

Lecture 8

Lecturer: Constantinos Daskalakis

Scribe: Govind Ramnarayan

1 Introduction to Mechanism Design

This lecture is an introduction to *Mechanism Design*, a scientific discipline that is also often called “*Reverse Game Theory*.” While game-theoretic analysis is commonly devoted to the study of a strategic environment that is already in place and explaining the outcomes resulting from strategic behavior within this environment, in mechanism design we start from the desired outcome, and ask if it is possible to design a strategic environment so that the desired outcome arises from strategic behavior within the designed environment. The strategic environment designed to bring about the desired outcome is called a “mechanism” and the study of how to design such an environment is called “mechanism design.”

From elections to kidney exchange platforms, online dating, sharing economy applications, spectrum auctions and online advertising, mechanisms can be recognized as drivers of political, social and economic activity. Our main focus within mechanism design will be on *auctions*. Perhaps the most familiar instances of auctions can be found in online marketplaces, such as eBay, where you bid to win an item. Online auctions are also determining what advertisements you see in banners or in sponsored search results. They are also commonly employed by governments to sell spectrum or drilling rights.

2 Single-Item Auctions

Suppose we have a single item to sell in an auction, and n bidders interested in buying it. Each bidder has a private *value* v_i for the item, and we will assume that her utility for the auction’s outcome takes the following, linear form:

$$u_i = \begin{cases} v_i - p_i & \text{if she wins the item and pays price } p_i; \\ -p_i & \text{if she does not win the item and pays price } p_i. \end{cases}$$

We will consider *sealed-bid auctions*, where each bidder privately sends a bid to the auctioneer. Based on these bids, the auctioneer decides who gets the item and decides on the prices charged to each bidder.

Definition 1. A sealed-bid auction is defined by a pair of functions $x : \mathbb{R}^n \rightarrow \Delta^{n+1} \equiv \{(\pi_0, \dots, \pi_n) \mid \pi_i \geq 0, \sum_i \pi_i = 1\}$, called the “allocation rule,” and $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$, called the “price rule,” and consists of the following steps:

1. Each bidder i privately communicates a bid b_i to the auctioneer (e.g., in a sealed envelope).
2. The auctioneer applies the allocation rule x to the bid vector \mathbf{b} to determine the probability $x_i(\mathbf{b})$ that the item is allocated to each bidder i ; $x_0(\mathbf{b})$ is the probability that the item is un-allocated.
3. The auctioneer applies the price rule p to the bid vector to determine the price $p_i(\mathbf{b})$ charged to each bidder i .

Recall that in mechanism design, we start with an objective, and then try to design a mechanism to achieve it. At this point we have defined what a sealed-bid auction is, but we have not yet said what our *objective* in designing one is. Naturally, there are many such objectives we could consider, but today our focus will be on *maximizing welfare*.

Definition 2. A sealed-bid auction with allocation rule x and price rule p derives welfare

$$\mathbb{E}_{\mathbf{b}} \left[\sum_i x_i(\mathbf{b}) \cdot v_i \right]$$

where the expectation is computed with respect to the (potentially randomized) bids that the bidders will submit to the auction, which are, in turn, chosen by the bidders based on their own value, the information they have about the others' values (if any), as well as based on what x and p are.

Remark 1. We were pretty casual in the above definition, taking expectation with respect to the strategically-chosen bids of the bidders. As it turns out, predicting bidders' bids in an auction is typically highly non-trivial. As such, it is extremely important to design the allocation and price rule of the auction in a manner that both makes it easy for the bidders' to reason about the situation and choose their bids, and makes it easy for us to predict their bidding strategies.

How should we choose the allocation and price rules of a sealed-bid auction to maximize welfare? A natural choice for the allocation rule is to give the good to the highest bidder: that is, $x_i(b_i) = 1$ if $i = \arg \max_j b_j$ (use some tie-breaking rule if there are multiple maximum bids), and $x_i(b_i) = 0$ otherwise. The choice of a price rule is not quite as clear. For example, we may choose to be generous and charge no one (in other words, set all prices to 0). Unfortunately, this is not a good idea, as then the bidders, thinking strategically about the situation, will determine that an optimal bid is to report $+\infty$. In this case, the allocation rule will choose the winner of the item according to its tie-breaking rule, but without any regard to which bidder really values the item the most. As such, the auction may miserably fail in maximizing welfare.

