Estimating Approximate Incentive Compatibility

Maria-Florina Balcan
Carnegie Mellon University
ninamf@cs.cmu.edu

Tuomas Sandholm
Carnegie Mellon University
Optimized Markets, Inc.
Strategic Machine, Inc.
Strategy Robot, Inc.
sandholm@cs.cmu.edu

Ellen Vitercik
Carnegie Mellon University
vitercik@cs.cmu.edu

June 6, 2019

Abstract

In practice, most mechanisms for selling, buying, matching, voting, and so on are not incentive compatible. We present techniques for estimating how far a mechanism is from incentive compatible. Given samples from the agents' type distribution, we show how to estimate the extent to which an agent can improve his utility by misreporting his type. We do so by first measuring the maximum utility an agent can gain by misreporting his type on average over the samples, assuming his true and reported types are from a finite subset—which our technique constructs—of the type space. The challenge is that by measuring utility gains over a finite subset of the type space, we might miss pairs of types θ and $\hat{\theta}$ where an agent with type θ can greatly improve his utility by reporting the type $\hat{\theta}$. Our technique discretizes the type space by constructing a learning-theoretic cover in a higher-dimensional space. The key technical contribution is proving that the maximum utility gain over this finite subset nearly matches the maximum utility gain overall, despite the volatility of the utility functions we study. We apply our tools to the single-item and combinatorial first-price auctions, generalized second-price auction, discriminatory auction, uniform-price auction, and second-price auction with spiteful bidders. To our knowledge, these are the first guarantees for estimating approximate incentive compatibility from the mechanism designer's perspective.

1 Introduction

Incentive compatibility [Hurwicz, 1972] is a fundamental concept in mechanism design. Under an incentive compatible mechanism, it is in every agent's best interest to report their type truthfully. Nonetheless, practitioners have long employed mechanisms that are not incentive compatible, also called *manipulable* mechanisms. This is the case in many settings for selling, buying, matching (such as school choice), voting, and so on. For example, most real-world auctions are implemented using the first-price mechanism. In multi-unit sales settings, the U.S. Treasury has used discriminatory auctions, a variant of the first-price auction, to sell treasury bills since 1929 [Krishna, 2002]. Similarly, electricity generators in the U.K. use discriminatory auctions to sell their output [Krishna, 2002]. Sponsored search auctions are typically implemented using variants of the generalized second-price auction [Edelman et al., 2007, Varian, 2007]. In the past year, many major display ad exchanges including AppNexus, Rubicon, and Index Exchange have transitioned to first-price auctions, driven by ad buyers who believe it offers a higher degree of transparency [Parkin, 2018, Harada, 2018]. Finally, nearly all fielded combinatorial auctions are manipulable. Essentially all combinatorial sourcing auctions are implemented using the first-price mechanism [Sandholm, 2013].

Combinatorial spectrum auctions are conducted using a variety of manipulable mechanisms. Even the "incentive auction" used to source spectrum licenses back from low-value broadcasters—which has sometimes been hailed as obviously incentive compatible—is manipulable once one takes into account the fact that many owners own multiple broadcasting stations, or the fact that stations do not only have the option to keep or relinquish their license, but also the option to move to (two) less desirable spectrum ranges [Nguyen and Sandholm, 2015].

Many reasons have been suggested why manipulable mechanisms are used in practice. First, the rules are often easier to explain. Second, incentive compatibility ceases to hold even in the relatively simple context of the Vickrey auction when determining one's own valuation is costly (for example, due to computation or information gathering effort) [Sandholm, 2000]. Third, bidders may have even more incentive to behave strategically when they can conduct computation or information gathering on each others' valuations, and if they can incrementally decide how to allocate valuation-determination effort [Larson and Sandholm, 2001, 2005]. Fourth, in combinatorial settings, well-known incentive-compatible mechanisms such as the Vickrey-Clarke-Groves (VCG) mechanism require bidders to submit bids for every bundle, which generally requires a prohibitive amount of valuation computation (solving a local planning problem for potentially every bundle, and each planning problem, itself, can be NP-complete) or information acquisition [Sandholm, 1993, Parkes, 1999, Conen and Sandholm, 2001, Sandholm and Boutilier, 2006]. Fifth, in settings such as sourcing, single-shot incentive compatible mechanisms such as the VCG are generally not incentive compatible when the bid-taker uses bids from one auction to adjust the parameters of later auctions in later years [Sandholm, 2013]. Sixth, incentive compatible mechanisms may leak the agents' sensitive private information [Rothkopf et al., 1990]. Seventh, incentive compatibility typically ceases to hold if agents are not risk neutral (i.e., the utility functions are not quasi-linear). There are also sound theoretical reasons why the designer sometimes prefers manipulable mechanisms. Specifically, there exist settings where the designer does better than under any incentive compatible mechanism if the agents cannot solve hard computational or communication problems, and equally well if they can [Conitzer and Sandholm, 2004, Othman and Sandholm, 2009].

Due in part to the ubiquity of manipulable mechanisms, a growing body of additional research has explored mechanisms that are not incentive compatible [Kothari et al., 2003, Archer et al., 2004, Conitzer and Sandholm, 2007, Dekel et al., 2010, Lubin and Parkes, 2012, Mennle and Seuken, 2014, Dütting et al., 2015, Azevedo and Budish, 2018, Feng et al., 2018, Golowich et al., 2018, Dütting et al., 2017]. A popular and widely-studied relaxation of incentive compatibility is γ-*incentive compatibility* [Kothari et al., 2003, Archer et al., 2004, Dekel et al., 2010, Lubin and Parkes, 2012, Mennle and Seuken, 2014, Dütting et al., 2015, Azevedo and Budish, 2018], which requires that no agent can improve his utility by more than γ when he misreports his type.

1.1 Our contributions

Much of the literature on γ-incentive compatibility rests on the strong assumption that the agents' type distribution is known. In reality, this information is rarely available. We relax this assumption and instead assume we only have samples from the distribution [Likhodedov and Sandholm, 2004, 2005, Sandholm and Likhodedov, 2015]. We present techniques with provable guarantees that the mechanism designer can use to estimate how far a mechanism is from incentive compatible. We analyze both the *ex-interim* and *ex-ante*[1] settings: in the *ex-interim* case, we bound the amount any agent can improve his utility by misreporting his type, in expectation over the other agents'

[1] We do not study *ex-post* approximate incentive compatibility because it is a worst-case, distribution-independent notion. Therefore, we cannot hope to measure the *ex-post* approximation factor using samples from the agents' type distribution.

types, no matter his true type. In the weaker *ex-ante* setting, the expectation is also over the agent's true type as well.

Our estimate is simple: it measures the maximum utility an agent can gain by misreporting his type on average over the samples, whenever his true and reported types are from a finite subset of the type space. We bound the difference between our incentive compatibility estimate and the true incentive compatibility approximation factor γ. We are the first paper to provide theoretical guarantees for estimating approximate incentive compatibility from the mechanism designer's perspective, to our knowledge. In settings where we can solve for the true approximation factor γ, we provide experiments demonstrating that our estimates quickly converge to γ.

We apply our estimation technique to a variety of auction classes. We begin with the first-price auction, in both single-item and combinatorial settings. Our guarantees can be used by display ad exchanges, for instance, to measure the extent to which incentive compatibility will be compromised if the exchange transitions to using first-price auctions [Parkin, 2018, Harada, 2018]. In the single-item setting, we prove that the difference between our estimate and the true incentive compatibility approximation factor is $\tilde{O}\left(\left(n + \kappa^{-1}\right)/\sqrt{N}\right)$, where n is the number of bidders, N is the number of samples, and $[0, \kappa]$ contains the range of the density functions defining agents' type distribution. We prove the same bound for the second-price auction with *spiteful bidders* [Brandt et al., 2007, Morgan et al., 2003, Sharma and Sandholm, 2010, Tang and Sandholm, 2012], where each bidder's utility not only increases when his surplus increases but also decreases when the other bidders' surpluses increase.

In a similar direction, we analyze the class of generalized second-price auctions [Edelman et al., 2007], where m sponsored search slots are for sale. The mechanism designer assigns a real-valued weight per bidder, collects a bid per bidder indicating their value per click, and allocates the slots in order of the bidders' weighted bids. In this setting, we prove that the difference between our incentive compatibility estimate and the true incentive compatibility approximation bound is $\tilde{O}\left(\left(n^{3/2} + \kappa^{-1}\right)/\sqrt{N}\right)$.

We also analyze multi-parameter mechanisms beyond the first-price combinatorial auction, namely, the uniform-price and discriminatory auctions, which are used extensively in markets around the world [Krishna, 2002]. In both, the auctioneer has m identical units of a single good to sell. Each bidder submits m bids indicating their value for each additional unit of the good. The number of goods the auctioneer allocates to each bidder equals the number of bids that bidder has in the top m bids. In both cases, we prove that the difference between our incentive compatibility estimate and the true incentive compatibility approximation bound is $\tilde{O}\left(\left(nm^2 + \kappa^{-1}\right)/\sqrt{N}\right)$.

A strength of our estimation techniques is that they are application-agnostic. For example, they can be used as a tool in *incremental mechanism design* [Conitzer and Sandholm, 2007], a subfield of *automated mechanism design* [Conitzer and Sandholm, 2002, Sandholm, 2003], where the mechanism designer gradually adds incentive compatibility constraints to her optimization problem until she has met a desired incentive compatibility guarantee. One line of work in the spirit of incremental mechanism design has studied mechanism design via deep learning [Feng et al., 2018, Dütting et al., 2017, Golowich et al., 2018]. The learning algorithm receives samples from the distribution over agents' types. The resulting allocation and payment functions are characterized by neural networks, and thus the corresponding mechanism may not be incentive compatible. In an attempt to make these mechanisms nearly incentive compatible, the authors of these works add constraints to the deep learning optimization problem enforcing that the resulting mechanism be incentive compatible over a set of buyer values sampled from the underlying, unknown distribution. However, they provide no guarantees indicating how far the resulting mechanism is from incentive compatible. One of the goals of this paper is to provide techniques for relating a mechanism's empirical incentive compatibility approximation factor to its expected incentive compatibility approximation factor.

Key challenges. To prove our guarantees, we must estimate the value γ defined such that no agent can misreport his type in order to improve his expected utility by more than γ, no matter his true type. We propose estimating γ by measuring the extent to which an agent can improve his utility by misreporting his type on average over the samples, whenever his true and reported types are from a finite subset[2] $\mathcal{G}$ of the type space. We denote this estimate as $\hat{\gamma}$. The challenge is that by searching over a subset of the type space, we might miss pairs of types θ and $\hat{\theta}$ where an agent with type θ can greatly improve his expected utility by misreporting his type as $\hat{\theta}$. Indeed, utility functions are often volatile in mechanism design settings. For example, under the first- and second-price auctions, nudging an agent's bid from below the other agents' largest bid to above will change the allocation, causing a jump in utility. Thus, there are two questions we must address: which finite subset should we search over and how do we relate $\hat{\gamma}$ to γ?

We provide two approaches to constructing the cover $\mathcal{G}$. The first is to run a greedy procedure based off a classic algorithm from learning theory. This approach is extremely versatile: it provides strong guarantees no matter the setting. However, depending on the domain, it may be difficult to implement. Meanwhile, implementing our second approach is straightforward: the cover is simply a uniform grid over the type space (assuming the type space equals $[0,1]^m$ for some integer m). The efficacy of this approach depends on a "niceness" property that holds under mild assumptions. To analyze this second approach, we must understand how the edge-length of the grid effects our error bound relating $\hat{\gamma}$ to γ. To do so, we rely on the notion of *dispersion*, introduced by Balcan et al. [2018a] in the context of online and batch learning as well as private optimization. Roughly speaking, a set of piecewise Lipschitz functions is (w,k)-dispersed if every ball of radius w in the domain contains at most k of the functions' discontinuities. Given a set of N samples from the distribution over agents' types, we analyze the set of N functions measuring the utility of an agent with type $\boldsymbol{\theta}$ and misreported type $\hat{\boldsymbol{\theta}}$ when the other agents' true and misreported types are represented by one of the N samples. We show that if these functions are (w,k)-dispersed, we can use a grid with edge-length w to discretize the agents' type space. We then prove that if the intrinsic complexity of the agents' utility functions are not too large (as measured by the learning-theoretic notion of *pseudo-dimension* [Pollard, 1984]), then $\hat{\gamma}$ quickly converges to γ as the number of samples grows.

Finally, we show that for a wide range of mechanism classes, dispersion holds under mild assumptions. As we describe in Section 3.1.2, this requires us to prove that with high probability, each function sequence from an infinite family of sequences is dispersed. This facet of our analysis is notably different from prior research by Balcan et al. [2018a]: in their applications, it is enough to show that with high probability, a single, finite sequence of functions is dispersed. Our proofs thus necessitate that we carefully examine the structure of the utility functions that we analyze.

1.2 Additional related research

Sample complexity of revenue maximization. A long line of research has studied revenue maximization via machine learning from a theoretical perspective [Balcan et al., 2008, Alon et al., 2017, Elkind, 2007, Cole and Roughgarden, 2014, Huang et al., 2015, Medina and Mohri, 2014, Morgenstern and Roughgarden, 2015, Roughgarden and Schrijvers, 2016, Devanur et al., 2016, Gonczarowski and Nisan, 2017, Bubeck et al., 2017, Morgenstern and Roughgarden, 2016, Balcan

[2]This is the approach also taken in mechanism design via deep learning [Dütting et al., 2017, Feng et al., 2018, Golowich et al., 2018], but without guarantees.

et al., 2016, Syrgkanis, 2017, Medina and Vassilvitskii, 2017, Balcan et al., 2018b, Gonczarowski and Weinberg, 2018, Cai and Daskalakis, 2017]. The mechanism designer receives samples from the type distribution which she uses to find a mechanism that is, ideally, nearly optimal in expectation. That research has only studied incentive compatible mechanism classes. Moreover, in this paper, it is not enough to provide generalization guarantees; we must both compute our estimate of the incentive compatibility approximation factor and bound our estimate's error. This type of error has nothing to do with revenue functions, but rather utility functions. These factors in conjunction mean that our research differs significantly from prior research on generalization guarantees in mechanism design.

In a related direction, Chawla et al. [2014, 2016, 2017] study counterfactual revenue estimation. Given two auctions, they provide techniques for estimating one of the auction's equilibrium revenue from the other auction's equilibrium bids. They also study social welfare in this context. Thus, their research is tied to selling mechanisms, whereas we study more general mechanism design problems from an application-agnostic perspective.

Strategy-proofness in the large. Azevedo and Budish [2018] propose a variation on approximate incentive compatibility called *strategy-proofness in the large (SP-L)*. SP-L requires that it is approximately optimal for agents to report their types truthfully in sufficiently large markets. As in our paper, SP-L is a condition on *ex-interim* γ-incentive compatibility. The authors argue that SP-L approximates, in large markets, attractive properties of a mechanism such as strategic simplicity and robustness. They categorize a number of mechanisms as either SP-L or not. For example, they show that the discriminatory auction is manipulable in the large whereas the uniform-price auction is SP-L. Measuring a mechanism's SP-L approximation factor requires knowledge of the distribution over agents' types, whereas we only require sample access to this distribution. Moreover, we do not make any large-market assumptions: our guarantees hold regardless of the number of agents.

Comparing mechanisms by their vulnerability to manipulation. Pathak and Sönmez [2013] analyze *ex-post* incentive compatibility without any connection to approximate incentive compatibility. They say that one mechanism M is at least as manipulable as another M' if every type profile that is vulnerable to manipulation under M is also vulnerable to manipulation under M'. They apply their formalism in the context of school assignment mechanisms, the uniform-price auction, the discriminatory auction, and several keyword auctions. We do not study *ex-post* approximate incentive compatibility because it is a worst-case, distribution-independent notion. Therefore, we cannot hope to measure an *ex-post* approximation factor using samples from the agents' type distribution. Rather, we are concerned with providing data-dependent bounds on the *ex-interim* and *ex-ante* approximation factors. Another major difference is that our work provides quantitative results on manipulability while theirs provides boolean comparisons as to the relative manipulability of mechanisms. Finally, our measure applies to all mechanisms while theirs cannot rank all mechanisms because in many settings, pairs of mechanisms are incomparable according to their boolean measure.

Incentive compatibility from a buyer's perspective. Lahaie et al. [2018] also provide tools for estimating approximate incentive compatibility, but from the buyer's perspective rather than the mechanism designer's perspective. As such, the type of information available to their estimation tools versus ours is different. Moreover, they focus on ad auctions, whereas we study mechanism design in general and apply our techniques to a wide range of settings and mechanisms.

2 Preliminaries and notation

There are n agents who each have a *type*. We denote agent i's type as θ_i, which is an element of a (potentially infinite) set Θ_i. A mechanism takes as input the agents' *reported types*, which it uses to choose an outcome. We denote agent i's reported type as $\hat{\theta}_i \in \Theta_i$. We denote all n agents' types as $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ and reported types as $\hat{\boldsymbol{\theta}} = \left(\hat{\theta}_1, \dots, \hat{\theta}_n\right)$. For $i \in [n]$, we use the standard notation $\boldsymbol{\theta}_{-i} \in \times_{j \neq i} \Theta_j$ to denote all $n-1$ agents' types except agent i. Using this notation, we denote the type profile $\boldsymbol{\theta}$ representing all n agents' types as $\boldsymbol{\theta} = (\theta_i, \boldsymbol{\theta}_{-i})$. Similarly, $\hat{\boldsymbol{\theta}} = \left(\hat{\theta}_i, \hat{\boldsymbol{\theta}}_{-i}\right)$. We assume there is a distribution $\mathcal{D}$ over all n agents' types, and thus the support of $\mathcal{D}$ is contained in $\times_{i=1}^n \Theta_i$. We use $\mathcal{D}|_{\theta_i}$ to denote the conditional distribution given θ_i, so the support of $\mathcal{D}|_{\theta_i}$ is contained in $\times_{j \neq i} \Theta_j$. We assume that we can draw samples $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots$ independently from $\mathcal{D}$. This is the same assumption made in a long line of work on mechanism design via machine learning [Balcan et al., 2008, Alon et al., 2017, Elkind, 2007, Cole and Roughgarden, 2014, Huang et al., 2015, Medina and Mohri, 2014, Morgenstern and Roughgarden, 2015, Roughgarden and Schrijvers, 2016, Devanur et al., 2016, Gonczarowski and Nisan, 2017, Bubeck et al., 2017, Morgenstern and Roughgarden, 2016, Balcan et al., 2016, Syrgkanis, 2017, Medina and Vassilvitskii, 2017, Balcan et al., 2018b, Gonczarowski and Weinberg, 2018, Cai and Daskalakis, 2017, Dütting et al., 2017, Feng et al., 2018, Golowich et al., 2018, Likhodedov and Sandholm, 2004, 2005, Sandholm and Likhodedov, 2015].

Given a mechanism M and agent $i \in [n]$, we use the notation $u_{i,M}\left(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}\right)$ to denote the utility agent i receives when the agents have types $\boldsymbol{\theta}$ and reported types $\hat{\boldsymbol{\theta}}$. We assume it maps to $[-1, 1]$. When $\boldsymbol{\theta}_{-i} = \hat{\boldsymbol{\theta}}_{-i}$, we use the simplified notation $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) = u_{i,M}\left((\theta_i, \boldsymbol{\theta}_{-i}), \left(\hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right)$.

