial auctions, using randomized tie breaking rule: on input a bid profile $v \in \Theta = \{0, 1, \ldots, B\}^{n\times(2^m-1)}$,

$M(v) \stackrel{\text{def}}{=}$

1. Randomly pick an allocation that maximizes the social welfare: $A \stackrel{\text{def}}{=} \arg\max_A \sum_{i=1}^n v_i(A_i)$. We do not require the uniform randomness. Instead, we only require that each maximizer has a positive probability to be chosen.
2. Charge the i-th bidder the following price:

$$P_{i^*} \stackrel{\text{def}}{=} \max_{A'} \sum_{j\neq i} v_j(A'_j) - \sum_{j\neq i} v_j(A_j) \ .$$

We have the following negative result which states that the fraction of the maximum social welfare we can guarantee is exponentially small in terms of the number of goods.

Theorem 6.3. *Fix the M defined above and any $\delta \in (0,1)$. There exists some δ-approximate knowledge profile $K = (K_1, \ldots, K_n)$, some $\mathsf{tv} \in K$, and some $\sigma \in \mathsf{NWDed}(\mathsf{tv})$,*

$$\mathrm{SW}\big(\mathsf{tv}, M(\sigma)\big) \leq O\left(\left(\frac{1-\delta}{1+\delta}\right)^{2^m-2}\right)\mathrm{MSW}(\mathsf{tv}) \ . \tag{4}$$

Remark 6.4. We remark as follows.

- We delay our proof of Theorem 6.3 to an extended version of this report.
- One can show that the above theorem is asymptotically tight: VCG will always guarantee (up to a constant) this fraction of the maximum social welfare.
- The result also holds for an arbitrary (including deterministic) tie breaking rule, but for more complicated reason.

We are currently working on the generalizations of the results in Section 6.1 to multiple goods.