A natural price rule to use is to have the winner of the auction pay their bid, and have the losers pay nothing. The resulting auction is called a *first-price auction*. As we will see, this type of auction is difficult to analyze. We will do this shortly, under certain assumptions, but before doing so let us take the opportunity to present a broader mathematical framework that will allow us to analyze auctions formally.

2.1 Games of Incomplete Information

Generally speaking, sealed-bid auctions are *games of incomplete information*; each party has a private value and needs to make a bid without knowledge of the other parties' values. Let us take a detour to define such strategic environments in more generality.

Definition 3. A game of strict incomplete information is specified by the following ingredients:

1. A set of players $N = \{1, \dots, n\}$.
2. A set of actions X_i for each player i . Set $X = X_1 \times \dots \times X_n$ is the set of action profiles.
3. A set of types T_i for each player i . An element $t_i \in T_i$ is the private information held by player i . In particular, the realized type $t_i \in T_i$ is known to player i , but the other players only know that it is some element of T_i . Set $T = T_1 \times \dots \times T_n$ is the set of type profiles.
4. For each player i , a utility, or payoff, function $u_i : T \times X \rightarrow \mathbb{R}$, where $u_i(t, x)$ is the utility derived by player i , if players' types are t and players' actions are x .
 - We focus on the case of independent private values where each player i 's utility is a function $u_i : T_i \times X \rightarrow \mathbb{R}$, in particular it depends on the player's own type and the actions chosen by all players, but not on the other players' types.

Definition 4 (Bayesian Setting). A Bayesian game of incomplete information has the same ingredients as a game of strict incomplete information of Definition 3. Additionally a distribution F supported on T is common knowledge, such that the realized type profile $t \sim F$. A Bayesian game of incomplete information and independent private values is defined similarly except that F is a product measure $F = F_1 \times \dots \times F_n$.

Now, we define notions of strategy and equilibrium in games of incomplete information. Informally, a strategy maps types to actions.

Definition 5. A (pure) strategy for player i in a game of incomplete information is a function $s_i : T_i \rightarrow X_i$.

Next, we define a few notions of equilibrium. We will focus on games of incomplete information and independent private values. Informally, an ex-post Nash equilibrium is a collection of strategies for the players of a game such that no player ever has incentive not to follow the recommendation of her strategy to map her type to an action, assuming the other players use the recommendations of their strategies. Formally,

Definition 6. A profile of strategies s_1, \dots, s_n is an ex-post Nash equilibrium of a game of incomplete information with independent private values if, for all i , all t_1, \dots, t_n , and all x'_i we have that

$$u_i(t_i, s_i(t_i), s_{-i}(t_{-i})) \geq u_i(t_i, x'_i, s_{-i}(t_{-i})).$$

In turn, a dominant strategy equilibrium is a collection of strategies such that no player ever has incentive not to follow the recommendation of her strategy regardless of what actions the other players use. Formally,

Definition 7. A profile of strategies s_1, \dots, s_n is a dominant strategy equilibrium of a game of incomplete information with independent private values if, for all i , t_i , x_{-i} , and x'_i we have that

$$u_i(t_i, s_i(t_i), x_{-i}) \geq u_i(t_i, x'_i, x_{-i}).$$

Finally, a Bayesian Nash equilibrium is a notion of equilibrium applicable in Bayesian games of incomplete information. It is similar to an ex-post Nash equilibrium, except that players take expectations with respect to the types of the other players. In particular, it is a collection of strategies such that no player ever has incentive not to follow the recommendation of her strategy to map her type to an action, in expectation with respect to the types of the other players and assuming they follow their strategies to map their types to actions. Formally,

Definition 8. A profile of strategies s_1, \dots, s_n is a Bayesian Nash equilibrium of a Bayesian game of incomplete information with independent private values if for all i , all t_i and all x'_i we have that

$$\mathbb{E}_{t_{-i}} u_i(t_i, s_i(t_i), s_{-i}(t_{-i})) \geq \mathbb{E}_{t_{-i}} u_i(t_i, x'_i, s_{-i}(t_{-i})).$$

2.2 The First-Price Auction

With our formal definitions in place, let us now analyze the first-price auction in a simple setting.

Theorem 1. Suppose n bidders participate in a first-price auction, and their values are sampled i.i.d. from $U[0, 1]$. Then the collection of strategies

$$s_i(v_i) = \left(1 - \frac{1}{n}\right) \cdot v_i, \quad \forall i,$$

is a Bayesian Nash equilibrium.