At a high level, a mechanism is incentive compatible if no agent can ever increase her utility by misreporting her type. A mechanism is γ-*incentive compatible* if each agent can increase her utility by an additive factor of at most γ by misreporting her type [Kothari et al., 2003, Archer et al., 2004, Conitzer and Sandholm, 2007, Dekel et al., 2010, Lubin and Parkes, 2012, Mennle and Seuken, 2014, Dütting et al., 2015, Azevedo and Budish, 2018, Feng et al., 2018, Golowich et al., 2018, Dütting et al., 2017]. In the main body, we concentrate on *ex-interim* approximate incentive compatibility [Azevedo and Budish, 2018, Lubin and Parkes, 2012].

Definition 2.1. A mechanism M is ex-interim γ-*incentive compatible* if for each $i \in [n]$ and all $\theta_i, \hat{\theta}_i \in \Theta_i$, agent i with type θ_i can increase her expected utility by an additive factor of at most γ by reporting her type as $\hat{\theta}_i$, so long as the other agents report truthfully. In other words, $\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}|_{\theta_i}}\left[u_{i,M}(\theta_i, \theta_i, \boldsymbol{\theta}_{-i})\right] \geq E_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}|_{\theta_i}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \gamma$.

In Appendix F, we study *ex-ante* approximate incentive compatibility, where the above definition holds in expectation over $\boldsymbol{\theta} \sim \mathcal{D}$.

3 Estimating approximate ex-interim incentive compatibility

In this section, we show how to estimate the *ex-interim* incentive compatibility approximation guarantee using data. We assume there is an unknown distribution $\mathcal{D}$ over agents' types, and we operate under the common assumption [Lubin and Parkes, 2012, Azevedo and Budish, 2018, Cai and Daskalakis, 2017, Yao, 2014, Cai et al., 2016, Goldner and Karlin, 2016, Babaioff et al., 2017, Hart and Nisan, 2012] that the agents' types are independently distributed. In other words, for each agent $i \in [n]$, there exists a distribution ϕ_i over Θ_i such that $\mathcal{D} = \times_{i=1}^n \phi_i$. (In Appendix F, we extend our analysis to approximate *ex-ante* incentive compatibility under no assumption about the underlying distribution, which is a weaker notion of incentive compatibility, but also a weaker assumption on the distribution.) For each agent $i \in [n]$, we receive a set $\mathcal{S}_{-i}$ of samples independently drawn from $\mathcal{D}_{-i} = \times_{j \neq i} \phi_j$. For each mechanism $M \in \mathcal{M}$, we show how to use the samples to estimate a value $\hat{\gamma}_M$ such that:

> *With probability* $1-\delta$ *over the draw of the* n *sets of samples* $\mathcal{S}_{-1}, \ldots, \mathcal{S}_{-n}$, *for any agent* $i \in [n]$ *and all pairs* $\theta_i, \hat{\theta}_i \in \Theta_i$, $\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \leq \hat{\gamma}_M$.

To this end, one simple approach, informally, is to estimate $\hat{\gamma}_M$ by measuring the extent to which any agent i with any type θ_i can improve his utility by misreporting his type, averaged over all profiles in $\mathcal{S}_{-i}$. In other words, we can estimate $\hat{\gamma}_M$ by solving the following optimization problem:

$$\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\}. \tag{1}$$

Unfortunately, in full generality, there might not be a finite-time procedure to solve this optimization problem, so in Section 3.1, we propose more nuanced approaches based on optimizing over finite subsets of $\Theta_i \times \Theta_i$. As a warm-up and a building block for our main theorems in that section, we prove that with probability $1-\delta$, for all mechanisms $M \in \mathcal{M}$,

$$\begin{aligned} &\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \right\} \\ \leq &\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + \epsilon_{\mathcal{M}}(N, \delta), \end{aligned} \tag{2}$$

where $\epsilon_{\mathcal{M}}(N, \delta)$ is an error term that converges to zero as the number N of samples grows. Its convergence rate depends on the *intrinsic complexity* of the utility functions corresponding to the mechanisms in $\mathcal{M}$, which we formalize using the learning-theoretic tool *pseudo-dimension.* We define pseudo-dimension below for an abstract class $\mathcal{A}$ of functions mapping a domain $\mathcal{X}$ to $[0,1]$.

Definition 3.1 (Pseudo-dimension [Pollard, 1984]). Let $\mathcal{S} = \left\{x^{(1)}, \ldots, x^{(N)}\right\} \subseteq \mathcal{X}$ be a set of elements from the domain of $\mathcal{A}$ and let $z^{(1)}, \ldots, z^{(N)} \in \mathbb{R}$ be a set of *targets*. We say that $z^{(1)}, \ldots, z^{(N)}$ *witness* the shattering of $\mathcal{S}$ by $\mathcal{A}$ if for all subsets $T \subseteq \mathcal{S}$, there exists some function $a_T \in \mathcal{A}$ such that for all $x^{(j)} \in T$, $a_T\left(x^{(j)}\right) \leq z^{(j)}$ and for all $x^{(j)} \notin T$, $a_T\left(x^{(j)}\right) > z^{(j)}$. If there exists some vector $\boldsymbol{z} \in \mathbb{R}^N$ that witnesses the shattering of $\mathcal{S}$ by $\mathcal{A}$, then we say that $\mathcal{S}$ is *shatterable* by $\mathcal{A}$. Finally, the pseudo-dimension of $\mathcal{A}$, denoted $\operatorname{Pdim}(\mathcal{A})$, is the size of the largest set that is shatterable by $\mathcal{A}$.

Theorem 3.1 provides an abstract generalization bound in terms of pseudo-dimension.

Theorem 3.1 (Pollard [1984]). *Let* Φ *be a distribution over* $\mathcal{X}$. *With probability* $1-\delta$ *over* $x^{(1)}, \ldots, x^{(N)} \sim \Phi$, *for all* $a \in \mathcal{A}$, $\left|\frac{1}{N}\sum_{j=1}^{N} a\left(x^{(j)}\right) - \mathbb{E}_{x \sim \Phi}[a(x)]\right| \leq \sqrt{\frac{2d}{N}\ln\frac{eN}{d}} + \sqrt{\frac{1}{2N}\ln\frac{2}{\delta}}$, *where* $d = \operatorname{Pdim}(\mathcal{A})$.

We now use Theorem 3.1 to prove that the error term $\epsilon_{\mathcal{M}}(N, \delta)$ in Equation (2) converges to zero as N increases. To this end, for any mechanism M, any agent $i \in [n]$, and any pair of types $\theta_i, \hat{\theta}_i \in \Theta_i$, let $u_{i,M,\theta_i,\hat{\theta}_i} : \times_{j \neq i} \Theta_j \to [-1, 1]$ be a function that maps the types $\boldsymbol{\theta}_{-i}$

of the other agents to the utility of agent i with type θ_i and reported type $\hat{\theta}_i$ when the other agents report their types truthfully. In other words, $u_{i,M,\theta_i,\hat{\theta}_i}(\boldsymbol{\theta}_{-i}) = u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$. Let $\mathcal{F}_{i,\mathcal{M}}$ be the set of all such functions defined by mechanisms M from the class $\mathcal{M}$. In other words, $\mathcal{F}_{i,\mathcal{M}} = \left\{ u_{i,M,\theta_i,\hat{\theta}_i} \,\middle|\, \theta_i, \hat{\theta}_i \in \Theta_i, M \in \mathcal{M} \right\}$. We now analyze the convergence rate of the error term $\epsilon_{\mathcal{M}}(N, \delta)$. The full proof is in Appendix C.

Theorem 3.2. *With probability $1 - \delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for all mechanisms $M \in \mathcal{M}$ and agents $i \in [n]$,*

$$\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}} \left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right) \right] \right\}$$
$$\leq \max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + \epsilon_{\mathcal{M}}(N, \delta),$$

where $\epsilon_{\mathcal{M}}(N, \delta) = 2\sqrt{\frac{2d_i}{N} \ln \frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N} \ln \frac{2n}{\delta}}$ and $d_i = \operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$.

Proof sketch. Fix an arbitrary bidder $i \in [n]$. By Theorem 3.1, we know that with probability $1 - \delta/n$, for every pair of types $\theta_i, \hat{\theta}_i \in \Theta_i$ and every mechanism $M \in \mathcal{M}$, the expected utility of agent i with type θ_i and reported type $\hat{\theta}_i$ (in expectation over $\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}$) is ϵ-close to his average utility (averaged over $\boldsymbol{\theta}_{-i} \in \mathcal{S}_{-i}$), where $\epsilon = \tilde{O}\left(\sqrt{d_i/N}\right)$. We use this fact to show that when agent i misreports his type, he cannot increase his expected utility by more than a factor of $2\epsilon + \max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\}$. □

3.1 Incentive compatibility guarantees via finite covers

In the previous section, we presented an empirical estimate of the *ex-interim* incentive compatibility approximation factor (Equation (1)) and we showed that it quickly converges to the true approximation factor. However, there may not be a finite-time procedure for computing Equation (1) in its full generality, restated below:

$$\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\}. \tag{3}$$

In this section, we address that challenge. A simple alternative approach is to fix a finite *cover* of $\Theta_i \times \Theta_i$, which we denote as $\mathcal{G} \subset \Theta_i \times \Theta_i$, and approximate Equation (3) by measuring the extent to which any agent i can improve his utility by misreporting his type when his true and reported types are elements of the cover $\mathcal{G}$, averaged over all profiles in $\mathcal{S}_{-i}$. In other words, we estimate Equation (3) as:

$$\max_{(\theta_i, \hat{\theta}_i) \in \mathcal{G}} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\}. \tag{4}$$

This is the approach also taken in the recent line of research on mechanism design via deep learning [Dütting et al., 2017, Feng et al., 2018, Golowich et al., 2018]. This raises two natural questions: how do we select the cover $\mathcal{G}$ and how close are the optimal solutions to Equations (3) and (4)? We provide two simple, intuitive approaches to selecting the cover $\mathcal{G}$. The first is to run a greedy procedure (see Section 3.1.1) and the second is to create a uniform grid over the type space (assuming $\Theta_i = [0, 1]^m$ for some integer m; see Section 3.1.2).

Algorithm 1 Greedy cover construction

Input: Mechanism $M \in \mathcal{M}$, set of samples $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \ldots, \boldsymbol{\theta}_{-i}^{(N)}\right\}$, accuracy parameter $\epsilon > 0$.

1: Let U be the set of vectors $U \leftarrow \left\{ \frac{1}{N} \begin{pmatrix} u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(1)}\right) \\ \vdots \\ u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(N)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(N)}\right) \end{pmatrix} : \theta_i, \hat{\theta}_i \in \Theta_i \right\}$.

2: Let $V \leftarrow \emptyset$ and $\mathcal{G} \leftarrow \emptyset$.

3: **while** $U \setminus \left(\bigcup_{\boldsymbol{v} \in V} B_1(\boldsymbol{v}, \epsilon)\right) \neq \emptyset$ **do**

4: Select an arbitrary vector $\boldsymbol{v}' \in U \setminus \left(\bigcup_{\boldsymbol{v} \in V} B_1(\boldsymbol{v}, \epsilon)\right)$.

5: Let $\theta_i, \hat{\theta}_i \in \Theta_i$ be the types such that $\boldsymbol{v}' = \frac{1}{N} \begin{pmatrix} u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(1)}\right) \\ \vdots \\ u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(N)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(N)}\right) \end{pmatrix}$.

6: Add $\boldsymbol{v}'$ to V and $\left(\theta_i, \hat{\theta}_i\right)$ to $\mathcal{G}$.

Output: The cover $\mathcal{G} \subseteq \Theta_i \times \Theta_i$.

3.1.1 Covering via a greedy procedure

In this section, we show how to construct the cover $\mathcal{G}$ of $\Theta_i \times \Theta_i$ greedily, based off a classic learning-theoretic algorithm. We then show that when we use the cover $\mathcal{G}$ to estimate the incentive compatibility approximation factor (via Equation (4)), the estimate quickly converges to the true approximation factor. This greedy procedure is summarized by Algorithm 1. For any $\boldsymbol{v} \in \mathbb{R}^N$, we use the notation $B_1(\boldsymbol{v}, \epsilon) = \{\boldsymbol{v}' : \|\boldsymbol{v}' - \boldsymbol{v}\|_1 \leq \epsilon\}$. To simplify notation, let U be the set of vectors defined in Algorithm 1:

$$U = \left\{ \frac{1}{N} \begin{pmatrix} u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(1)}\right) \\ \vdots \\ u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(N)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(N)}\right) \end{pmatrix} : \theta_i, \hat{\theta}_i \in \Theta_i \right\}.$$

Note that the solution to Equation (3) equals $\max_{\boldsymbol{v} \in U} \sum_{i=1}^N v[i]$. The algorithm greedily selects a set of vectors $V \subseteq U$, or equivalently, a set of type pairs $\mathcal{G} \subseteq \Theta_i \times \Theta_i$ as follows: while $U \setminus \left(\bigcup_{\boldsymbol{v} \in V} B_1(\boldsymbol{v}, \epsilon)\right)$ is non-empty, it chooses an arbitrary vector $\boldsymbol{v}'$ in the set, adds it to V, and adds the pair $\left(\theta_i, \hat{\theta}_i\right) \in \Theta_i \times \Theta_i$ defining the vector $\boldsymbol{v}'$ to $\mathcal{G}$. Classic results from learning theory [Anthony and Bartlett, 2009] guarantee that this greedy procedure will repeat for at most $(8eN/(\epsilon d_i))^{2d_i}$ iterations, where $d_i = \text{Pdim}(\mathcal{F}_{i,\mathcal{M}})$.

We now relate the true incentive compatibility approximation factor to the solution to Equation (4) when the cover is constructed using Algorithm 1. The full proof is in Appendix C.

Theorem 3.3. *Given a set $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \ldots, \boldsymbol{\theta}_{-i}^{(N)}\right\}$, a mechanism $M \in \mathcal{M}$, and accuracy parameter $\epsilon > 0$, let $\mathcal{G}(\mathcal{S}_{-i}, M, \epsilon)$ be the cover returned by Algorithm 1. With probability $1 - \delta$ over the draw* *of the* n *sets* $\mathcal{S}_{-i} \sim \mathcal{D}_{-i}^N$*, for every mechanism* $M \in \mathcal{M}$ *and every agent* $i \in [n]$*,*

$$\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}} \left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right) \right] \right\}$$

$$\leq \max_{\left(\theta_i, \hat{\theta}_i\right) \in \mathcal{G}(\mathcal{S}_{-i}, M, \epsilon)} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + \epsilon + \tilde{O}\left(\sqrt{\frac{d_i}{N}}\right), \tag{5}$$

where $d_i = \operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$*. Moreover, with probability 1,* $\left|\mathcal{G}\left(\mathcal{S}_{-i}, M, \epsilon\right)\right| \leq \left(8eN/\left(\epsilon d_i\right)\right)^{2d_i}$*.*

Proof sketch. Inequality (5) follows from Theorem 3.2, the fact that V is an ϵ-cover of U (where U and V are the sets defined in Algorithm 1), and the equivalence of $\mathcal{G}\left(\mathcal{S}_{-i}, M, \epsilon\right)$ and V. The bound on $\left|\mathcal{G}\left(\mathcal{S}_{-i}, M, \epsilon\right)\right|$ follows from classic learning-theoretic results [Anthony and Bartlett, 2009]. □

3.1.2 Covering via a uniform grid

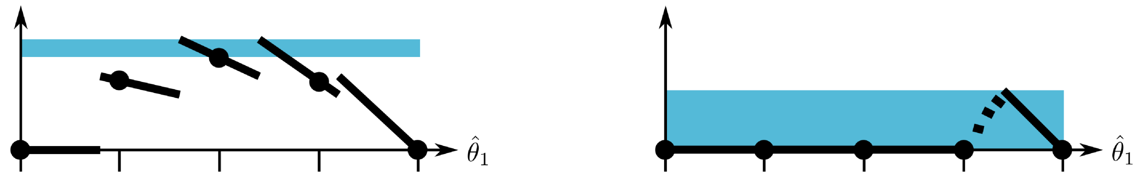
The greedy approach in Section 3.1.1 is extremely versatile: no matter the type space Θ_i, when we use the resulting cover to estimate the incentive compatibility approximation factor (via Equation (4)), the estimate quickly converges to the true approximation factor. However, implementing the greedy procedure (Algorithm 1) might be computationally challenging because at each round, it is necessary to check if $U \setminus \left(\bigcup_{\boldsymbol{v} \in V} B_1\left(\boldsymbol{v}, \epsilon\right)\right)$ is nonempty and if so, select a vector from the set. In this section, we propose an alternative, extremely simple approach to selecting a cover $\mathcal{G}$: using a uniform grid over the type space. The efficacy of this approach depends on a "niceness" assumption that holds under mild assumptions, as we prove in Section 3.2. Throughout this section, we assume that $\Theta_i = [0,1]^m$ for some integer m. We propose covering the type space using a *w-grid* $\mathcal{G}_w$ over $[0,1]^m$, by which we mean a finite set of vectors in $[0,1]^m$ such that for all $\boldsymbol{p} \in [0,1]^m$, there exists a vector $\boldsymbol{p}' \in \mathcal{G}_w$ such that $\|\boldsymbol{p} - \boldsymbol{p}'\|_1 \leq w$. For example, if $m = 1$, we could define $\mathcal{G}_w = \left\{0, \frac{1}{\lceil 1/w \rceil}, \frac{2}{\lceil 1/w \rceil}, \dots, 1\right\}$. We will estimate the expected incentive compatibility approximation factor using Equation (4) with $\mathcal{G} = \mathcal{G}_w \times \mathcal{G}_w$. Throughout the rest of the paper, we discuss how the choice of w effects the error bound. To do so, we will use the notion of *dispersion*, defined below.

Definition 3.2 (Balcan et al. [2018a])**.** Let $a_1, \dots, a_N : \mathbb{R}^d \to \mathbb{R}$ be a set of functions where each a_i is piecewise Lipschitz with respect to the ℓ_1-norm over a partition $\mathcal{P}_i$ of $\mathbb{R}^d$. We say that $\mathcal{P}_i$ splits a set $A \subseteq \mathbb{R}^d$ if A intersects with at least two sets in $\mathcal{P}_i$. The set of functions is (w,k)-*dispersed* if for every point $\boldsymbol{p} \in \mathbb{R}^d$, the ball $\left\{\boldsymbol{p}' \in \mathbb{R}^d : \|\boldsymbol{p} - \boldsymbol{p}'\|_1 \leq w\right\}$ is split by at most k of the partitions $\mathcal{P}_1, \dots, \mathcal{P}_N$.

The smaller w is and the larger k is, the more "dispersed" the functions' discontinuities are. Moreover, the more jump discontinuities a set of functions has, the more difficult it is to approximately optimize its average using a grid. We illustrate this phenomenon in Example 3.1.

Example 3.1. Suppose there are two agents and M is the first-price single-item auction. For any trio of types $\theta_1, \theta_2, \hat{\theta}_1 \in [0,1]$, $u_{1,M}\left(\theta_1, \hat{\theta}_1, \theta_2\right) = \mathbf{1}_{\left\{\hat{\theta}_1 > \theta_2\right\}}\left(\theta_1 - \hat{\theta}_1\right)$. Suppose $\theta_1 = 1$. Figure 1a displays the function $\frac{1}{4}\sum_{\theta_2 \in \mathcal{S}_2} u_{1,M}\left(1, \hat{\theta}_1, \theta_2\right)$ where $\mathcal{S}_2 = \left\{\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}\right\}$ and Figure 1b displays the function $\frac{1}{4}\sum_{\theta_2 \in \mathcal{S}_2'} u_{1,M}\left(1, \hat{\theta}_1, \theta_2\right)$ where $\mathcal{S}_2' = \left\{\frac{31}{40}, \frac{32}{40}, \frac{33}{40}, \frac{34}{40}\right\}$.

In Figure 1, we evaluate each function on the grid $\mathcal{G} = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$. In Figure 1a, the maximum



(a) Average of dispersed functions. (b) Average of functions that are not dispersed.

Figure 1: Illustration of Example 3.1. The lines in Figure 1a depict the average of four utility functions corresponding to the first-price auction. The maximum of this average falls at the top of the blue region. We evaluate the function in Figure 1a on the grid $\left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$, as depicted by the dots. The maximum over the grid falls at the bottom of the blue region. In Figure 1b, we illustrate the same concepts but for a different set of utility functions that are not as dispersed as those in Figure 1a. As illustrated by the blue regions, the approximation over the grid is better for the dispersed functions than the non-dispersed functions.

over $\mathcal{G}$ better approximates the the maximum over $[0,1]$ compared to Figure 1b. In other words,

$$\max_{\hat{\theta}_1 \in [0,1]} \sum_{\theta_2 \in \mathcal{S}_2} u_{1,M}\left(1, \hat{\theta}_1, \theta_2\right) - \max_{\hat{\theta}_1 \in \mathcal{G}} \sum_{\theta_2 \in \mathcal{S}_2} u_{1,M}\left(1, \hat{\theta}_1, \theta_2\right) \tag{6}$$

$$< \max_{\hat{\theta}_1 \in [0,1]} \sum_{\theta_2 \in \mathcal{S}_2'} u_{1,M}\left(1, \hat{\theta}_1, \theta_2\right) - \max_{\hat{\theta}_1 \in \mathcal{G}} \sum_{\theta_2 \in \mathcal{S}_2'} u_{1,M}\left(1, \hat{\theta}_1, \theta_2\right). \tag{7}$$

Intuitively, this is because the functions we average over in Figure 1a are more dispersed than the functions we average over in Figure 1b. The differences described by Equations (6) and (7) are represented by the shaded regions in Figures 1a and 1b, respectively.

We now state a helpful lemma which we use to prove this section's main theorem. Informally, it shows that we can measure the average amount that any agent can improve his utility by misreporting his type, even if we discretize his type space, so long as his utility function applied to the samples demonstrates dispersion. The full proof is in Appendix C.

Lemma 3.4. *Suppose that for each agent $i \in [n]$, there exist $L_i, k_i, w_i \in \mathbb{R}$ such that with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for each mechanism $M \in \mathcal{M}$ and agent $i \in [n]$, the following conditions hold:*

1. *For any type $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

2. *For any reported type $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, the functions $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

Then with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} \sim \mathcal{D}_{-i}^N$, for all mechanisms $M \in \mathcal{M}$ and agents $i \in [n]$,

$$\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\}$$

$$\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + 4L_i w_i + \frac{8k_i}{N}. \tag{8}$$

Proof sketch. By definition of dispersion, the following conditions hold with probability $1-\delta$:

Condition 1. For all mechanisms $M \in \mathcal{M}$, agents $i \in [n]$, types $\boldsymbol{\theta}_i \in [0,1]^m$, and pairs of reported types $\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_i' \in [0,1]^m$, if $\left\|\hat{\boldsymbol{\theta}}_i - \hat{\boldsymbol{\theta}}_i'\right\|_1 \leq w_i$, then $\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i', \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}$.

Condition 2. For all mechanisms $M \in \mathcal{M}$, agents $i \in [n]$, reported types $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, and pairs of types $\boldsymbol{\theta}_i, \boldsymbol{\theta}_i' \in [0,1]^m$, if $\|\boldsymbol{\theta}_i - \boldsymbol{\theta}_i'\|_1 \leq w_i$, then $\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i', \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}$.

We claim that Inequality (8) holds so long as Conditions 1 and 2 hold. To see why, suppose they do both hold, and fix an arbitrary agent $i \in [n]$, mechanism $M \in \mathcal{M}$, and pair of types $\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m$. Consider the average amount agent i with type $\boldsymbol{\theta}_i$ can improve his utility by misreporting his type as $\hat{\boldsymbol{\theta}}_i$:

$$\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right). \tag{9}$$

By definition of the grid $\mathcal{G}_{w_i}$, we know there are points $\boldsymbol{p}, \hat{\boldsymbol{p}} \in \mathcal{G}_{w_i}$ such that $\|\boldsymbol{\theta}_i - \boldsymbol{p}\|_1 \leq w_i$ and $\left\|\hat{\boldsymbol{\theta}}_i - \hat{\boldsymbol{p}}\right\|_1 \leq w_i$. Based on Conditions 1 and 2, we show that if we "snap" $\boldsymbol{\theta}_i$ to $\boldsymbol{p}$ and $\hat{\boldsymbol{\theta}}_i$ to $\hat{\boldsymbol{p}}$, we will not increase Equation (9) by more than an additive factor of $O\left(L_i w_i + k_i/N\right)$. Since $\boldsymbol{p}$ and $\hat{\boldsymbol{p}}$ are elements of $\mathcal{G}_{w_i}$, Inequality (8) holds so long as Conditions 1 and 2 hold. Since Conditions 1 and 2 hold with probability $1-\delta$, the lemma statement holds. □

In Section 3.2, we prove that under mild assumptions, for a wide range of mechanisms, Conditions 1 and 2 in Lemma 3.4 hold with $L_i = 1$, $w_i = O\left(1/\sqrt{N}\right)$, and $k_i = \tilde{O}\left(\sqrt{N}\right)$, ignoring problem-specific multiplicands. Thus, we find that $4L_i w_i + 8k_i/N$ quickly converges to 0 as N grows.

Theorem 3.2 and Lemma 3.4 immediately imply this section's main theorem. At a high level, it states that any agent's average utility gain when restricted to types on the grid is a close estimation of the true incentive compatibility approximation factor, so long as his utility function applied to the samples demonstrates dispersion. The full proof is in Appendix C.

Theorem 3.5. *Suppose that for each agent $i \in [n]$, there exist $L_i, k_i, w_i \in \mathbb{R}$ such that with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for each mechanism $M \in \mathcal{M}$ and agent $i \in [n]$, the following conditions hold:*

1. *For any $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

2. *For any $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, the functions $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

Then with probability $1-2\delta$, for every mechanism $M \in \mathcal{M}$ and every agent $i \in [n]$,

$$\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \Theta_i}\left\{\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right]\right\}$$
$$\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}}\left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon,$$

where $\epsilon = 4L_iw_i + \frac{8k_i}{N} + 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ *and* $d_i = \mathrm{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$.

Proof sketch. Theorem 3.2 guarantees that with probability at most δ, the extent to which any agent i can improve his utility by misreporting his type, averaged over all profiles in $\mathcal{S}_{-i}$, does not approximate the true incentive compatibility approximation factor, as summarized below:

Bad event 1. For some mechanism $M \in \mathcal{M}$ and agent $i \in [n]$,

$$\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \Theta_i} \left\{\mathbb{E}_{\boldsymbol{\theta}_{-i}\sim\mathcal{D}_{-i}}\left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right]\right\}$$
$$> \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon_{\mathcal{M}}(N, \delta), \tag{10}$$

where $\epsilon_{\mathcal{M}}(N,\delta) = 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ and $d_i = \mathrm{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$.

By Lemma 3.4, we also know that with probability at most δ, we cannot approximate Equation (10) by discretizing the agent's type space, as summarized by the following bad event:

Bad event 2. For some mechanism $M \in \mathcal{M}$ and agent $i \in [n]$,

$$\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\}$$
$$> \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + 4L_iw_i + \frac{8k_i}{N}. \tag{11}$$

With probability $1-2\delta$, neither bad event occurs, so Equation (11) approximates Equation (10). □

In Appendix C, we show that if, given a set of samples $\mathcal{S}_{-i}$, one can measure the dispersion parameters $L_i, k_i, w_i \in \mathbb{R}$ such that Conditions 1 and 2 in Theorem 3.5 hold, then the theorem's inequality holds. In Section 4, we demonstrate this technique via experiments.

We conclude with one final comparison of the greedy approach in Section 3.1.1 and the uniform grid approach in this section. In Section 3.1.1, we use the functional form of the utility functions in order to bound the cover size, as quantified by pseudo-dimension. When we use the approach based on dispersion, we do not use these functional forms to the fullest extent possible; we only use simple facts about the functions' discontinuities and Lipschitzness.

3.2 Dispersion and pseudo-dimension guarantees

We now provide dispersion and pseudo-dimension guarantees for a variety of mechanism classes. The theorems in this section allow us to instantiate the bounds from the previous section and thus understand how well our empirical incentive compatibility approximation factor matches the true approximation factor.

Given an agent $i \in [n]$, a true type $\boldsymbol{\theta}_i \in [0,1]^m$, and a set of samples $\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)} \sim \mathcal{D}_{-i}$, we often find that the discontinuities of each function $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(j)}\right)$ are highly dependent on the vector $\boldsymbol{\theta}_{-i}^{(j)}$. For example, under the first-price single-item auction, the discontinuities of the

(a) The function $u_{1,M}\left(1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$.

(b) The function $u_{1,M}\left(\frac{1}{2}, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$.

Figure 2: Agent 1's utility function under a first-price single-item auction when the other agents' values and bids are defined by $\boldsymbol{\theta}_{-1}^{(j)} \in [0,1]^{n-1}$. The utility is $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right) = \mathbf{1}_{\left\{\hat{\theta}_1 > \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty\right\}}\left(\theta_1 - \hat{\theta}_1\right)$.

function $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(j)}\right)$ occur when $\hat{\theta}_i = \left\|\boldsymbol{\theta}_{-i}^{(j)}\right\|_\infty$. Since each vector $\boldsymbol{\theta}_{-i}^{(j)}$ is a random draw from the distribution $\mathcal{D}_{-i}$, these functions will not be dispersed if these random draws are highly-concentrated. For this reason, we focus on type distributions with *κ-bounded density functions*. A density function $\phi : \mathbb{R} \to \mathbb{R}$ is κ-bounded if $\max\{\phi(x)\} \leq \kappa$. For example, the density function of any normal distribution with standard deviation $\sigma \in \mathbb{R}$ is $\frac{1}{\sigma\sqrt{2\pi}}$-bounded.

The challenge we face in this section is that it is not enough to show that for some $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are dispersed. Rather, we must prove that for all type vectors, the dispersion property holds. This facet of our analysis is notably different from prior work by Balcan et al. [2018a]: in their applications, it is enough to show that with high probability, a single, finite sequence of functions is dispersed. In contrast, we show that under mild assumptions, with high probability, each function sequence from an infinite family is dispersed.

3.2.1 First-price single-item auction

To develop intuition, we begin by analyzing the first-price auction for a single item. We then analyze the first-price combinatorial auction in Section 3.2.2. Under this auction, each agent $i \in [n]$ has a value $\theta_i \in [0,1]$ for the item and submits a bid $\hat{\theta}_i \in [0,1]$. The agent with the highest bid wins the item and pays his bid. Therefore, agent i's utility function is $u_{i,M}\left(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}\right) = \mathbf{1}_{\left\{\hat{\theta}_i > \left\|\hat{\boldsymbol{\theta}}_{-i}\right\|_\infty\right\}}\left(\theta_i - \hat{\theta}_i\right)$.

To prove our dispersion guarantees, we begin with the following helpful lemma.

Lemma 3.6. *For any agent $i \in [n]$, any set $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\}$, and any type $\theta_i \in [0,1]$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and their discontinuities fall in the set $\left\{\left\|\boldsymbol{\theta}_{-i}^{(j)}\right\|_\infty\right\}_{j \in [N]}$.*

Proof. Suppose $i = 1$ and choose an arbitrary sample $j \in [N]$. For any value $\theta_1 \in [0,1]$ and bid $\hat{\theta}_1 \in [0,1]$, we know that $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right) = \mathbf{1}_{\left\{\hat{\theta}_1 > \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty\right\}}\left(\theta_1 - \hat{\theta}_1\right)$, as illustrated in Figure 2. Therefore, no matter the value of θ_1, the function $u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is piecewise 1-Lipschitz with a discontinuity at $\left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$. $\square$

We now use Lemma 3.6 to prove our first dispersion guarantee. The full proof is in Appendix D.

Theorem 3.7. *Suppose each agent's type has a κ-bounded density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$ and types $\theta_i \in [0,1]$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and $\left(O\left(1/\left(\kappa\sqrt{N}\right)\right), \tilde{O}\left(n\sqrt{N}\right)\right)$-dispersed.*

Proof sketch. Based on Lemma 3.6, in order to prove the theorem, it is enough to show that with probability $1-\delta/n$, any interval of width $O\left(1/\left(\kappa\sqrt{N}\right)\right)$ contains at most $\tilde{O}\left(n\sqrt{N}\right)$ points from the set $\left\{\left\|\boldsymbol{\theta}_{-i}^{(j)}\right\|_\infty\right\}_{j\in[N]}$. We do so by relying on the κ-bounded assumption, which allows us to analyze the concentration of the discontinuities within any interval. □

Theorem 3.8. *For all agents $i \in [n]$, reported types $\hat{\theta}_i \in [0,1]$, and type profiles $\boldsymbol{\theta}_{-i} \in [0,1]^{n-1}$, the function $u_{i,M}\left(\cdot, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz.*

Proof. Since the reported types $\left(\hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ are fixed, the allocation is fixed. Thus, $u_{i,M}\left(\cdot, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ is either a constant function if $\hat{\theta}_i \leq \|\boldsymbol{\theta}_{-i}\|_\infty$ or a linear function if $\hat{\theta}_i > \|\boldsymbol{\theta}_{-i}\|_\infty$. □

Next, we prove the following pseudo-dimension bound. The full proof is in Appendix D.

Theorem 3.9. *For any agent $i \in [n]$, the pseudo-dimension of the class $\mathcal{F}_{i,M}$ is 2.*

Proof sketch. First, we prove $\operatorname{Pdim}(\mathcal{F}_{i,M}) \leq 2$. For a contradiction, suppose there exists a set $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \boldsymbol{\theta}_{-i}^{(2)}, \boldsymbol{\theta}_{-i}^{(3)}\right\}$ that is shattered by $\mathcal{F}_{i,M}$. Without loss of generality, assume $\left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty < \left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty < \left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty$. Since $\mathcal{S}_{-i}$ is shatterable, there exist $z^{(1)}, z^{(2)}, z^{(3)} \in \mathbb{R}$ witnessing the shattering.

We split the proof into two cases: $z^{(3)} > 0$ or $z^{(3)} \leq 0$. In this sketch, we analyze the former case. Since $\mathcal{S}_{-i}$ is shatterable, there is a value $\theta_i \in [0,1]$ and bid $\hat{\theta}_i \in [0,1]$ such that $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) \geq z^{(1)}$, $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(2)}\right) < z^{(2)}$, and $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(3)}\right) \geq z^{(3)}$. Since $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(3)}\right) \geq z^{(3)} > 0$ and $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(3)}\right) = 0$ if $\left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty \geq \hat{\theta}_i$, it must be that $\hat{\theta}_i > \left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty$. Therefore, $\theta_i - \hat{\theta}_i \geq z^{(1)}$ and $\theta_i - \hat{\theta}_i < z^{(2)}$, so $z^{(1)} < z^{(2)}$. Similarly, there must also be $\theta_i' \in [0,1]$ and $\hat{\theta}_i' \in [0,1]$ such that $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(1)}\right) < z^{(1)}$, $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(2)}\right) \geq z^{(2)}$, and $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(3)}\right) \geq z^{(3)}$. By a similar argument, this means $z^{(1)} > z^{(2)}$, which is a contradiction. We prove that in the case where $z^{(3)} \leq 0$, we also arrive at a contradiction. Therefore, $\operatorname{Pdim}(\mathcal{F}_{i,M}) \leq 2$. Furthermore, we exhibit a set of size 2 that is shattered by $\mathcal{F}_{i,M}$, which means that $\operatorname{Pdim}(\mathcal{F}_{i,M}) \geq 2$. Therefore, the theorem statement holds. □

3.2.2 First-price combinatorial auction

Under this auction, there are ℓ items for sale and each agent's type $\boldsymbol{\theta}_i \in [0,1]^{2^\ell}$ indicates his value for each bundle $b \subseteq [\ell]$. We denote his value and bid for bundle b as $\theta_i(b)$ and $\hat{\theta}_i(b)$, respectively. The allocation $(b_1^*, \dots, b_n^*)$ is the solution to the *winner determination problem*:

$$\begin{array}{ll} \text{maximize} & \sum_{i=1}^n \hat{\theta}_i(b_i) \\ \text{subject to} & b_i \cap b_{i'} = \emptyset \quad \forall i, i' \in [n], i \neq i'. \end{array}$$

Each agent $i \in [n]$ pays $\hat{\theta}_i(b_i^*)$.

We begin with dispersion guarantees. The full proof of the following theorem is in Appendix D.

Theorem 3.10. *Suppose that for each pair of agents $i, i' \in [n]$ and each pair of bundles $b, b' \subseteq [\ell]$, the values $\theta_i(b)$ and $\theta_{i'}(b')$ have a κ-bounded joint density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$ and types $\boldsymbol{\theta}_i \in [0,1]^{2^\ell}$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and $\left(O\left(1/\left(\kappa\sqrt{N}\right)\right), \tilde{O}\left((n+1)^{2\ell}\sqrt{N\ell}\right)\right)$-dispersed.*

Proof sketch. Consider agent 1. Fix an arbitrary sample $j \in [N]$ and pair of allocations $(b_1, \dots, b_n)$ and $(b_1', \dots, b_n')$. We know that $(b_1', \dots, b_n')$ will *not* be the allocation so long as $\hat{\boldsymbol{\theta}}_1 \in [0,1]^{2^\ell}$ is chosen such that $\hat{\theta}_1(b_1) + \sum_{i=2}^n \theta_i^{(j)}(b_i) > \hat{\theta}_1(b_1') + \sum_{i=2}^n \theta_i^{(j)}(b_i')$. This means that across all $\boldsymbol{\theta}_1 \in [0,1]^{2^\ell}$, there is a fixed set $\mathcal{H}_j$ of $\binom{(n+1)^\ell}{2}$ hyperplanes (one per pair of allocations) such that for any connected component C of $[0,1]^{2^\ell} \setminus \mathcal{H}_j$, the allocation given bids $\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is invariant across all $\hat{\boldsymbol{\theta}}_1 \in C$. When the allocation is fixed, agent 1's utility is 1-Lipschitz as a function of $\hat{\boldsymbol{\theta}}_1$.