Proof. Let us consider the perspective of an arbitrary bidder i whose value has been realized to some arbitrary v_i . It suffices to verify that bidding $(1 - \frac{1}{n}) \cdot v_i$ maximizes the expected utility for bidder i if all the other bidders use strategy $s_j(v_j) = (1 - \frac{1}{n}) \cdot v_j$ to map their values to bids, where the expectation is with respect to their values drawn uniformly from $[0, 1]$.

To argue this, let us first notice that, if all the other bidders use strategy $s_j(v_j) = (1 - \frac{1}{n}) \cdot v_j$, then it is certainly sub-optimal for bidder i to bid higher than $1 - \frac{1}{n}$. Hence, w.l.o.g. we can restrict our attention to the interval $[0, 1 - \frac{1}{n}]$ for the purposes of identifying bidder i 's optimal bid.

Next, suppose that bidder i bids some $b_i \in [0, 1 - \frac{1}{n}]$. The expected utility of bidder i then is:

$$\begin{aligned}\mathbb{E}u_i &= (v_i - b_i) \cdot \Pr[\forall j (b_i \geq b_j)] \\ &= (v_i - b_i) \cdot \Pr\left[\forall j \left(b_i \geq \left(1 - \frac{1}{n}\right) v_j\right)\right] \\ &= (v_i - b_i) \cdot \left(\frac{n}{n-1} b_i\right)^{n-1},\end{aligned}$$

where in the transition to the last line we used that values v_j are uniformly distributed in $[0, 1]$, and that we are optimizing with respect to $b_i \in [0, 1 - \frac{1}{n}]$. Selecting b_i to maximize the above quantity gives us that $b_i = (1 - \frac{1}{n})v_i$. \square

The setting of Theorem 1 was fairly simple. Specifically, each bidder had their value drawn from the same distribution. What if different bidders have their values drawn from different distributions? In this case, the Bayesian Nash equilibrium strategies can get quite complicated quite fast. For example, in [1] Kaplan and Zamir analyze the setting where two bidder's values are drawn uniformly from the intervals $[0, 5]$ and $[6, 7]$ respectively. Kaplan and Zamir show that the following strategies form a Bayesian Nash equilibrium: bidder 1, whose value is uniform in $[0, 5]$, bids her value if it falls in $[0, 3]$, otherwise for all $b \in (3, 13/3]$, bidder 1 bids b if her value is:

$$v_1(b) = \frac{36}{(2b-6)\left(\frac{1}{5}\right)e^{(9/4)+(6/(6-2b))} + 24 - 4b}.$$

For all $b \in (3, 13/3]$, bidder 2 bids b when her value is

$$v_2(b) = 6 + \frac{36}{(2b-6)(20)e^{-(9/4)-(6/(6-2b))} - 4b}.$$

See the figure below for a visual diagram of the bidders' Bayesian Nash equilibrium strategies. Interest-

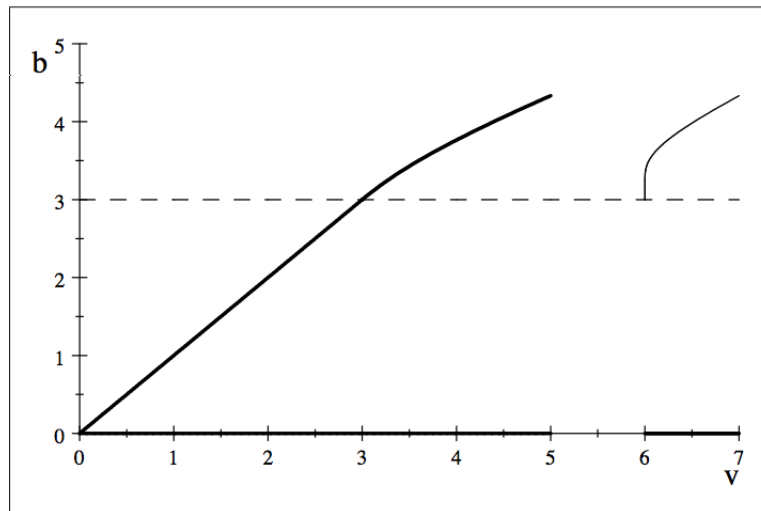


Figure 1: Equilibrium 1. The thicker line is buyer 1's bid function.

ingly, under this Bayesian Nash equilibrium it is not always the case that the item is won by the bidder who values it the most. Indeed, while the value of bidder 2 is always higher than that of bidder 1, bidder 1 wins the item with constant probability.