Next, consider the set $\mathcal{S}_{-1} = \left\{\boldsymbol{\theta}_{-1}^{(1)}, \dots, \boldsymbol{\theta}_{-1}^{(N)}\right\}$ of type profiles and the corresponding sets $\mathcal{H}_1, \dots, \mathcal{H}_N$ of hyperplanes. These hyperplanes can be partitioned into $\binom{(n+1)^\ell}{2}$ buckets consisting of parallel hyperplanes with offsets independently drawn from κ-bounded distributions. Within each bucket, the offsets are (w', k')-dispersed with high probability with $w' = O\left(1/\left(\kappa\sqrt{N}\right)\right)$ and $k' = \tilde{O}\left(\sqrt{N\ell}\right)$. Since the hyperplanes within each bucket are parallel and since their offsets are dispersed, for any ball $\mathcal{B}$ of radius w' in $[0,1]^{2^\ell}$, at most k' hyperplanes from each set intersect $\mathcal{B}$. The theorem statement holds by a union bound over the n agents and $\binom{(n+1)^\ell}{2}$ buckets. □

Theorem 3.11. *For all agents $i \in [n]$, reported types $\hat{\boldsymbol{\theta}}_i \in [0,1]^{2^\ell}$, and type profiles $\boldsymbol{\theta}_{-i} \in [0,1]^{(n-1)2^\ell}$, the function $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz.*

Proof. So long as all bids are fixed, the allocation is fixed, so $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz. □

Next, we prove the following pseudo-dimension bound.

Theorem 3.12. *For any agent $i \in [n]$, the pseudo-dimension of the class $\mathcal{F}_{i,M}$ is $O\left(\ell 2^\ell \log n\right)$.*

Proof. As we saw in the proof of Theorem 3.10, for any type profile $\boldsymbol{\theta}_{-i} \in [0,1]^{(n-1)2^\ell}$, there is a set $\mathcal{H}$ of $(n+1)^{2\ell}$ hyperplanes such that for any connected component C of $[0,1]^{2^\ell} \setminus \mathcal{H}$, the auction's allocation given bids $\left(\hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right)$ is invariant across all $\hat{\boldsymbol{\theta}}_i \in C$. So long as the allocation is fixed, $u_{i,M}(\cdot, \cdot, \boldsymbol{\theta}_{-i})$ is a linear function of $\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$. Therefore, the pseudo-dimension bound follows from Theorem B.1 in Appendix B, which relates the class's pseudo-dimension to the number of hyperplanes splitting the type space into regions where the utility function is linear. □

3.2.3 Generalized second-price auction

A *generalized second-price auction* allocates m advertising slots to a set of $n > m$ agents. Each slot s has a probability $\alpha_{s,i}$ of being clicked if agent i's advertisement is in that slot. We assume $\alpha_{s,i}$ is fixed and known by the mechanism designer. The mechanism designer assigns a weight

$\omega_i \in (0,1]$ per agent i. Each agent has a value $\theta_i \in [0,1]$ for a click and submits a bid $\hat{\theta}_i \in [0,1]$. The mechanism allocates the first slot to the agent with the highest weighted bid $\omega_i\hat{\theta}_i$, the second slot to the agent with the second highest weighted bid, and so on. Let $\pi(s)$ be the agent allocated slot s. If slot s is clicked on, agent $\pi(s)$ pays the lowest amount that would have given him slot s, which is $\omega_{\pi(s+1)}\hat{\theta}_{\pi(s+1)}/\omega_{\pi(s)}$. Agent $\pi(s)$'s expected utility is thus $u_{\pi(s),M}\left(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}\right) = \alpha_{s,\pi(s)}\left(\theta_{\pi(s)} - \omega_{\pi(s+1)}\hat{\theta}_{\pi(s+1)}/\omega_{\pi(s)}\right)$.

For $r \in \mathbb{Z}_{\geq 1}$, let $\mathcal{M}_r$ be the set of auctions defined by agent weights from the set $\{1/r, 2/r, \ldots, 1\}$. We begin by proving dispersion guarantees. The full proof is in Appendix D.

Theorem 3.13. *Suppose each agent's type has a κ-bounded density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \ldots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$, types $\theta_i \in [0,1]$, and mechanisms $M \in \mathcal{M}_r$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \ldots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 0-Lipschitz and $\left(O\left(1/\left(r\kappa\sqrt{N}\right)\right), \tilde{O}\left(\sqrt{n^3N}\right)\right)$-dispersed.*

Proof sketch. Consider agent 1. Choose an arbitrary sample $j \in [N]$ and mechanism $M \in \mathcal{M}_r$ with weights $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)$. Let $\omega_{i_1}\theta_{i_1}^{(j)} \leq \cdots \leq \omega_{i_{n-1}}\theta_{i_{n-1}}^{(j)}$ be the weighted types of all agents except agent 1. Consider the n intervals delineated by these $n-1$ weighted values and suppose we vary $\hat{\theta}_1$ in such a way that $\omega_1\hat{\theta}_1$ remains within one interval. The allocation will be invariant, so agent 1's utility will be a constant function of $\hat{\theta}_1$. Therefore, no matter the value of $\theta_1 \in [0,1]$, the functions $u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(1)}\right), \ldots, u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(N)}\right)$ are piecewise constant with discontinuities in the set $\mathcal{B}_{1,\boldsymbol{\omega}} = \left\{\omega_i\theta_i^{(j)}/\omega_1 : i \in \{2, \ldots, n\}, j \in [N]\right\}$. We use the κ-bounded assumption to show that with probability $1-\delta$, for all $\boldsymbol{\omega}$ and agents $i \in [n]$, at most $\tilde{O}\left(n^{3/2}\sqrt{N}\right)$ of the values in each set $\mathcal{B}_{i,\boldsymbol{\omega}}$ fall within any interval of length $O\left(1/\left(r\kappa\sqrt{N}\right)\right)$. □

Theorem 3.14. *For all agents $i \in [n]$, all reported types $\hat{\theta}_i \in [0,1]$, all type profiles $\boldsymbol{\theta}_{-i} \in [0,1]^{n-1}$, and all generalized second-price auctions M, the function $u_{i,M}\left(\cdot, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz.*

Proof. Since the reported types $\left(\hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ are fixed, the allocation is fixed. Let $\pi(s)$ be the agent who is allocated slot s. The function $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ is either a constant function of θ_i if agent i is not allocated a slot or a linear function of θ_i with a slope of $\alpha_{\pi^{-1}(i),i}$ otherwise. □

We now provide the following pseudo-dimension guarantee. The full proof is in Appendix D.

Theorem 3.15. *For any agent $i \in [n]$ and $r \in \mathbb{Z}_{\geq 1}$,* $\operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}_r}\right) = O\left(n \log n\right)$.

Proof sketch. Suppose $i = 1$. Fix $\boldsymbol{\theta}_{-1} = (\theta_2, \ldots, \theta_n) \in [0,1]^{n-1}$. We denote agent 1's utility when the agents' true types are $(\theta_1, \boldsymbol{\theta}_{-1})$ and reported types are $\left(\hat{\theta}_1, \boldsymbol{\theta}_{-1}\right)$ as the function $u_{\boldsymbol{\theta}_{-1}}\left(\theta_1, \hat{\theta}_1, \omega_1, \ldots, \omega_n\right)$, which maps $\mathbb{R}^{n+2}$ to $[-1,1]$. We show that we can split $\mathbb{R}^{n+2}$ into regions where the allocation is fixed as we vary $\left(\theta_1, \hat{\theta}_1, \omega_1, \ldots, \omega_n\right)$ over any one region. When the allocation is fixed, $u_{\boldsymbol{\theta}_{-1}}$ is a linear function of $(\theta_1, \omega_2/\omega_1, \ldots, \omega_n/\omega_1)$. We prove that this partition of $\mathbb{R}^{n+2}$ is delineated by a small number of polynomials of degree at most 2 over these $n+2$ variables. This fact allows us to bound the number of regions making up the partition. Since the utility function is simple within each region, we use our bound on the number of regions to prove the theorem. □

3.2.4 Discriminatory auction

Under the discriminatory auction, there are m identical units of a single item for sale. For each agent $i \in [n]$, his type $\boldsymbol{\theta}_i \in [0,1]^m$ indicates how much he is willing to pay for each additional unit. Thus, $\theta_i[1]$ is the amount he is willing to pay for one unit, $\theta_i[1]+\theta_i[2]$ is the amount he is willing to pay for two units, and so on. We assume that $\theta_i[1] \geq \theta_i[2] \geq \cdots \geq \theta_i[m]$. The auctioneer collects nm bids $\hat{\theta}_i[\mu]$ for $i \in [n]$ and $\mu \in [m]$. If exactly m_i of agent i's bids are among the m highest of all nm bids, then agent i is awarded m_i units and pays $\sum_{\mu=1}^{m_i} \hat{\theta}_i[\mu]$.

We begin with dispersion guarantees. The full proof of the following theorem is in Appendix D.

Theorem 3.16. *Suppose that each agent's value for each marginal unit has a κ-bounded density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for all agents $i \in [n]$ and types $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and $\left(O\left(1/\left(\kappa\sqrt{N}\right)\right), \tilde{O}\left(nm^2\sqrt{N}\right)\right)$-dispersed.*

Proof sketch. Consider agent 1 and suppose we sort $\left\{\theta_i^{(j)}[\mu] : i \in \{2, \dots, n\}, j \in [N], \mu \in [m]\right\}$. So long as agent 1's bid falls between these sorted bids, the allocations will be fixed across all $j \in [N]$, and thus each $u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is 1-Lipschitz. Across all $\boldsymbol{\theta}_1 \in [0,1]^m$, the partition $\mathcal{P}_j$ splitting $u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ into Lipschitz portions is delineated by the set of hyperplanes

$$\left\{\theta_i^{(j)}[\mu] - \hat{\theta}_1[\mu'] = 0 : i \in \{2, \dots, n\}, \mu, \mu' \in [m]\right\}.$$

We partition these hyperplanes into $m^2(n-1)$ buckets consisting of parallel hyperplanes with offsets independently drawn from κ-bounded distributions. The remainder of the proof is similar to that of Theorem 3.10. □

Theorem 3.17. *For all agents $i \in [n]$, reported types $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, and type profiles $\boldsymbol{\theta}_{-i} \in [0,1]^{(n-1)m}$, the function $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz.*

Proof. So long as all bids are fixed, the allocation is fixed, so $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz. □

Next, we prove the following pseudo-dimension bound.

Theorem 3.18. *For any agent $i \in [n]$, the pseudo-dimension of the class $\mathcal{F}_{i,M}$ is $O(m \log(nm))$.*

Proof. As we saw in the proof of Theorem 3.16, for any $\boldsymbol{\theta}_{-i} \in [0,1]^{(n-1)m}$, there is a set $\mathcal{H}$ of $O\left(m^2 n\right)$ hyperplanes such that for any connected component C of $[0,1]^m \setminus \mathcal{H}$, the allocation is fixed across all $\hat{\boldsymbol{\theta}}_i \in C$. So long as the allocation is fixed, $u_{i,M}(\cdot, \cdot, \boldsymbol{\theta}_{-i})$ is a linear function of $\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$. Therefore, the pseudo-dimension bound follows directly from Theorem B.1 in Appendix B. □

3.2.5 Uniform-price auction

Under this auction, the allocation rule is the same as in the discriminatory auction (Section 3.2.4). However, all m units are sold at a "market-clearing" price, meaning the total amount demanded equals the total amount supplied. See the formal definition in Appendix D.1. We obtain the same bounds as we do in Section 3.2.4, so we state the theorems in Appendix D.1.

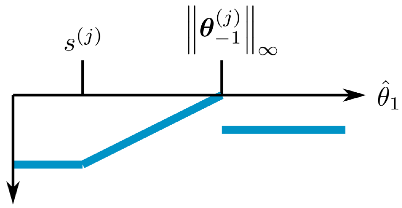


Figure 3: Graph of $u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ for a spiteful agent under a second-price auction with $\alpha_1 = \theta_1 = \frac{1}{2}$, $\|\boldsymbol{\theta}_{-1}\|_\infty = \frac{3}{4}$, and $s^{(j)} = \frac{1}{4}$.

3.2.6 Second-price auction with spiteful agents

We conclude by studying *spiteful agents* [Brandt et al., 2007, Morgan et al., 2003, Sharma and Sandholm, 2010, Tang and Sandholm, 2012], where each bidder's utility not only increases when his surplus increases, but also decreases when the other bidders' surpluses increase. Formally, given a *spite parameter* $\alpha_i \in [0,1]$, agent i's utility under the second-price auction M is $u_{i,M}\left(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}\right) = \alpha_i \mathbf{1}_{\left\{\hat{\theta}_i > \left\|\hat{\boldsymbol{\theta}}_{-i}\right\|_\infty\right\}}\left(\theta_i - \left\|\hat{\boldsymbol{\theta}}_{-i}\right\|_\infty\right) - (1-\alpha_i)\sum_{i' \neq i} \mathbf{1}_{\left\{\hat{\theta}_{i'} > \left\|\hat{\boldsymbol{\theta}}_{-i'}\right\|_\infty\right\}}\left(\theta_{i'} - \left\|\hat{\boldsymbol{\theta}}_{-i'}\right\|_\infty\right)$. The closer α_i is to zero, the more spiteful bidder i is.

We begin with the following lemma, which follows from the form of $u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ (see Figure 3). The proof is in Appendix D.

Lemma 3.19. *For any agent $i \in [n]$, any $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\}$, and any type $\theta_i \in [0,1]$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and their discontinuities fall in the set $\left\{\left\|\boldsymbol{\theta}_{-i}^{(j)}\right\|_\infty\right\}_{j \in [N]}$.*

Lemma 3.19 implies the following dispersion guarantee.

Theorem 3.20. *Suppose each agent's type has a κ-bounded density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$ and types $\theta_i \in [0,1]$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and $\left(O\left(1/\left(\kappa\sqrt{N}\right)\right), \tilde{O}\left(n\sqrt{N}\right)\right)$-dispersed.*

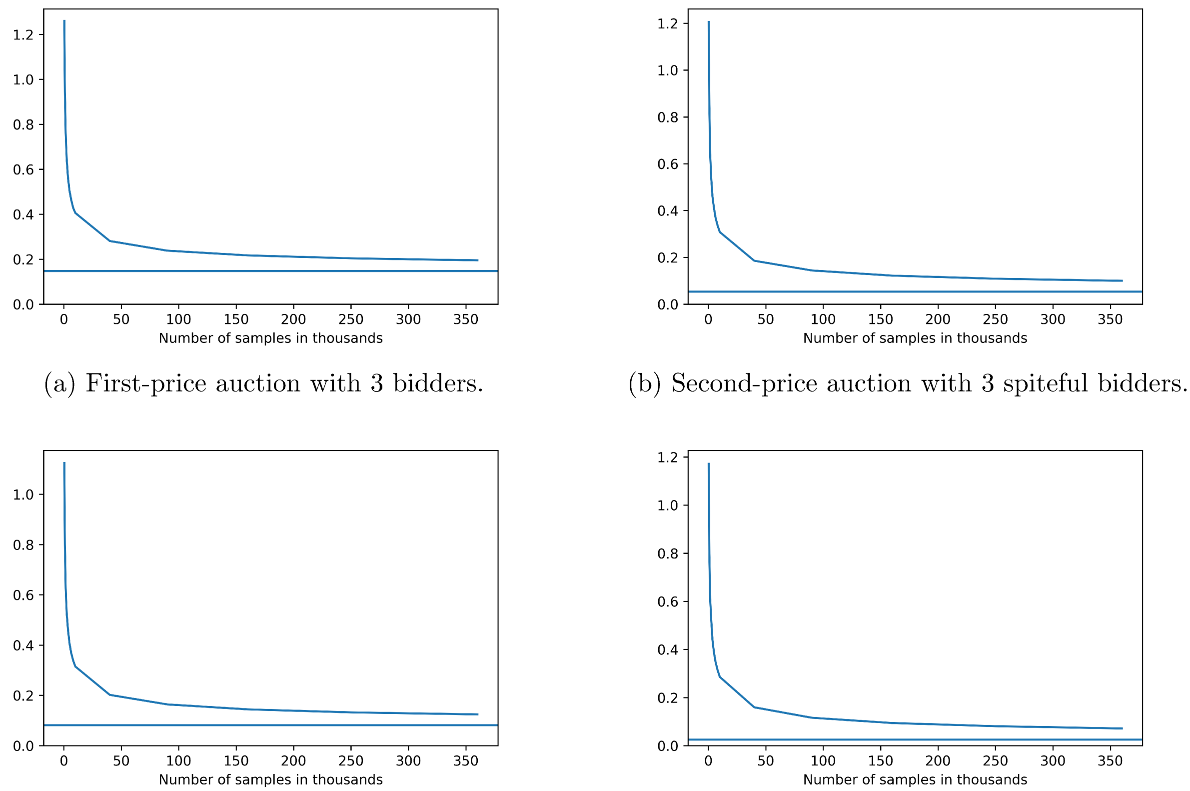
Proof. This follows from Lemma 3.19 exactly as Theorem 3.7 follows from Lemma 3.6. □

Theorem 3.21. *For all agents $i \in [n]$, reported types $\hat{\theta}_i \in [0,1]$, and type profiles $\boldsymbol{\theta}_{-i} \in [0,1]^{n-1}$, the function $u_{i,M}\left(\cdot, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz.*

Proof. Since the reported types $\left(\hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$ are fixed, the allocation is fixed. If $\hat{\theta}_i > \|\boldsymbol{\theta}_{-i}\|_\infty$, then $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) = \alpha_i\left(\theta_i - \|\boldsymbol{\theta}_{-i}\|_\infty\right)$, which is an α_i-Lipschitz function of θ_i. Otherwise, it is a constant function of θ_i. □

We conclude with the following pseudo-dimension bound.

Theorem 3.22. *For any agent $i \in [n]$, the pseudo-dimension of the class $\mathcal{F}_{i,M}$ is $O(1)$.*



(a) First-price auction with 3 bidders.

(b) Second-price auction with 3 spiteful bidders.

(c) First-price auction with 5 bidders.

(d) Second-price auction with 5 spiteful bidders.

Figure 4: Plots illustrating our estimates converging to the true incentive compatibility approximation factors.

Proof. For any vector $\boldsymbol{\theta}_{-i} \in [0,1]^{(n-1)}$, let s be the second-highest component. Letting $\mathcal{H} = \{\hat{\theta}_i = \|\boldsymbol{\theta}_{-i}\|_\infty, \hat{\theta}_i = s\}$ we know that for any interval I of $[0,1] \setminus \mathcal{H}$, the allocation is fixed across all $\hat{\theta}_i \in C$. So long as the allocation is fixed, $u_{i,M}(\cdot,\cdot,\boldsymbol{\theta}_{-i})$ is a linear function of $\left(\theta_i, \hat{\theta}_i\right)$. Therefore, the pseudo-dimension bound follows directly from Theorem B.1 in Appendix B. $\square$

4 Experiments

We now present several experiments illustrating the strength of our diserpersion-based estimation technique. We choose mechanisms M and distributions $\mathcal{D}$ where we can evaluate the expression

$$\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}} \left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right) \right] \right\}. \tag{12}$$

This expression can be challenging or even impossible to evaluate even with knowledge of $\mathcal{D}$.