2.3 The Second-Price, a.k.a. Vickrey, Auction

Recall that, in a first-price auction, the equilibrium strategies may depend on the number of bidders, as well as the bidders' information about each other. Furthermore, these strategies can easily get quite complex and analytically challenging to compute, and the winner may not even be the one who values the item the most. This motivates us to consider a different type of auction.

The idea in second-price auction, which is also called Vickrey auction, is to charge the winner the second highest bid. This idea may seem a bit strange at first (why charge the highest bidder the second highest bid when they are willing to pay the highest bid?) but this auction, at least in disguise, is quite prevalent in practice. Indeed, the outcome of the second-price auction matches that of the commonly used ascending price auction. Furthermore, as we will see next, this auction enjoys some very nice properties.

Theorem 2. *In a second-price auction, it is a dominant strategy for every bidder to bid truthfully, i.e. to use strategy $s_i(v_i) = v_i$. In other words, truthful bidding maximizes the utility of bidder i no matter her value and no matter what the other bidders bid.*

Proof. Let us suppose that bidder i is bidding truthfully, that is, bidding her true value v_i . We will prove that there is no incentive for i to change her bid no matter what the other bidders' bids are. The proof will proceed in two cases.

- **Case 1: Bidder i does not win the item by bidding v_i .** In this case, the highest bid must be greater than or equal to v_i . Bidder i 's utility is 0, and she cannot possibly receive positive utility via some other bid; if i manages to win the item via some other bid, then her bid must be at least the highest bid, which is at least v_i , so the bidder's utility would be non-positive!
- **Case 2: Bidder i does win the item by bidding v_i .** In this case, v_i is greater than or equal to the second highest bid, and the bidder receives non-negative utility by bidding her value. If the bidder were to change her bid to any other bid above the second-highest bid, bidder i would still win the item and would not affect the price, which equals the second-highest bid. If, on the other hand, the bidder changed her bid to a bid below the second-highest bid, then she would lose the item and make her utility 0. In sum, no bid can possibly improve her utility.

□

It is thus trivial for bidders in a second-price auction to compute optimal strategies, even without distributional information about the other bidders' values. This is in sharp contrast to the first-price auction, where there is no meaningful way to choose optimal strategies without distributional information about the other bidders' values, and the Bayesian Nash equilibrium strategies are analytically challenging, and sensitive to both the number of bidders and the type distribution!

Another important property of the second-price auction is that truthful bidding guarantees a bidder non-negative utility, no matter how the others bid.

Lemma 1. *In a second-price auction, every truthful bidder is guaranteed non-negative utility.*

Proof. Again, the proof is trivial and proceeds by case analysis. Fix a truthful bidder i and denote her value by v_i . If she does *not* win, then her utility is 0. If she *does* win, then hers must be the highest bid, and her price is the second-highest bid, so her utility is non-negative. □

Theorem 3 ([2]). *The second-price auction satisfies the following properties:*

1. Dominant-Strategy Incentive-Compatibility (DSIC): *Truthful bidding is a dominant strategy Nash equilibrium.*
2. Individually Rationality (IR): *Truthful bidding guarantees non-negative utility to all bidders. In fact, if a bidder bids truthfully, her utility is non-negative no matter how the others bid.*
3. Welfare Maximization: *Under truthful bidding social welfare is maximized, i.e. the item is allocated to the bidder with the highest value for the item.*

4. Computational efficiency: *The auction can be implemented in polynomial (in fact, linear) time.*

To summarize, the above properties are criteria for a good auction. In future lectures we will tackle more complex allocation problems and address more complex objectives, such as revenue.

3 Mechanism Design Theory

In future lectures, we will consider more general mechanism design settings and objectives. To prepare the ground for this investigation, we provide next a more general mathematical formalism that we can use to design and analyze mechanisms.

Definition 9. *A mechanism design setting is defined by the following ingredients:*

1. *A set of bidders/players/agents $N = \{1, \dots, n\}$.*
2. *A set of alternatives A . This could be a very general set, e.g. where to build a hospital, which bidder gets which cloud resources, which candidate gets elected, etc.*
3. *For each bidder i , a type $t_i : A \rightarrow \mathbb{R}$. A bidder type assigns a value to each alternative. As such, it is sometimes called a valuation function, or simply a value. Set $T = T_1 \times \dots \times T_n$ is the set of type profiles. Sometimes a distribution F over T is common knowledge, such that $t = (t_1, \dots, t_n) \sim F$. In this case, the setting is called Bayesian.*

A mechanism for a mechanism design setting as above has the following ingredients:

1. *An action space X_i for each bidder i . Set $X = X_1 \times \dots \times X_n$ is the set of action profiles.*
2. *An allocation function $a : X \rightarrow A$, mapping action profiles to alternatives.*
3. *A price function $p_i : X \rightarrow \mathbb{R}$ for each bidder i , mapping action profiles to the price charged to bidder i .*

Finally, a mechanism is called direct iff $X_i = T_i$ for all i .

A mechanism induces a game of incomplete information and independent private values, where players' utilities are as follows:

$$u_i(t_i, x_1, \dots, x_n) = t_i(a(x_1, \dots, x_n)) - p_i(x_1, \dots, x_n).$$

As such, mechanisms can be analyzed by studying the properties of the incomplete information games they induce, specifically, by studying properties of their Dominant Strategy, ex-post Nash, or Bayesian Nash equilibria.

When we study *direct* mechanisms, we will be interested in whether truth-telling (i.e. strategy $s_i(t_i) = t_i$, for all i) is an equilibrium. Depending on what kind of equilibrium truth-telling may be, we get the following types of direct mechanisms.

Definition 10. *A direct mechanism (a, p) is Dominant Strategy Incentive Compatible (DSIC) iff truth-telling is a dominant strategy equilibrium, i.e. for all i , t_i , t'_i and t_{-i} :*

$$t_i(a(t_i, t_{-i})) - p_i(t_i, t_{-i}) \geq t_i(a(t'_i, t_{-i})) - p_i(t'_i, t_{-i}).$$

Definition 11. *A direct mechanism (a, p) is Bayesian Incentive Compatible (BIC) iff truth-telling is a Bayesian Nash equilibrium, i.e. for all i , t_i :*

$$\mathbb{E}_{t_{-i}}[t_i(a(t_i, t_{-i})) - p_i(t_i, t_{-i})] \geq \mathbb{E}_{t_{-i}}[t_i(a(t'_i, t_{-i})) - p_i(t'_i, t_{-i})].$$

3.1 Implementation

In this section, we discuss what it means for a mechanism to *implement* a social-choice function $f : T \rightarrow A$, mapping each type profile to the desired alternative under that type profile. The definition is fairly intuitive: a mechanism implements a social-choice function if there is an equilibrium under which the alternative chosen by the allocation rule of the mechanism coincides with that desired by social-choice function, even though the social-choice function computes on the true type profile, while the allocation rule computes on the action profile. Formally,

Definition 12. *We say that a mechanism (a, p) implements social-choice function f in dominant strategies if for some dominant strategy equilibrium $s = (s_1, \dots, s_n)$ of the incomplete information game induced by the mechanism, we have that, for all t_1, \dots, t_n ,*

$$a(s_1(t_1), \dots, s_n(t_n)) \equiv f(t_1, \dots, t_n),$$

where note that the left-hand-side is the outcome of the mechanism under equilibrium strategies, and the right-hand-side is the outcome desired by the social-choice function.

We define ex-post Nash implementation and Bayesian Nash implementation similarly, with the only modification being that we require s be an ex-post Nash equilibrium or a Bayesian Nash equilibrium, respectively.

Remark 2. *Note that our definition requires $a(s_1(t_1), \dots, s_n(t_n)) \equiv f(t_1, \dots, t_n)$ to hold for some equilibrium $s = (s_1, \dots, s_n)$, and not all equilibria. In practice we prefer that it holds for all equilibria.*

Remark 3. *Note that, in single-item settings, Vickrey auction implements the maximum social welfare function in dominant strategies, as $s_i(t_i) = t_i$ is a dominant strategy equilibrium, and maximum social welfare is achieved under this equilibrium.*

3.2 The Revelation Principle

Looking forward, as we investigate more complex mechanism design settings and objectives, could it be that non-truthful or indirect mechanisms are more powerful in terms of implementing social-choice functions compared to direct and truthful ones? The “Revelation Principle” tells us that, in very general settings, the answer is “no.”