We analyze the first-price single-item auction with $n \in \{3,5\}$ and the second-price auction with $n \in \{3,5\}$, $\alpha_1 = 1/2$, and $\alpha_2 = \alpha_3 = 0$. In Appendix E, we also show plots for the case where $n = 10$. When there are fewer bidders, each bidder can generally gain more utility by bidding strategically, so Equation (12) will be larger. In all settings, the values are drawn i.i.d. from Uniform(0,1). Under the the first-price auction with $n = 3$, Equation (12) equals 0.15. Under the

the first-price auction with $n = 5$, Equation (12) equals 0.08. Under the the second-price auction with $n = 3$, Equation (12) equals 0.05. Finally, under the the second-price auction with $n = 5$, Equation (12) equals 0.03. These values correspond to the horizontal lines in Figure 4. In all four cases, we range the sample size N up to 360,000 and follow the following procedure. For each bidder $i \in [n]$, we draw $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$. In both cases, we know from Lemmas 3.6 and 3.19 that the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are 1-Lipschitz and their discontinuities fall within the set $T_{-i} = \left\{\left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty, \dots, \left\|\boldsymbol{\theta}_{-i}^{(N)}\right\|_\infty\right\}$. We set $k_i = \lfloor 3\sqrt{N}/2 \rfloor$ and we define w_i to be the largest w such that at most k_i points from T_{-i} fall within any interval of width $2w$. In other words, $w_i = 2 \cdot \sup_{w \in (0,1]} \{w : |[a, a+w] \cap T_{-i}| \leq k_i, \forall a \in [0,1]\}$. Figure 4 displays the bound from Theorem C.5 (a variation on Theorem 3.5) with $\delta = 0.01$. In all settings, we observe that our estimate quickly converges to the true incentive compatibility approximation factor.

5 Conclusion

In this paper, we provided techniques for estimating how far a mechanism is from *ex-interim* and *ex-ante* incentive compatible. We introduced an empirical variant of approximate incentive compatibility which intuitively measures the maximum utility an agent can gain by misreporting his type, on average over the samples. We bounded the difference between our empirical incentive compatibility estimate and the true incentive compatibility approximation factor. To do so, we relied on a subtle mixture of tools from learning theory, including dispersion and pseudo-dimension. We thus derived strong guarantees for many important manipulable mechanisms, including the first-price auction, generalized second-price auction, discriminatory auction, uniform price auction, and second-price auction under spiteful agents.

As in prior research on auction design via machine learning, we assumed that we could receive independent samples from the type distribution. One direction for future work would be to relax this assumption in the spirit of incentive-aware learning [Epasto et al., 2018]. Another would be to relax the assumption that each agent's type is independent from the others' in the *ex-interim* setting.

Acknowledgments

We thank Zhe Feng, Noah Golowich, Harikrishna Narasimhan, and David Parkes for a valuable discussion on approximate incentive compatibility in the context of deep learning, a motivating factor in our study of this general setting.

This material is based on work supported by the National Science Foundation under grants IIS-1618714, IIS-1718457, IIS-1617590, IIS-1901403, CCF-1733556, CCF-1422910, CCF-1535967, the ARO under award W911NF-17-1-0082, an Amazon Research Award, a Microsoft Research Faculty Fellowship, and a Google Research Award.

References

Noga Alon, Moshe Babaioff, Yannai A Gonczarowski, Yishay Mansour, Shay Moran, and Amir Yehudayoff. Submultiplicative Glivenko-Cantelli and uniform convergence of revenues. *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2017.

Martin Anthony and Peter L Bartlett. *Neural network learning: Theoretical foundations*. Cambridge University Press, 2009.

Aaron Archer, Christos Papadimitriou, Kunal Talwar, and Éva Tardos. An approximate truthful mechanism for combinatorial auctions with single parameter agents. *Internet Mathematics*, 1(2): 129–150, 2004.

Eduardo M Azevedo and Eric Budish. Strategy-proofness in the large. *The Review of Economic Studies*, 86(1):81–116, 2018.

Moshe Babaioff, Yannai A Gonczarowski, and Noam Nisan. The menu-size complexity of revenue approximation. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2017.

Maria-Florina Balcan, Avrim Blum, Jason Hartline, and Yishay Mansour. Reducing mechanism design to algorithm design via machine learning. *Journal of Computer and System Sciences*, 74: 78–89, December 2008.

Maria-Florina Balcan, Tuomas Sandholm, and Ellen Vitercik. Sample complexity of automated mechanism design. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2016.

Maria-Florina Balcan, Travis Dick, and Ellen Vitercik. Dispersion for data-driven algorithm design, online learning, and private optimization. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*, 2018a.

Maria-Florina Balcan, Tuomas Sandholm, and Ellen Vitercik. A general theory of sample complexity for multi-item profit maximization. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2018b.

Felix Brandt, Tuomas Sandholm, and Yoav Shoham. Spiteful bidding in sealed-bid auctions. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI)*, Hyderabad, India, 2007. Early version in GTDT-05.

Sébastien Bubeck, Nikhil R Devanur, Zhiyi Huang, and Rad Niazadeh. Online auctions and multi-scale online learning. *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2017.

R. C. Buck. Partition of space. *Amer. Math. Monthly*, 50:541–544, 1943. ISSN 0002-9890.

Yang Cai and Constantinos Daskalakis. Learning multi-item auctions with (or without) samples. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*, 2017.

Yang Cai, Nikhil R. Devanur, and S. Matthew Weinberg. A duality based unified approach to Bayesian mechanism design. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2016.

Shuchi Chawla, Jason Hartline, and Denis Nekipelov. Mechanism design for data science. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2014.

Shuchi Chawla, Jason Hartline, and Denis Nekipelov. A/B testing of auctions. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2016.

Shuchi Chawla, Jason D. Hartline, and Denis Nekipelov. Mechanism redesign. *arXiv preprint arXiv:1708.04699*, 2017.

Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2014.

Wolfram Conen and Tuomas Sandholm. Preference elicitation in combinatorial auctions: Extended abstract. In *Proceedings of the ACM Conference on Electronic Commerce (ACM-EC)*, pages 256–259, Tampa, FL, 2001. More detailed description of algorithmic aspects in IJCAI-01 Workshop on Economic Agents, Models, and Mechanisms, pp. 71–80.

Vincent Conitzer and Tuomas Sandholm. Complexity of mechanism design. In *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 103–110, Edmonton, Canada, 2002.

Vincent Conitzer and Tuomas Sandholm. Computational criticisms of the revelation principle. In *The Conference on Logic and the Foundations of Game and Decision Theory (LOFT)*, Leipzig, Germany, 2004. Earlier versions: AMEC-03, EC-04.

Vincent Conitzer and Tuomas Sandholm. Incremental mechanism design. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI)*, Hyderabad, India, 2007.

Ofer Dekel, Felix Fischer, and Ariel D Procaccia. Incentive compatible regression learning. *Journal of Computer and System Sciences*, 76(8):759–777, 2010.

Nikhil R Devanur, Zhiyi Huang, and Christos-Alexandros Psomas. The sample complexity of auctions with side information. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2016.

Paul Dütting, Felix Fischer, Pichayut Jirapinyo, John K Lai, Benjamin Lubin, and David C Parkes. Payment rules through discriminant-based classifiers. *ACM Transactions on Economics and Computation (TEAC)*, 3(1):5, 2015.

Paul Dütting, Zhe Feng, Harikrishna Narasimhan, and David C Parkes. Optimal auctions through deep learning. *arXiv preprint arXiv:1706.03459*, 2017.

Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *The American Economic Review*, 97(1):242–259, March 2007. ISSN 0002-8282.

Edith Elkind. Designing and learning optimal finite support auctions. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2007.

Alessandro Epasto, Mohammad Mahdian, Vahab Mirrokni, and Song Zuo. Incentive-aware learning for large markets. 2018.

Zhe Feng, Harikrishna Narasimhan, and David C Parkes. Deep learning for revenue-optimal auctions with budgets. In *Autonomous Agents and Multi-Agent Systems*, 2018.

Kira Goldner and Anna R Karlin. A prior-independent revenue-maximizing auction for multiple additive bidders. In *International Workshop On Internet And Network Economics (WINE)*, 2016.

Noah Golowich, Harikrishna Narasimhan, and David C Parkes. Deep learning for multi-facility location mechanism design. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2018.

Yannai A Gonczarowski and Noam Nisan. Efficient empirical revenue maximization in single-parameter auction environments. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, pages 856–868, 2017.

Yannai A Gonczarowski and S Matthew Weinberg. The sample complexity of up-to-ε multi-dimensional revenue maximization. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*, 2018.

Matt Harada. The ad exchanges place in a first-price world. *Marketing Technology Insights*, 2018. URL `https://martechseries.com/mts-insights/guest-authors/ad-exchanges-place-first-price-world/`.

Sergiu Hart and Noam Nisan. Approximate revenue maximization with multiple items. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2012.

Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. Making the most of your samples. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2015.

L. Hurwicz. On informationally decentralized systems. In C.B McGuire and R. Radner, editors, *Decision and Organization*. Amsterdam: North Holland, 1972.

Anshul Kothari, David Parkes, and Subhash Suri. Approximately-strategyproof and tractable multi-unit auctions. In *Proceedings of the ACM Conference on Electronic Commerce (ACM-EC)*, pages 166–175, San Diego, CA, 2003.

Vijay Krishna. *Auction Theory*. Academic Press, 2002.

Sébastien Lahaie, Andrés Munoz Medina, Balasubramanian Sivan, and Sergei Vassilvitskii. Testing incentive compatibility in display ad auctions. 2018.

Kate Larson and Tuomas Sandholm. Costly valuation computation in auctions. In *Theoretical Aspects of Rationality and Knowledge (TARK VIII)*, pages 169–182, Siena, Italy, July 2001.

Kate Larson and Tuomas Sandholm. Mechanism design and deliberative agents. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 650–656, Utrecht, The Netherlands, 2005.

Anton Likhodedov and Tuomas Sandholm. Methods for boosting revenue in combinatorial auctions. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 232–237, San Jose, CA, 2004.

Anton Likhodedov and Tuomas Sandholm. Approximating revenue-maximizing combinatorial auctions. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, Pittsburgh, PA, 2005.

Benjamin Lubin and David C Parkes. Approximate strategyproofness. *Current Science*, pages 1021–1032, 2012.

Andres Munoz Medina and Mehryar Mohri. Learning theory and algorithms for revenue optimization in second price auctions with reserve. In *International Conference on Machine Learning (ICML)*, 2014.

Andrés Muñoz Medina and Sergei Vassilvitskii. Revenue optimization with approximate bid predictions. *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2017.

Timo Mennle and Sven Seuken. An axiomatic approach to characterizing and relaxing strategyproofness of one-sided matching mechanisms. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2014.

John Morgan, Kenneth Steiglitz, and George Reis. The spite motive and equilibrium behavior in auctions. *Contributions to Economic Analysis & Policy*, 2(1), 2003.

Jamie Morgenstern and Tim Roughgarden. On the pseudo-dimension of nearly optimal auctions. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2015.

Jamie Morgenstern and Tim Roughgarden. Learning simple auctions. In *Conference on Learning Theory (COLT)*, 2016.

Tri-Dung Nguyen and Tuomas Sandholm. Multi-option descending clock auction. Technical report, 2015. Mimeo.

Abraham Othman and Tuomas Sandholm. Better with byzantine: Manipulation-optimal mechanisms. In *International Symposium on Algorithmic Game Theory (SAGT)*, Cyprus, 2009.

David Parkes. Optimal auction design for agents with hard valuation problems. In *Agent-Mediated Electronic Commerce Workshop at the International Joint Conference on Artificial Intelligence*, Stockholm, Sweden, 1999.

Rachel Parkin. A year in first-price. *Ad Exchanger*, 2018. URL `https://adexchanger.com/the-sell-sider/a-year-in-first-price/`.

Parag A Pathak and Tayfun Sönmez. School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, 103(1):80–106, 2013.

David Pollard. *Convergence of Stochastic Processes*. Springer, 1984.

Michael Rothkopf, Thomas Teisberg, and Edward Kahn. Why are Vickrey auctions rare? *Journal of Political Economy*, 98(1):94–109, 1990.

Tim Roughgarden and Okke Schrijvers. Ironing in the dark. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2016.

Tuomas Sandholm. An implementation of the contract net protocol based on marginal cost calculations. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 256–262, Washington, D.C., July 1993.

Tuomas Sandholm. Issues in computational Vickrey auctions. *International Journal of Electronic Commerce*, 4(3):107–129, 2000. Early version in ICMAS-96.

Tuomas Sandholm. Automated mechanism design: A new application area for search algorithms. In *International Conference on Principles and Practice of Constraint Programming*, pages 19–36, Cork, Ireland, 2003.

Tuomas Sandholm. Very-large-scale generalized combinatorial multi-attribute auctions: Lessons from conducting $60 billion of sourcing. In Zvika Neeman, Alvin Roth, and Nir Vulkan, editors, *Handbook of Market Design*. Oxford University Press, 2013.

Tuomas Sandholm and Craig Boutilier. Preference elicitation in combinatorial auctions. In Peter Cramton, Yoav Shoham, and Richard Steinberg, editors, *Combinatorial Auctions*, pages 233–263. MIT Press, 2006. Chapter 10.

Tuomas Sandholm and Anton Likhodedov. Automated design of revenue-maximizing combinatorial auctions. *Operations Research*, 63(5):1000–1025, 2015. Special issue on Computational Economics. Subsumes and extends over a AAAI-05 paper and a AAAI-04 paper.

Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge University Press, 2014.

Ankit Sharma and Tuomas Sandholm. Asymmetric spite in auctions. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2010.

Vasilis Syrgkanis. A sample complexity measure with applications to learning optimal auctions. *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2017.

Pingzhong Tang and Tuomas Sandholm. Optimal auctions for spiteful bidders. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2012.

Hal R. Varian. Position auctions. *International Journal of Industrial Organization*, pages 1163–1178, 2007.

Andrew Chi-Chih Yao. An n-to-1 bidder reduction for multi-item auctions and its applications. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2014.

A Helpful lemmas

Lemma A.1 (Anthony and Bartlett [2009]). *Suppose $f_1, \ldots, f_p$ are polynomials of degree at most d in $v \leq p$ variables. Then*

$$\left|\left\{\begin{pmatrix}\operatorname{sign}(f_1(\boldsymbol{x}))\\ \vdots \\ \operatorname{sign}(f_p(\boldsymbol{x}))\end{pmatrix} : \boldsymbol{x} \in \mathbb{R}^v\right\}\right| \leq 2\left(\frac{2epd}{v}\right)^v.$$

Lemma A.2 (Shalev-Shwartz and Ben-David [2014]). *Let $\alpha \geq 1$ and $\beta > 0$. If $x < \alpha \log x + \beta$, then $x < 4\alpha \log(2\alpha) + 2\beta$.*

A.1 Dispersion lemmas

We now include a few lemmas about dispersion from work by Balcan et al. [2018a] as well as a few variations we prove ourselves for our specific applications.

Lemma A.3 (Balcan et al. [2018a]). *Let $\mathcal{B} = \{\beta_1, \ldots, \beta_r\} \subset \mathbb{R}$ be a collection of samples where each β_i is drawn from a distribution with a κ-bounded density function. For any $\delta \geq 0$, the following statements hold with probability at least $1 - \delta$:*

1. *If the β_i are independent, then every interval of width w contains at most $k = O(rw\kappa + \sqrt{r\log(1/\delta)})$ samples. In particular, for any $\alpha \geq 1/2$ we can take $w = 1/(\kappa r^{1-\alpha})$ and $k = O(r^\alpha\sqrt{\log(1/\delta)})$.*

2. *If the samples can be partitioned into P buckets $\mathcal{B}_1, \dots, \mathcal{B}_P$ such that each $\mathcal{B}_i$ contains independent samples and $|\mathcal{B}_i| \leq M$, then every interval of width w contains at most $k = O(PMw\kappa + \sqrt{M\log(P/\delta)})$. In particular, for any $\alpha \geq 1/2$ we can take $w = 1/(\kappa M^{1-\alpha})$ and $k = O(PM^\alpha\sqrt{\log(P/\delta)})$.*

Lemma A.4 (Balcan et al. [2018a])**.** *Suppose X is a random variable with κ-bounded density function and suppose $c \neq 0$ is a constant. Then cX has a $\frac{\kappa}{|c|}$-bounded density function.*

Lemma A.5. *Suppose X and Y are random variables taking values in $[0,1]$ and suppose that their joint distribution has a κ-bounded density function. Then the distribution of $X+Z$ has a κ-bounded density function.*

Proof. Let $Z = X+Y$. We will perform change of variables using the function $g(x,y) = (x, x+y)$. Let $g^{-1}(x,z) = h(x,z) = (x, z-x)$. Then

$$J_h(x,z) = \det\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = 1.$$

Therefore, $f_{X,Z}(x,z) = f_{X,Y}(x, z-x)$. This means that $f_Z(z) = \int_0^1 f_{X,Y}(x, z-x)\,dx \leq \kappa$, so the theorem statement holds. □

Lemma A.6. *Suppose X and Y are random variables taking values in $[0,1]$ and suppose that their joint distribution has a κ-bounded density function. Then the distribution of $X-Z$ has a κ-bounded density function.*

Proof. Let $Z = X-Y$. We will perform change of variables using the function $g(x,y) = (x, x-y)$. Let $g^{-1}(x,z) = h(x,z) = (x, x-z)$. Then

$$J_h(x,z) = \det\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = -1.$$

Therefore, $f_{X,Z}(x,z) = f_{X,Y}(x, x-z)$. This means that $f_Z(z) = \int_0^1 f_{X,Z}(x,z)\,dx = \int_0^1 f_{X,Y}(x, x-z)\,dx \leq \kappa$, so the theorem statement holds. □

Definition A.1 (Hyperplane delineation)**.** Let Ψ be a set of hyperplanes in $\mathbb{R}^m$ and let $\mathcal{P}$ be a partition of a set $\mathbb{R}^m$. Let $K_1, \dots, K_q$ be the connected components of $\mathbb{R}^m \setminus \Psi$. Suppose every set in $\mathcal{P}$ is the union of some collection of sets $K_{i_1}, \dots, K_{i_j}$ together with their limit points. Then we say that the set Ψ *delineates* $\mathcal{P}$.

Lemma A.7 (Balcan et al. [2018a])**.** *Let $u_1, \dots, u_N$ be a set of piecewise L-Lipschitz functions mapping $\mathbb{R}^m$ to $\mathbb{R}$, drawn i.i.d. from a distribution $\mathcal{D}$. For each $j \in [N]$, let $\mathcal{P}_j$ be the partition of $[0,1]^m$ such that over any $R \in \mathcal{P}_j$, u_j is L-Lipschitz. Suppose the hyperplane sets $\Psi_1, \dots, \Psi_N$ delineate the partitions $\mathcal{P}_1, \dots, \mathcal{P}_N$. Moreover, suppose the multi-set union of $\Psi_1, \dots, \Psi_N$ can be partitioned into P multi-sets $\mathcal{B}_1, \dots, \mathcal{B}_P$ such that for each multi-set $\mathcal{B}_i$:*

1. *The hyperplanes in $\mathcal{B}_i$ are parallel with probability 1 over the draw of $u_1, \dots, u_N$.*

2. *The offsets of the hyperplanes in $\mathcal{B}_i$ are independently drawn from κ-bounded distributions.*

With probability at least $1-\delta$ over the draw of $u_1, \dots, u_N$, the functions are

$$\left(\frac{1}{2\kappa\sqrt{\max|\mathcal{B}_i|}}, O\left(P\sqrt{\max|\mathcal{B}_i|\ln\frac{P}{\delta}}\right)\right)\text{-dispersed.}$$

B Delineable utility functions

To prove some of our pseudo-dimension guarantees, we use the notion of *delineability* introduced by Balcan et al. [2018b]. They used it in the context of profit-maximization, so here we adapt it to our setting.