Theorem 4. *If there is an arbitrary mechanism that implements some social-choice function f in dominant strategies, there is also a direct, DSIC mechanism that implements f , under truthful bidding. Moreover, the payments of the players in the direct mechanism are identical to those in the original mechanism, at equilibrium, point-wise for each type profile $t = (t_1, \dots, t_n)$.*

The idea behind the proof is *simulation*: given some mechanism M that implements a social-choice function under some equilibrium $s = (s_1, \dots, s_n)$, we can always define a new direct mechanism M' that simply asks each bidder i to report their type t_i to the mechanism, then simulates each bidder's equilibrium strategy to compute an action $s_i(t_i)$, and then feeds the resulting action profile into mechanism M to compute prices and allocations. Formally,

Proof. Let $s = (s_1, \dots, s_n)$ be a dominant strategy equilibrium of some mechanism $M = (a, p)$ such that, for all type profiles $t = (t_1, \dots, t_n)$,

$$a(s_1(t_1), \dots, s_n(t_n)) \equiv f(t_1, \dots, t_n).$$

Define a new, direct mechanism M' as follows:

$$\begin{aligned} a'(t_1, \dots, t_n) &= a(s_1(t_1), \dots, s_n(t_n)) \\ p'_i(t_1, \dots, t_n) &= p_i(s_1(t_1), \dots, s_n(t_n)), \quad \forall i \end{aligned}$$

For all i , since s_i is a dominant strategy for player i in the original mechanism M , we have that for every t_i, x_{-i}, x'_i :

$$t_i(a(s_i(t_i), x_{-i})) - p_i(s_i(t_i), x_{-i}) \geq t_i(a(x'_i, x_{-i})) - p_i(x'_i, x_{-i}).$$

So, in particular, the inequality holds if we were to set $x_{-i} = s_{-i}(t_{-i})$ and $x'_i = s_i(t'_i)$ for any t_{-i} and t'_i . Plugging these choices into the above inequality and substituting the definitions of a' and p' , we get that for all t_i, t'_i, t_{-i} :

$$t_i(a'(t_i, t_{-i})) - p'_i(t_i, t_{-i}) \geq t_i(a'(t'_i, t_{-i})) - p'_i(t'_i, t_{-i}).$$

Hence, mechanism M' is DSIC. Moreover, $a'(t_1, \dots, t_n) = f(t_1, \dots, t_n)$, so under truthful bidding M' implements f . Finally, at equilibrium, M and M' charge the same prices on a type profile by type profile basis. \square

We can provide analogous statements for ex-post Nash and Bayesian Nash implementation.

Theorem 5. *If there is an arbitrary mechanism that implements some social-choice function f in ex-post (resp. Bayesian) Nash equilibrium, then there is also a direct, DSIC (resp. Bayesian IC) mechanism that implements f . Moreover, the payments of the players in the direct mechanism are identical to those in the original mechanism, at equilibrium, point-wise for each type profile $t = (t_1, \dots, t_n)$.*

Proof. The claim for ex-post Nash implementation can be proven by invoking Theorem 4 and the following lemma, whose easy proof we omit.

Lemma 2 (From Ex-Post Nash to Dominant Strategy Equilibrium). *Let $s = (s_1, \dots, s_n)$ be an ex-post Nash equilibrium of some incomplete information game $(X_1, \dots, X_n; T_1, \dots, T_n; u_1, \dots, u_n)$. For all i , define the restricted action space $X'_i = \{s_i(t_i) | t_i \in T_i\}$. Then $s = (s_1, \dots, s_n)$ is a dominant strategy equilibrium of the game $(X'_1, \dots, X'_n; T_1, \dots, T_n; u_1, \dots, u_n)$.*

We continue with the proof of Theorem 5. Suppose M implements f under ex-post Nash equilibrium $s = (s_1, \dots, s_n)$. Restrict the action set X_i of each player i in mechanism M to the set $X'_i = \{s_i(t_i) | t_i \in T_i\}$, and keep the allocation and price rules of the mechanism the same. Lemma 2 implies that $s = (s_1, \dots, s_n)$ is a dominant strategy equilibrium of the resulting mechanism. Now we invoke the revelation principle for dominant strategy implementation (Theorem 4) to conclude that there exists a direct DSIC mechanism that implements f under truthtelling equilibrium. This concludes the proof of the claim for ex-post Nash implementation.

The claim for Bayesian Nash implementation is essentially identical to that of Theorem 4 and proceeds by simulation. There is no need to invoke Lemma 2. \square

References

- [1] Todd R Kaplan and Shmuel Zamir. Asymmetric first-price auctions with uniform distributions: analytic solutions to the general case. *Economic Theory*, 50(2):269–302, 2012.
- [2] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.