Given a mechanism M and agent $i \in [n]$, let $\mathcal{F}_{i,M} = \left\{ u_{i,M,\theta_i,\hat{\theta}_i} \middle| \theta_i, \hat{\theta}_i \in \Theta_i \right\}$, where $u_{i,M,\theta_i,\hat{\theta}_i}$ is the function introduced in Section 3. Moreover, for a fixed type profile $\boldsymbol{\theta}_{-i} \in \times_{j \neq i} \Theta_j$, let $u_{\boldsymbol{\theta}_{-i}} : \Theta_i^2 \to [-1, 1]$ be a function that maps a pair of types $\theta_i, \hat{\theta}_i \in \Theta_i$ to $u_{i,M,\theta_i,\hat{\theta}_i}(\boldsymbol{\theta}_{-i})$.

Definition B.1 ((m,t)-delineable)**.** Given a mechanism M and agent $i \in [n]$, we say the function class $\mathcal{F}_{i,M}$ is (m,t)-*delineable* if:

1. Each agent's type space is a subset of $[0,1]^m$; and
2. For any $\boldsymbol{\theta}_{-i} \in [0,1]^{m(n-1)}$, there is a set $\mathcal{H}$ of t hyperplanes such that for any connected component $\mathcal{C}$ of $[0,1]^{2m} \setminus \mathcal{H}$, the function $u_{\boldsymbol{\theta}_{-i}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$ is linear over $\mathcal{C}$.

The following theorem is similar to Balcan et al.'s main theorem [Balcan et al., 2018b], though we have adapted it to our setting. We include the proof for completeness.

Theorem B.1. *If $\mathcal{F}_{i,M}$ is (m,t)-delineable, the pseudo dimension of $\mathcal{F}_{i,M}$ is $O(m \log(mt))$.*

Proof. Suppose $\operatorname{Pdim}(\mathcal{F}_{i,M}) = N$. By definition, there exists a set $\mathcal{S} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\}$ that is shattered by $\mathcal{F}_{i,M}$. Let $z^{(1)}, \dots, z^{(N)} \in \mathbb{R}$ be the points that witness this shattering. Again, by definition, we know that for any $T \subseteq [N]$, there exists a pair of types $\boldsymbol{\theta}_i^{(T)}, \hat{\boldsymbol{\theta}}_i^{(T)} \in [0,1]^m$ such that $u_{i,M,\boldsymbol{\theta}_i^{(T)},\hat{\boldsymbol{\theta}}_i^{(T)}}\left(\boldsymbol{\theta}_{-i}^{(j)}\right) \geq z^{(j)}$ if and only if $j \in T$. Let $\Theta^* = \left\{\left(\boldsymbol{\theta}_i^{(T)}, \hat{\boldsymbol{\theta}}_i^{(T)}\right) : T \subseteq [N]\right\}$. To show that the pseudo-dimension N of $\mathcal{F}_{i,M}$ is $O(m \log(mt))$, we will show that $|\Theta^*| = 2^N < 4m^2 t^{2m} N^{4m}$, which means that $N = O(m \log(mt))$.

To this end, for $\boldsymbol{\theta}_{-i}^{(j)} \in \mathcal{S}$, let $\mathcal{H}^{(j)}$ be the set of t hyperplanes such that for any connected component $\mathcal{C}$ of $[0,1]^{2m} \setminus \mathcal{H}^{(j)}$, $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$ is linear over $\mathcal{C}$. We now consider the *overlay* of all N partitions $[0,1]^{2m} \setminus \mathcal{H}^{(1)}, \dots, [0,1]^{2m} \setminus \mathcal{H}^{(N)}$. Formally, this overlay is made up of the sets $\mathcal{C}_1, \dots, \mathcal{C}_\tau$ which are the connected components of $[0,1]^{2m} \setminus \left(\bigcup_{j=1}^N \mathcal{H}^{(j)}\right)$. For each set $\mathcal{C}_\ell$ and each $j \in [N]$, $\mathcal{C}_\ell$ is completely contained in a single connected component of $[0,1]^{2m} \setminus \mathcal{H}^{(j)}$, which means that $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$ is linear over $\mathcal{C}_\ell$. As we know from work by Buck [1943], since $\left|\mathcal{H}^{(j)}\right| \leq t$ for all $j \in [N]$, $\tau < 2m(Nt)^{2m}$.

Now, consider a single connected component $\mathcal{C}_\ell$ of $[0,1]^{2m} \setminus \left(\bigcup_{j=1}^N \mathcal{H}^{(j)}\right)$. For any sample $\boldsymbol{\theta}_{-i}^{(j)} \in \mathcal{S}$, we know that $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$ is linear over $\mathcal{C}_\ell$. Let $\boldsymbol{a}_\ell^{(j)} \in \mathbb{R}^{2m}$ and $b_\ell^{(j)} \in \mathbb{R}$ be the weight vector and offset such that $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) = \boldsymbol{a}_\ell^{(j)} \cdot \left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) + b_\ell^{(j)}$ for all $\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) \in \mathcal{C}_\ell$. We know that there is a hyperplane $\boldsymbol{a}_\ell^{(j)} \cdot \left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) + b_\ell^{(j)} = z^{(j)}$ where on one side of the hyperplane, $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) \leq z^{(j)}$ and on the other side, $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) > z^{(j)}$. Let $\mathcal{H}_{\mathcal{C}_\ell}$ be all N hyperplanes for all N samples $\left(\mathcal{H}_{\mathcal{C}_\ell} = \left\{\boldsymbol{a}_\ell^{(j)} \cdot \left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) + b_\ell^{(j)} = z^{(j)} : j \in [N]\right\}\right)$. Notice that in any connected component $\mathcal{C}$ of $\mathcal{C}_\ell \setminus \mathcal{H}_{\mathcal{C}_\ell}$, for all $j \in [N]$, $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$ is either greater than $z^{(j)}$ or less than $z^{(j)}$

(but not both) for all $\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) \in \mathcal{C}$. Thus, at most one vector $\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right) \in \Theta^*$ can come from $\mathcal{C}$. In total, the number of connected components of $\mathcal{C}_\ell \setminus \mathcal{H}_{\mathcal{C}_\ell}$ is smaller than $2mN^{2m}$. The same holds for every partition $\mathcal{C}_\ell$. Thus, the total number of regions where for all $j \in [N]$, $u_{\boldsymbol{\theta}_{-i}^{(j)}}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i\right)$ is either greater than $z^{(j)}$ or less than $z^{(j)}$ (but not both) is smaller than $2mN^{2m} \cdot 2m(Nt)^{2m}$. Therefore, we may bound $|\Theta^*| = 2^N < 2mN^{2m} \cdot 2m(Nt)^{2m}$. By Lemma A.2, we have that $N = O(m \log(mt))$. $\square$

C Proofs about approximate ex-interim incentive compatibility (Section 3)

Theorem 3.2. *With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \ldots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for all mechanisms $M \in \mathcal{M}$ and agents $i \in [n]$,*

$$\begin{aligned}
&\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right]\right\} \\
\leq &\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon_{\mathcal{M}}(N, \delta),
\end{aligned}$$

where $\epsilon_{\mathcal{M}}(N,\delta) = 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ and $d_i = \operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$.

Proof. Fix an arbitrary agent $i \in [n]$. By Theorem 3.1, we know that with probability at least $1-\delta/n$ over the draw of $\mathcal{S}_{-i}$, for all mechanisms $M \in \mathcal{M}$ and all $\theta_i, \hat{\theta}_i \in \Theta_i$,

$$\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \leq \sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + \sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$$

and

$$\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \leq \sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + \sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}.$$

Therefore,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \\
= \;&\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \\
&+ \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \\
&+ \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \\
\leq \;&\max_{\theta_i, \hat{\theta}_i \in \Theta_i} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon_{\mathcal{M}}(N, \delta).
\end{aligned}$$

Since the above inequality holds for all $i \in [n]$ with probability $1-\delta/n$, a union bound implies that with probability $1-\delta$, for all $i \in [n]$ and all mechanism $M \in \mathcal{M}$,

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_{-i}\sim\mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}\right)\right]\\
&\leq \max_{\theta_i,\hat{\theta}_i\in\Theta_i}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}+\epsilon_{\mathcal{M}}(N,\delta),
\end{aligned}
$$

where $\epsilon_{\mathcal{M}}(N,\delta)=2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}}+2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ and $d_i=\operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$. $\square$

Theorem 3.3. *Given a set* $\mathcal{S}_{-i}=\left\{\boldsymbol{\theta}_{-i}^{(1)},\ldots,\boldsymbol{\theta}_{-i}^{(N)}\right\}$*, a mechanism* $M\in\mathcal{M}$*, and accuracy parameter* $\epsilon>0$*, let* $\mathcal{G}\left(\mathcal{S}_{-i},M,\epsilon\right)$ *be the cover returned by Algorithm 1. With probability* $1-\delta$ *over the draw of the* n *sets* $\mathcal{S}_{-i}\sim\mathcal{D}_{-i}^N$*, for every mechanism* $M\in\mathcal{M}$ *and every agent* $i\in[n]$*,*

$$
\begin{aligned}
&\max_{\theta_i,\hat{\theta}_i\in\Theta_i}\left\{\mathbb{E}_{\boldsymbol{\theta}_{-i}\sim\mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}\right)\right]\right\}\\
&\leq \max_{\left(\theta_i,\hat{\theta}_i\right)\in\mathcal{G}\left(\mathcal{S}_{-i},M,\epsilon\right)}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}+\epsilon+\tilde{O}\left(\sqrt{\frac{d_i}{N}}\right), \qquad (5)
\end{aligned}
$$

where $d_i=\operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$*. Moreover, with probability 1,* $\left|\mathcal{G}\left(\mathcal{S}_{-i},M,\epsilon\right)\right|\leq\left(8eN/\left(\epsilon d_i\right)\right)^{2d_i}$*.*

Proof. Let θ_i^* and $\hat{\theta}_i^*$ be the types in Θ_i that maximize

$$
\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right).
$$

By definition of the set $\mathcal{G}\left(\mathcal{S}_{-i},M,\epsilon\right)$, we know there exists a pair $\left(\theta_i',\hat{\theta}_i'\right)\in\mathcal{G}\left(\mathcal{S}_{-i},M,\epsilon\right)$ such that

$$
\begin{aligned}
&\max_{\theta_i,\hat{\theta}_i\in\Theta_i}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
&=\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i^*,\hat{\theta}_i^*,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i^*,\theta_i^*,\boldsymbol{\theta}_{-i}^{(j)}\right)\\
&\leq\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i',\hat{\theta}_i',\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i',\theta_i',\boldsymbol{\theta}_{-i}^{(j)}\right)+\epsilon\\
&\leq \max_{\theta_i,\hat{\theta}_i\in\mathcal{G}\left(\mathcal{S}_{-i},M,\epsilon\right)}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}+\epsilon.
\end{aligned}
$$

Therefore, by this inequality and Theorem 3.2, we know that with probability at least $1-\delta$,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_{-i}\sim\mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}\right)\right]\\
&-\max_{\left(\theta_i,\hat{\theta}_i\right)\in\mathcal{G}(\mathcal{S}_{-i},M,\epsilon)}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
=&\mathbb{E}_{\boldsymbol{\theta}_{-i}\sim\mathcal{D}_{-i}}\left[u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}\right)\right]\\
&-\max_{\theta_i,\hat{\theta}_i\in\Theta_i}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
&+\max_{\theta_i,\hat{\theta}_i\in\Theta_i}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
&-\max_{\left(\theta_i,\hat{\theta}_i\right)\in\mathcal{G}(\mathcal{S}_{-i},M,\epsilon)}\left\{\frac{1}{N}\sum_{j=1}^{N}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(j)}\right)-u_{i,M}\left(\theta_i,\theta_i,\boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
&\leq\epsilon+\tilde{O}\left(\sqrt{\frac{d_i}{N}}\right).
\end{aligned}$$

The bound on $|\mathcal{G}(\mathcal{S}_{-i},M,\epsilon)|$ follows from Lemma C.1. ☐

Lemma C.1. *For any $\mathcal{G}$ returned by Algorithm 1, $|\mathcal{G}|\leq(8eN/(\epsilon d_i))^{2d_i}$, where $d_i=\mathrm{Pdim}(\mathcal{F}_{i,\mathcal{M}})$.*

Proof. Let $\mathcal{N}_1(\epsilon,U)$ be the ϵ-covering number of U and let $\mathcal{P}_1(\epsilon,U)$ be the ϵ-packing number of U. In other words, $\mathcal{N}_1(\epsilon,U)$ is the size of the smallest set $V'\subseteq\mathbb{R}^N$ such that for all $\boldsymbol{v}\in U$, there is a vector $\boldsymbol{v}'\in V'$ such that $\|\boldsymbol{v}-\boldsymbol{v}'\|_1\leq\epsilon$, and $\mathcal{P}_1(\epsilon,U)$ is the size of the largest set $P\subseteq U$ such that for all $\boldsymbol{v},\boldsymbol{v}'\in P$, if $\boldsymbol{v}\neq\boldsymbol{v}'$, then $\|\boldsymbol{v}-\boldsymbol{v}'\|_1\geq\epsilon$. As Anthony and Bartlett [2009] prove, $\mathcal{P}(\epsilon,U)\leq\mathcal{N}(\epsilon/2,U)$.

By construction, for every pair of vectors $\boldsymbol{v}$ and $\boldsymbol{v}'$ in the set V defined in Algorithm 1, we know that $\|\boldsymbol{v}-\boldsymbol{v}'\|_1\geq\epsilon$. Therefore, $|V|\leq\mathcal{P}(\epsilon,U)\leq\mathcal{N}(\epsilon/2,U)$. In the following claim, we prove that $\mathcal{N}(\epsilon/2,U)\leq(8eN/(\epsilon d_i))^{2d_i}$, which proves the theorem.

Claim C.2. $\mathcal{N}(\epsilon/2,U)\leq(8eN/(\epsilon d_i))^{2d_i}$.

Proof. Let U' be the set of vectors

$$U'=\left\{\frac{1}{N}\begin{pmatrix}u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(1)}\right)\\ \vdots\\ u_{i,M}\left(\theta_i,\hat{\theta}_i,\boldsymbol{\theta}_{-i}^{(N)}\right)\end{pmatrix}:\theta_i,\hat{\theta}_i\in\Theta_i\right\}.$$

Note that $U\subseteq\{\boldsymbol{v}-\boldsymbol{v}':\boldsymbol{v},\boldsymbol{v}'\in U'\}$. We claim that $\mathcal{N}(\epsilon/2,U)\leq\mathcal{N}(\epsilon/4,U')^2\leq(8eN/(\epsilon d_i))^{2d_i}$, where the second inequality follows from a theorem by Anthony and Bartlett [2009]. To see why the first inequality holds, suppose C' is an $\frac{\epsilon}{4}$-cover of U' with minimal size. Let $C=\{\boldsymbol{c}-\boldsymbol{c}':\boldsymbol{c},\boldsymbol{c}'\in C'\}$. Then for all $\boldsymbol{v}\in U$, we know that $\boldsymbol{v}=\boldsymbol{v}'-\boldsymbol{v}''$ for some $\boldsymbol{v}',\boldsymbol{v}''\in U'$. Moreover, we know that there are vectors $\boldsymbol{c}',\boldsymbol{c}''\in C'$ such that $\|\boldsymbol{v}'-\boldsymbol{c}'\|_1\leq\epsilon/4$ and $\|\boldsymbol{v}''-\boldsymbol{c}''\|_1\leq\epsilon/4$. Therefore, $\|\boldsymbol{v}-(\boldsymbol{c}'-\boldsymbol{c}'')\|_1=\|(\boldsymbol{v}'-\boldsymbol{v}'')-(\boldsymbol{c}'-\boldsymbol{c}'')\|_1\leq\epsilon/2$. Since $\boldsymbol{c}'-\boldsymbol{c}''\in C$, we have that C is a cover of U. Therefore, $\mathcal{N}(\epsilon/2,U)\leq|C|\leq|C'|^2=\mathcal{N}(\epsilon/4,U')^2\leq(8eN/(\epsilon d_i))^{2d_i}$. ☐

$\square$

Lemma C.3. *Let $\mathcal{M}$ be a class of mechanisms. For each $i \in [n]$, let $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\}$ be a set of type profiles for all agents except agent i and let $L_i, k_i, w_i \in \mathbb{R}$ be real values such that for each mechanism $M \in \mathcal{M}$, the following conditions hold:*

1. *For any $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*
2. *For any $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, the functions $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

Then for all $M \in \mathcal{M}$ and $i \in [n]$,

$$\begin{aligned}
&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} \\
&\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + 4L_i w_i + \frac{8k_i}{N}.
\end{aligned} \tag{13}$$

Proof. By definition of dispersion, we know that the following conditions hold:

Condition 1. For all mechanisms $M \in \mathcal{M}$, all agents $i \in [n]$, all types $\boldsymbol{\theta}_i \in [0,1]^m$, and all reported types $\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_i' \in [0,1]^m$, if $\left\|\hat{\boldsymbol{\theta}}_i - \hat{\boldsymbol{\theta}}_i'\right\|_2 \leq w_i$, then

$$\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i', \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}. \tag{14}$$

Condition 2. For all mechanisms $M \in \mathcal{M}$, all agents $i \in [n]$, all reported types $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, and all types $\boldsymbol{\theta}_i, \boldsymbol{\theta}_i' \in [0,1]^m$, if $\|\boldsymbol{\theta}_i - \boldsymbol{\theta}_i'\|_2 \leq w_i$, then

$$\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i', \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}. \tag{15}$$

We claim that the inequality from the lemma statement (Equation (13)) holds so long as Conditions 1 and 2 hold.

Claim C.4. *If Conditions 1 and 2 hold, then for all $M \in \mathcal{M}$ and $i \in [n]$,*

$$\begin{aligned}
&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} \\
&\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + 4L_i w_i + \frac{8k_i}{N}.
\end{aligned}$$

Proof of Claim C.4. Fix an agent $i \in [n]$ and let $\boldsymbol{\theta}_i$ and $\hat{\boldsymbol{\theta}}_i$ be two fixed, arbitrary vectors in $[0,1]^m$. By definition of $\mathcal{G}_w$, there must be a point $\boldsymbol{p} \in \mathcal{G}_w$ such that $\|\boldsymbol{\theta}_i - \boldsymbol{p}\|_1 \leq w_i$. By Equation (14),

$$\left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \leq L_i w_i + \frac{2k_i}{N}$$

and by Equation (15),

$$\left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{p}, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \leq L_i w_i + \frac{2k_i}{N}.$$

Similarly, there must be a point $\hat{\boldsymbol{p}} \in \mathcal{G}_w$ such that $\left\|\hat{\boldsymbol{p}} - \hat{\boldsymbol{\theta}}_i\right\|_1 \leq w_i$. By Equation (14),

$$\left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \leq L_i w_i + \frac{2k_i}{N}$$

and by Equation (15),

$$\left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{p}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \leq L_i w_i + \frac{2k_i}{N}.$$

Therefore,

$$\begin{aligned}
& \left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \\
\leq & \left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \\
+ & \left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{p}, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \\
+ & \left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{p}, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{p}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \\
+ & \left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{p}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \\
+ & \left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| \\
\leq & \left| \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{p}, \boldsymbol{p}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{p}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right| + 4L_i w_i + \frac{8k_i}{N}.
\end{aligned}$$

Since $\boldsymbol{p}, \hat{\boldsymbol{p}} \in \mathcal{G}_{w_i}$, the claim holds. □

The lemma thus follows from Claim C.4. $\square$

Theorem C.5. *Let $\mathcal{M}$ be a mechanism class. Given n sets of samples $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^{N}$, for each $i \in [n]$ let $L_i, k_i, w_i \in \mathbb{R}$ be defined such that for each $M \in \mathcal{M}$, the following conditions hold:*

1. *For any $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*
2. *For any $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, the functions $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

Then with probability $1-\delta$, for every $M \in \mathcal{M}$, every agent $i \in [n]$, and every pair $\theta_i, \hat{\theta}_i \in \Theta_i$,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] \\
&\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + 4L_i w_i + \frac{8k_i}{N} + \epsilon_{\mathcal{M}}(N, \delta)
\end{aligned}$$

where $\epsilon_{\mathcal{M}}(N,\delta) = 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ and $d_i = \operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$.

Proof. From Theorem 3.2, we know that with probability $1-\delta$, for every mechanism $M \in \mathcal{M}$ and every agent $i \in [n]$,

$$\begin{aligned}
&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m}\left\{\mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}}\left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right]\right\} \\
&\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon_{\mathcal{M}}(N, \delta).
\end{aligned}$$

Moreover, from Lemma C.3, we know that the following inequality holds:

$$\begin{aligned}
&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} \\
&\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + 4L_i w_i + \frac{8k_i}{N}.
\end{aligned}$$

Therefore, with probability $1-\delta$, for every mechanism $M \in \mathcal{M}$ and every agent $i \in [n]$,

$$\begin{aligned}
&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}} \left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right) \right] \right\} \\
=& \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}} \left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right) \right] \right\} \\
&- \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} \\
&+ \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} \\
&- \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} \\
&+ \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} \\
\leq& \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + 4L_i w_i + \frac{8k_i}{N} + \epsilon_{\mathcal{M}}(N, \delta).
\end{aligned}$$

□

Lemma 3.4. *Suppose that for each agent $i \in [n]$, there exist $L_i, k_i, w_i \in \mathbb{R}$ such that with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for each mechanism $M \in \mathcal{M}$ and agent $i \in [n]$, the following conditions hold:*

1. *For any type $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

2. *For any reported type $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, the functions $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

Then with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} \sim \mathcal{D}_{-i}^N$, for all mechanisms $M \in \mathcal{M}$ and agents $i \in [n]$,

$$\begin{aligned}
&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} \\
\leq& \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{ \frac{1}{N} \sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + 4L_i w_i + \frac{8k_i}{N}. \qquad (8)
\end{aligned}$$

Proof. By definition of dispersion, we know that with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, the following conditions hold:

Condition 1. For all mechanisms $M \in \mathcal{M}$, all agents $i \in [n]$, all types $\boldsymbol{\theta}_i \in [0,1]^m$, and all reported types $\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_i' \in [0,1]^m$, if $\left\|\hat{\boldsymbol{\theta}}_i - \hat{\boldsymbol{\theta}}_i'\right\|_1 \leq w_i$, then

$$\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i', \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}. \tag{16}$$

Condition 2. For all mechanisms $M \in \mathcal{M}$, all agents $i \in [n]$, all reported types $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, and all types $\boldsymbol{\theta}_i, \boldsymbol{\theta}_i' \in [0,1]^m$, if $\|\boldsymbol{\theta}_i - \boldsymbol{\theta}_i'\|_1 \leq w_i$, then

$$\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i', \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}. \tag{17}$$

From Claim C.4, we know that Equation (8) holds so long as Conditions 1 and 2 hold. Since Conditions 1 and 2 hold with probability $1-\delta$, the lemma statement holds. □

Theorem 3.5. *Suppose that for each agent $i \in [n]$, there exist $L_i, k_i, w_i \in \mathbb{R}$ such that with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for each mechanism $M \in \mathcal{M}$ and agent $i \in [n]$, the following conditions hold:*

1. *For any $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*
2. *For any $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, the functions $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed.*

Then with probability $1-2\delta$, for every mechanism $M \in \mathcal{M}$ and every agent $i \in [n]$,

$$\begin{aligned}&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \Theta_i}\left\{\mathbb{E}_{\boldsymbol{\theta}_{-i}\sim\mathcal{D}_{-i}}\left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right]\right\}\\ \leq &\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}}\left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon,\end{aligned}$$

where $\epsilon = 4L_iw_i + \frac{8k_i}{N} + 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ and $d_i = \text{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$.

Proof. Theorem 3.2 guarantees that with probability at most δ, the extent to which any agent i can improve his utility by misreporting his type, averaged over all profiles in $\mathcal{S}_{-i}$, does not approximate the true incentive compatibility approximation factor, as summarized below:

Bad event 1. For some mechanism $M \in \mathcal{M}$ and agent $i \in [n]$,

$$\begin{aligned}&\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \Theta_i}\left\{\mathbb{E}_{\boldsymbol{\theta}_{-i}\sim\mathcal{D}_{-i}}\left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right]\right\}\\ > &\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m}\left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon_{\mathcal{M}}(N,\delta),\end{aligned}$$

where $\epsilon_{\mathcal{M}}(N,\delta) = 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ and $d_i = \text{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$.

By Lemma 3.4, we also know that with probability at most δ, we cannot approximate Equation (10) by discretizing the agent's type space, as summarized by the following bad event:

Bad event 2. For some mechanism $M \in \mathcal{M}$ and agent $i \in [n]$,

$$\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\}$$
$$> \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + 4L_i w_i + \frac{8k_i}{N}.$$

By a union bound, the probability that either Bad Event 1 or Bad Event 2 occurs is at most 2δ. Therefore, the probability that neither occurs is at least $1 - 2\delta$. If neither occurs, then for every mechanism $M \in \mathcal{M}$ and every agent $i \in [n]$,

$$\max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i} \sim \mathcal{D}_{-i}} \left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right) \right] \right\}$$
$$\leq \max_{\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + \epsilon,$$

where $\epsilon = 4L_i w_i + \frac{4k_i}{N} + 2\sqrt{\frac{2d_i}{N} \ln \frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N} \ln \frac{2n}{\delta}}$ and $d_i = \operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}}\right)$. □

D Proofs about dispersion and pseudo-dimension (Section 3.2)

Theorem 3.7. *Suppose each agent's type has a κ-bounded density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \ldots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$ and types $\theta_i \in [0,1]$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \ldots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and $\left(O\left(1/\left(\kappa\sqrt{N}\right)\right), \tilde{O}\left(n\sqrt{N}\right)\right)$-dispersed.*

Proof. Consider an arbitrary bidder, and without loss of generality, suppose that bidder is bidder 1. Next, choose an arbitrary sample $\boldsymbol{\theta}_{-1}^{(j)} = \left(\theta_2^{(j)}, \ldots, \theta_n^{(j)}\right)$. For any value $\theta_1 \in [0,1]$ and bid $\hat{\theta}_1 \in [0,1]$, we know that $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right) = \mathbf{1}_{\left\{\hat{\theta}_1 > \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty\right\}}\left(\theta_1 - \hat{\theta}_1\right)$, as illustrated in Figure 2. Therefore, if $\hat{\theta}_1 \leq \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$, then $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is a constant function of $\hat{\theta}_1$, whereas if $\hat{\theta}_1 > \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$, then $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is a linear function of $\hat{\theta}_1$ with a slope of -1. Therefore, for all $\theta_i \in [0,1]$, the function $u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is piecewise 1-Lipschitz with a discontinuity at $\left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$.

We now prove that with probability $1 - \frac{\delta}{n}$, for all $\theta_1 \in [0,1]$, the functions

$$u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(1)}\right), \ldots, u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(N)}\right)$$

are (w,k)-dispersed. Since for any $\theta_1 \in [0,1]$ the function $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ will only have a discontinuity at a point in the set $\left\{\theta_2^{(j)}, \ldots, \theta_n^{(j)}\right\}$, it is enough to prove that with probability $1 - \frac{\delta}{n}$, at most k points in the set $\mathcal{B} = \bigcup_{j=1}^N \left\{\theta_2^{(j)}, \ldots, \theta_n^{(j)}\right\}$ fall within any interval of width w. The theorem statement then holds by a union bound over all n bidders.

Claim D.1. *With probability $1 - \frac{\delta}{n}$, at most k points in the set $\bigcup_{j=1}^{N} \left\{\theta_2^{(j)}, \dots, \theta_n^{(j)}\right\}$ fall within any interval of width w.*

Proof of Claim D.1. For $i \in \{2, \dots, n\}$, let $\mathcal{B}_i = \left\{\theta_i^{(j)}\right\}_{j \in [N]}$. The claim follows from Lemma A.3 in Appendix A.1. $\square$

$\square$

Theorem 3.9. *For any agent $i \in [n]$, the pseudo-dimension of the class $\mathcal{F}_{i,M}$ is 2.*

Proof. First, we prove that $\mathrm{Pdim}\left(\mathcal{F}_{i,M}\right) \leq 2$.

Claim D.2. *The pseudo-dimension of $\mathcal{F}_{i,M}$ is at most 2.*

Proof of Claim D.2. For a contradiction, suppose there exists a set $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \boldsymbol{\theta}_{-i}^{(2)}, \boldsymbol{\theta}_{-i}^{(3)}\right\}$ that is shattered by $\mathcal{F}_{i,M}$. Without loss of generality, assume that $\left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty < \left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty < \left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty$. Since $\mathcal{S}_{-i}$ is shatterable, there exists three values $z^{(1)}, z^{(2)}, z^{(3)} \in \mathbb{R}$ that witness the shattering of $\mathcal{S}_{-i}$ by $\mathcal{F}_{i,M}$. We split the proof into two cases: one where $z^{(3)} > 0$ and one where $z^{(3)} \leq 0$.

Case 1: $z^{(3)} > 0$. Since $\mathcal{S}_{-i}$ is shatterable, there is a value $\theta_i \in [0,1]$ and bid $\hat{\theta}_i \in [0,1]$ such that $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) \geq z^{(1)}$, $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(2)}\right) < z^{(2)}$, and $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(3)}\right) \geq z^{(3)}$. Recall that for any $\boldsymbol{\theta}_{-i} \in [0,1]^{n-1}$, $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) = \mathbf{1}_{\left\{\hat{\theta}_i > \|\boldsymbol{\theta}_{-i}\|_\infty\right\}}\left(\theta_i - \hat{\theta}_i\right)$. Since $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(3)}\right) \geq z^{(3)} > 0$, it must be that $\hat{\theta}_i > \left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty$. Therefore, $\theta_i - \hat{\theta}_i \geq z^{(1)}$ and $\theta_i - \hat{\theta}_i < z^{(2)}$, which means that $z^{(1)} < z^{(2)}$.

Next, there must also be a value $\theta_i' \in [0,1]$ and bid $\hat{\theta}_i' \in [0,1]$ such that $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(1)}\right) < z^{(1)}$, $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(2)}\right) \geq z^{(2)}$, and $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(3)}\right) \geq z^{(3)}$. Again, since $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(3)}\right) \geq z^{(3)} > 0$, it must be that $\hat{\theta}_i' > \left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty$. Therefore, $\theta_i' - \hat{\theta}_i' < z^{(1)}$ and $\theta_i' - \hat{\theta}_i' \geq z^{(2)}$, which means that $z^{(2)} > z^{(1)}$, which is a contradiction.

Case 2: $z^{(3)} \leq 0$. Since $\mathcal{S}_{-i}$ is shatterable, there is a value $\theta_i \in [0,1]$ and bid $\hat{\theta}_i \in [0,1]$ such that $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) \geq z^{(1)}$, $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(2)}\right) < z^{(2)}$, and $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(3)}\right) < z^{(3)}$. Since $u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(3)}\right) < z^{(3)} \leq 0$, it must be that $\hat{\theta}_i > \left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty$. Therefore, $\theta_i - \hat{\theta}_i \geq z^{(1)}$ and $\theta_i - \hat{\theta}_i < z^{(2)}$, which means that $z^{(1)} < z^{(2)}$.

Next, there must also be a value $\theta_i' \in [0,1]$ and bid $\hat{\theta}_i' \in [0,1]$ such that $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(1)}\right) < z^{(1)}$, $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(2)}\right) \geq z^{(2)}$, and $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(3)}\right) < z^{(3)}$. Again, since $u_{i,M}\left(\theta_i', \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(3)}\right) < z^{(3)} \leq 0$, it must be that $\hat{\theta}_i' > \left\|\boldsymbol{\theta}_{-i}^{(3)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty > \left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty$. Therefore, $\theta_i' - \hat{\theta}_i' < z^{(1)}$ and $\theta_i' - \hat{\theta}_i' \geq z^{(2)}$, which means that $z^{(2)} > z^{(1)}$, which is a contradiction.

Since we arrive at a contradiction in both cases, we conclude that $\mathrm{Pdim}\left(\mathcal{F}_{i,M}\right) \leq 2$. $\square$

We now prove that $\mathrm{Pdim}\left(\mathcal{F}_{i,M}\right) \geq 2$.

Claim D.3. *The pseudo-dimension of $\mathcal{F}_{i,M}$ is at least 2.*

θ_i	$\hat{\theta}_i$	$u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right)$	$u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(2)}\right)$
1	1	0	0
	7/8	1/8	1/8
	3/4	1/4	1/4
	1/2	1/2	0

Table 1: Shattering of the example from Claim D.3.

Proof of Claim D.3. To prove this claim, we exhibit a set of size 2 that is shattered by $\mathcal{F}_{i,M}$. Let $\boldsymbol{\theta}_{-i}^{(1)}$ and $\boldsymbol{\theta}_{-i}^{(2)}$ be two sets of values such that $\left\|\boldsymbol{\theta}_{-i}^{(1)}\right\|_\infty = 1/3$ and $\left\|\boldsymbol{\theta}_{-i}^{(2)}\right\|_\infty = 2/3$. Table 1 illustrates that $z^{(1)} = 3/16$ and $z^{(2)} = 1/16$ witness the shattering of $\left\{\boldsymbol{\theta}_{-i}^{(1)}, \boldsymbol{\theta}_{-i}^{(2)}\right\}$ by $\mathcal{F}_{i,M}$. ☐

Together, Claims D.2 and D.3 prove that $\operatorname{Pdim}(\mathcal{F}_{i,M}) = 2$. ☐

Theorem 3.10. *Suppose that for each pair of agents $i, i' \in [n]$ and each pair of bundles $b, b' \subseteq [\ell]$, the values $\theta_i(b)$ and $\theta_{i'}(b')$ have a κ-bounded joint density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$ and types $\boldsymbol{\theta}_i \in [0,1]^{2^\ell}$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and $\left(O\left(1/\left(\kappa\sqrt{N}\right)\right), \tilde{O}\left((n+1)^{2\ell}\sqrt{N\ell}\right)\right)$-dispersed.*

Proof. Fix an arbitrary agent, and without loss of generality, suppose that agent is agent 1. Next, fix an arbitrary sample $\boldsymbol{\theta}_{-1}^{(j)}$ and pair of allocations $(b_1, \dots, b_n)$ and $(b_1', \dots, b_n')$. We know that $(b_1', \dots, b_n')$ will *not* be the allocation of the first-price combinatorial auction so long as agent 1's reported type $\hat{\boldsymbol{\theta}}_1 \in [0,1]^{2^\ell}$ is chosen such that

$$\hat{\theta}_1(b_1) + \sum_{i=2}^n \theta_i^{(j)}(b_i) > \hat{\theta}_1(b_1') + \sum_{i=2}^n \theta_i^{(j)}(b_i').$$

This means that across all $\boldsymbol{\theta}_1 \in [0,1]^{2^\ell}$, there is fixed a set Ψ_j of $\binom{(n+1)^\ell}{2}$ hyperplanes (one per pair of allocations) such that for any connected component C of $[0,1]^{2^\ell} \setminus \Psi_j$, the allocation of the first-price combinatorial auction is invariant across all $\hat{\boldsymbol{\theta}}_1 \in C$. Namely,

$$\Psi_j = \left\{\hat{\theta}_1(b_1) - \hat{\theta}_1(b_1') = \sum_{i=2}^n \theta_i^{(j)}(b_i') - \theta_i^{(j)}(b_i) : (b_1, \dots, b_n) \text{ and } \left(b_1', \dots, b_n'\right) \text{ are allocations}\right\}.$$

So long as the allocation is fixed, agent 1's utility is 1-Lipschitz in $\hat{\boldsymbol{\theta}}_1$.

For each pair of allocations $\boldsymbol{b} = (b_1, \dots, b_n)$ and $\boldsymbol{b}' = (b_1', \dots, b_n')$, let

$$\mathcal{B}_{\boldsymbol{b},\boldsymbol{b}'} = \left\{\hat{\theta}_1(b_1) - \hat{\theta}_1(b_1') = \sum_{i=2}^n \theta_i^{(j)}(b_i') - \theta_i^{(j)}(b_i) : j \in [N]\right\}.$$

Within each set, the hyperplanes are parallel with probability 1. By Lemmas A.5 and A.6 in Appendix A.1, their offsets are independently drawn from distributions with κ-bounded density functions. Therefore, by Lemma A.7 in Appendix A.1, with probability $1 - \frac{\delta}{n}$, for all $\boldsymbol{\theta}_1 \in [0,1]^{2^\ell}$, the functions $u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(1)}\right), \dots, u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(N)}\right)$ are (w,k)-dispersed with $w = O\left(\frac{1}{\kappa\sqrt{N}}\right)$ and $k = O\left((n+1)^{2\ell}\sqrt{N \log \frac{n(n+1)^{2\ell}}{\delta}}\right)$. The theorem statement holds by a union bound over all n agents. $\square$

Theorem 3.13. *Suppose each agent's type has a κ-bounded density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$, types $\theta_i \in [0,1]$, and mechanisms $M \in \mathcal{M}_r$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 0-Lipschitz and $\left(O\left(1/\left(r\kappa\sqrt{N}\right)\right), \tilde{O}\left(\sqrt{n^3 N}\right)\right)$-dispersed.*

Proof. Let $w = O\left(\frac{1}{r\kappa\sqrt{N}}\right)$ and $k = O\left(n^{3/2}\sqrt{N \log \frac{nr}{\delta}}\right)$. Consider an arbitrary agent, and without loss of generality, suppose that agent is agent 1. Next, choose an arbitrary sample $\boldsymbol{\theta}_{-1}^{(j)} = \left(\theta_2^{(j)}, \dots, \theta_n^{(j)}\right)$ and mechanism $M \in \mathcal{M}_r$ with agent weights $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n) \in \left\{\frac{1}{r}, \frac{2}{r}, \dots, 1\right\}^n$. Let $\omega_{i_1}\theta_{i_1}^{(j)} \leq \cdots \leq \omega_{i_{n-1}}\theta_{i_{n-1}}^{(j)}$ be the weighted types of all agents except agent 1. Consider the n intervals delineated by these $n-1$ weighted values: $I_1 = \left[0, \omega_{i_1}\theta_{i_1}^{(j)}\right), \dots, I_n = \left[\omega_{i_{n-1}}\theta_{i_{n-1}}^{(j)}, 1\right]$ with

$$I_\tau = \left[\omega_{i_{\tau-1}}\theta_{i_{\tau-1}}^{(j)}, \omega_{i_\tau}\theta_{i_\tau}^{(j)}\right).$$

Suppose we vary $\hat{\theta}_1$ in such a way that $\omega_1\hat{\theta}_1$ remains within a single interval I_τ. Said another way, consider varying $\hat{\theta}_1$ over the interval

$$\left[\frac{\omega_{i_{\tau-1}}\theta_{i_{\tau-1}}^{(j)}}{\omega_1}, \frac{\omega_{i_\tau}\theta_{i_\tau}^{(j)}}{\omega_1}\right).$$

The allocation will be constant, which means agent 1's utility will be a constant function of $\hat{\theta}_1$. Therefore, no matter the value of θ_1, $u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(1)}\right), \dots, u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(N)}\right)$ are piecewise constant with discontinuities in the set

$$\mathcal{B}_{1,\boldsymbol{\omega}} = \left\{\frac{\omega_{i'}\theta_{i'}^{(j)}}{\omega_1} : i' \in \{2, \dots, n\}, j \in [N]\right\}.$$

This means that in order to prove the theorem, it is enough to show that with probability $1-\delta$, for all $\boldsymbol{\omega} \in \left\{\frac{1}{r}, \frac{2}{r}, \dots, 1\right\}^n$ and all agents $i \in [n]$, at most k of the values in each set

$$\mathcal{B}_{i,\boldsymbol{\omega}} = \left\{\frac{\omega_{i'}\theta_{i'}^{(j)}}{\omega_i} : i' \in [n] \setminus \{i\}, j \in [N]\right\}$$

fall within any interval of length w. This follows from the following claim and a union bound.

Claim D.4. *For a fixed $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n) \in \left\{\frac{1}{r}, \frac{2}{r}, \dots, 1\right\}^n$ and agent $i \in [n]$, with probability $1 - \frac{\delta}{nr^n}$, at most k points from $\mathcal{B}_{i,\boldsymbol{\omega}}$ fall within any interval of width w.*

Proof of Claim D.4. For each $i' \in [n] \setminus \{i\}$, let

$$\mathcal{B}_{i,i',\boldsymbol{\omega}} = \left\{ \frac{\omega_{i'} \theta_{i'}^{(j)}}{\omega_i} : j \in [N] \right\}.$$

Since $\frac{\omega_{i'}}{\omega_i} \geq \frac{1}{r}$, Lemma A.4 in Appendix A.1 implies that the points in $\mathcal{B}_{i,i',\boldsymbol{\omega}}$ are independent draws from a $r\kappa$-bounded distribution. The claim then follows from Lemma A.3 in Appendix A.1. □

□

Theorem 3.15. *For any agent $i \in [n]$ and $r \in \mathbb{Z}_{\geq 1}$,* $\operatorname{Pdim}\left(\mathcal{F}_{i,\mathcal{M}_r}\right) = O\left(n \log n\right)$.

Proof. Fix an arbitrary agent, and without loss of generality, suppose that agent is agent 1. Fix a vector $\boldsymbol{\theta}_{-1} = (\theta_2, \dots, \theta_n) \in [0,1]^{n-1}$. We denote the utility of agent 1 when the agents' types are $(\theta_1, \boldsymbol{\theta}_{-1})$ and reported types are $\left(\hat{\theta}_1, \boldsymbol{\theta}_{-1}\right)$ as $u_{\boldsymbol{\theta}_{-1}}\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right)$, which maps $\mathbb{R}^{n+2}$ to $[-1,1]$.

We now show that we can split $\mathbb{R}^{n+2}$ into regions R where the allocation of the generalized second-price auction is fixed across all $\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right) \in R$, given the fixed set of reported types $\boldsymbol{\theta}_{-1} = (\theta_2, \dots, \theta_n)$. Let $\bar{\pi}$ be a permutation of the n agents, and let $R_{\bar{\pi}} \subseteq [0,1]^{n+2}$ be the set of all vectors $\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right)$ where the ordering of the elements $\left\{\omega_1 \hat{\theta}_1, \omega_2 \theta_2, \dots, \omega_n \theta_n\right\}$ matches $\bar{\pi}$. Specifically, for all $\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right) \in R_{\bar{\pi}}$, if $i' = \operatorname{argmax}\left\{\omega_1 \hat{\theta}_1, \omega_2 \theta_2, \dots, \omega_n \theta_n\right\}$, then $\bar{\pi}(1) = i'$ and if $i'' = \operatorname{argmin}\left\{\omega_1 \hat{\theta}_1, \omega_2 \theta_2, \dots, \omega_n \theta_n\right\}$, then $\bar{\pi}(n) = i''$. By definition of the generalized second-price auction, this means that the agent who receives the first slot is agent $\bar{\pi}(1)$ and the agent who receives the m^{th} is agent $\bar{\pi}(m)$. Therefore, for all $\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right) \in R_{\bar{\pi}}$, the allocation of the generalized second-price auction given agent weights $(\omega_1, \dots, \omega_n)$ and reported types $\left(\hat{\theta}_1, \boldsymbol{\theta}_{-1}\right)$ is invariant. Next, consider the transformation $\phi : \left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right) \mapsto \left(\theta_1, \frac{\omega_2}{\omega_1}, \dots, \frac{\omega_n}{\omega_1}\right)$. Since the vector $\boldsymbol{\theta}_{-1}$ is fixed, the function $u_{\boldsymbol{\theta}_{-1}}\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right)$ is a fixed linear function of $\phi\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right)$ across all $\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right) \in R_{\bar{\pi}}$.

Next, let $\mathcal{S}_{-1} = \left\{\boldsymbol{\theta}_{-1}^{(1)}, \dots, \boldsymbol{\theta}_{-1}^{(N)}\right\}$ be a set of vectors that is shatterable by the set of functions $\mathcal{F}_{1,\mathcal{M}_r}$, where $\boldsymbol{\theta}_{-1}^{(j)} = \left(\theta_2^{(j)}, \dots, \theta_n^{(j)}\right)$. This means there is a set of values $z^{(1)}, \dots, z^{(N)} \in \mathbb{R}$ that witnesses the shattering of $\mathcal{S}_{-1}$ by $\mathcal{F}_{1,\mathcal{M}_r}$. Therefore, for every $T \subseteq [N]$, there exists a pair of values $\theta_T, \hat{\theta}_T$ and a mechanism $M_T \in \mathcal{M}_r$ such that $u_{1,M_T,\theta_T,\hat{\theta}_T}\left(\boldsymbol{\theta}_{-1}^{(j)}\right) \leq z^{(j)}$ if and only if $j \in T$. Let $(\omega_{1,T}, \dots, \omega_{n,T})$ be the set of agent weights corresponding to each such mechanism M_T. We define the set $\mathcal{R}$ to be the set of 2^N vectors $\mathcal{R} = \left\{\left(\theta_T, \hat{\theta}_T, \omega_{1,T}, \dots, \omega_{n,T}\right) : T \subseteq [N]\right\} \subset \mathbb{R}^{n+2}$.

For each $j \in [N]$, let $\mathcal{H}^{(j)}$ be the set of $\binom{n}{2} + n - 1$ hypersurfaces

$$\mathcal{H}^{(j)} = \left\{\omega_1 \hat{\theta}_1 - \omega_i \theta_i^{(j)} = 0 : i \in \{2, \dots, n\}\right\} \cup \left\{\omega_i \theta_i - \omega_{i'} \theta_{i'}^{(j)} = 0 : i, i' \in \{2, \dots, n\}\right\}.$$

As we saw, as we range $\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right)$ over a single connected component of $\mathbb{R}^{n+2} \setminus \mathcal{H}^{(j)}$, the allocation of the generalized second-price auction with agent weights $(\omega_1, \dots, \omega_n)$ and reported types $\left(\hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is invariant. If we define $\mathcal{H} = \bigcup_{j=1}^N \mathcal{H}^{(j)}$, then as we range $\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right)$

over a single connected component of $\mathbb{R}^{n+2} \setminus \mathcal{H}$, the allocations are fixed across all $j \in [N]$. By Lemma A.1 in Appendix A, the number of connected components is at most

$$2\left(\frac{4eN\left(\binom{n}{2}+n-1\right)}{n+2}\right)^{n+2} = 2\left(2eN(n-1)\right)^{n+2}.$$

Next, fix a connected component C of $\mathbb{R}^{n+2} \setminus \mathcal{H}$. We know that for each $j \in [N]$, there is a constant $\alpha_j \in [0,1]$ and a bidder $i_j \neq 1$ such that $u_{\boldsymbol{\theta}_{-1}^{(j)}}\left(\theta_1, \hat{\theta}_1, \omega_1, \dots, \omega_n\right) = \alpha_j\left(\theta_1 - \frac{\omega_{i_j}\theta_{i_j}}{\omega_1}\right)$ (where $\alpha_j = 0$ if agent i does not receive any slot). Consider the N hypersurfaces $\mathcal{H}_C = \left\{\alpha_j\left(\theta_1 - \frac{\omega_{i_j}t_{i_j}^{(j)}}{\omega_1}\right) = z^{(j)} : j \in [N]\right\}$. We know that at most one vector in $\mathcal{R}$ can come from each connected component of $C \setminus \mathcal{H}_C$, of which there are at most $2\left(\frac{4eN}{n}\right)^n$. In total, this means that $2^N = |\mathcal{R}| \leq 4\left(2eN(n-1)\right)^{n+2}\left(\frac{4eN}{n}\right)^n$, so by Lemma A.2 in Appendix A, $N = O(n\log n)$. $\square$

Theorem 3.16. *Suppose that each agent's value for each marginal unit has a κ-bounded density function. With probability $1-\delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, for all agents $i \in [n]$ and types $\boldsymbol{\theta}_i \in [0,1]^m$, the functions $u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and $\left(O\left(1/\left(\kappa\sqrt{N}\right)\right), \tilde{O}\left(nm^2\sqrt{N}\right)\right)$-dispersed.*

Proof. Fix an arbitrary agent, and without loss of generality, suppose that agent is agent 1. Let $\theta'_1 \leq \theta'_2 \leq \cdots \leq \theta'_{Nm(n-1)}$ be the sorted values of $\left\{\theta_i^{(j)}[\mu] : i \in \{2, \dots, n\}, j \in [N], \mu \in [m]\right\}$, and define $\theta'_0 = 0$ and $\theta'_{Nm(n-1)+1} = 1$. So long as agent 1's bids fall between these sorted bids, the allocation will be fixed across all samples. More formally, for each $\mu \in [m]$, so long as $\theta'_i < \hat{\theta}_1[\mu] < \theta'_{i+1}$ for some $i \in \{0, \dots, Nm(n-1)\}$, the resulting allocation will be fixed across all N samples. Now, for a given sample $\boldsymbol{\theta}_{-1}^{(j)}$, consider a region $R \subseteq [0,1]^m$ where the allocation of the discriminatory auction given bids $\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is invariant across all $\hat{\boldsymbol{\theta}}_1 \in R$. In particular, let m_R be the number of units allocated to agent 1. For a fixed $\boldsymbol{\theta}_1 \in [0,1]^m$, agent 1's utility will be linear in $\hat{\boldsymbol{\theta}}_1$, because across all $\hat{\boldsymbol{\theta}}_1 \in R$, $u_{i,M}\left(\boldsymbol{\theta}_1, \hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_{-1}^{(j)}\right) = \sum_{\mu=1}^{m_R} \theta_1[\mu] - \hat{\theta}_1[\mu]$. Therefore, the functions $u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(1)}\right), \dots, u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(N)}\right)$ are piecewise 1-Lipschitz.

Note that across all $\boldsymbol{\theta}_1 \in [0,1]^m$, the partition $\mathcal{P}_j$ splitting $u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ into Lipschitz portions is invariant. In particular, no matter the value of $\boldsymbol{\theta}_1$, the partition $\mathcal{P}_j$ is delineated by the set of $m^2(n-1)$ hyperplanes $\mathcal{H}_j = \left\{\theta_i^{(j)}[\mu] - \hat{\theta}_1[\mu'] = 0 : i \in \{2, \dots, n\}, \mu, \mu' \in [m]\right\}$. For each $i \in \{2, \dots, n\}$ and pair $\mu, \mu' \in [m]$, let $\mathcal{B}_{i,\mu,\mu'}$ be the set of hyperplanes

$$\mathcal{B}_{i,\mu,\mu'} = \left\{\theta_i^{(j)}[\mu] - \hat{\theta}_1[\mu'] = 0 : j \in [N]\right\}.$$

Within each set, the hyperplanes are parallel with probability 1 and their offsets are independently drawn from κ-bounded distributions. Therefore, by Lemma A.7 in Appendix A.1, with probability $1 - \frac{\delta}{n}$, for all $\boldsymbol{\theta}_1 \in [0,1]^m$, the functions $u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(1)}\right), \dots, u_{1,M}\left(\boldsymbol{\theta}_1, \cdot, \boldsymbol{\theta}_{-1}^{(N)}\right)$ are $\left(O\left(\frac{1}{\kappa\sqrt{N}}\right), \tilde{O}\left(nm^2\sqrt{N}\right)\right)$-dispersed. The theorem holds by a union bound over the agents. $\square$

Lemma 3.19. *For any agent $i \in [n]$, any $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\}$, and any type $\theta_i \in [0,1]$, the functions $u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\theta_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$ are piecewise 1-Lipschitz and their discontinuities fall in the set $\left\{\left\|\boldsymbol{\theta}_{-i}^{(j)}\right\|_\infty\right\}_{j\in[N]}$.*

Proof. Consider an arbitrary bidder, and without loss of generality, suppose that bidder is bidder 1. Next, choose an arbitrary sample $\boldsymbol{\theta}_{-1}^{(j)} = \left(\theta_2^{(j)}, \dots, \theta_n^{(j)}\right)$. Letting $j^* = \operatorname{argmax}\left\{\theta_2^{(j)}, \dots, \theta_n^{(j)}\right\}$ and letting $s^{(j)}$ be the second-largest component of $\boldsymbol{\theta}_{-1}^{(j)}$, we know that either agent 1 will win and pay $\left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$ or agent j^* will win and pay the maximum of $\hat{\theta}_1$ and $s^{(j)}$. Therefore, for any value $\theta_1 \in [0,1]$ and bid $\hat{\theta}_1 \in [0,1]$, we know that

$$u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right) = \begin{cases} (\alpha_1 - 1)\left(\left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty - s^{(j)}\right) & \text{if } \hat{\theta}_1 \le s^{(j)} \\ (\alpha_1 - 1)\left(\left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty - \hat{\theta}_1\right) & \text{if } s^{(j)} < \hat{\theta}_1 \le \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty \\ \alpha_1\left(\theta_1 - \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty\right) & \text{if } \hat{\theta}_1 > \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty. \end{cases} \tag{18}$$

This means that if $\hat{\theta}_1 > \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$, then $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is a constant function of $\hat{\theta}_1$, whereas if $\hat{\theta}_1 \le \left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$, then $u_{1,M}\left(\theta_1, \hat{\theta}_1, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is a continuous function of $\hat{\theta}_1$ with a slope of either 0 (if $\hat{\theta}_1 < s^{(j)}$) or $\alpha_1 - 1$ (if $\hat{\theta}_1 \ge s^{(j)}$). (See Figure 3 in Appendix D.) Since $|\alpha_1 - 1| \le 1$, this means that for all $\theta_i \in [0,1]$, the function $u_{1,M}\left(\theta_1, \cdot, \boldsymbol{\theta}_{-1}^{(j)}\right)$ is piecewise 1-Lipschitz with a discontinuity at $\left\|\boldsymbol{\theta}_{-1}^{(j)}\right\|_\infty$. □

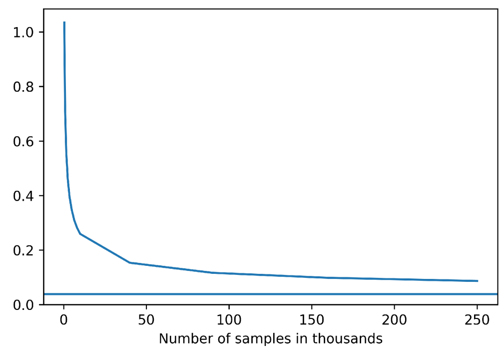
D.1 Uniform-price auctions

Under the uniform-price auction, the allocation rule is the same as in the discriminatory auction (Section 3.2.4). However, all m units are sold at a "market-clearing" price, meaning that the total amount demanded is equal to the total amount supplied. We adopt the convention [Krishna, 2002] that the market-clearing price equals the highest losing bid. Let $\boldsymbol{c}_{-i} \in \mathbb{R}^m$ be the vector $\hat{\boldsymbol{\theta}}_{-i}$ of competing bids facing agent i, sorted in decreasing order and limited to the top m bids. This means that $c_{-i}[1] = \left\|\hat{\boldsymbol{\theta}}_{-i}\right\|_\infty$ is the highest of the other bids, $c_{-i}[2]$ is the second-highest, and so on. Agent i will win exactly one unit if and only if $\hat{\theta}_i[1] > c_{-i}[m]$ and $\hat{\theta}_i[2] < c_{-i}[m-1]$; that is, his bid must be higher than the lowest competing bid but not the second-lowest. Similarly, agent i will win exactly two units if and only if $\hat{\theta}_i[2] > c_{-i}[m-1]$ and $\hat{\theta}_i[3] < c_{-i}[m-2]$. More generally, agent i will win exactly $m_i > 0$ units if and only if $\hat{\theta}_i[m_i] > c_{-i}[m - m_i + 1]$ and $\hat{\theta}_i[m_i + 1] < c_{-i}[m - m_i]$. The market-clearing price, which equals the highest losing bid, is $p = \max\left\{\hat{\theta}_i[m_i + 1], c_{-i}[m - m_i + 1]\right\}$. In a uniform-price auction, agent i pays $m_i p$.

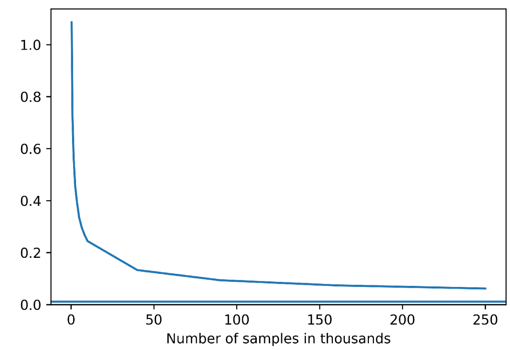
Theorem D.5. *Suppose that each agent's value for each marginal unit has a κ-bounded density function. With probability $1 - \delta$ over the draw of the n sets $\mathcal{S}_{-i} = \left\{\boldsymbol{\theta}_{-i}^{(1)}, \dots, \boldsymbol{\theta}_{-i}^{(N)}\right\} \sim \mathcal{D}_{-i}^N$, we have that for all agents $i \in [n]$ and types $\boldsymbol{\theta}_i \in [0,1]^m$, the functions*

$$u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$$

are piecewise 1-Lipschitz and $\left(O\left(\frac{1}{\kappa\sqrt{N}}\right), O\left(nm^2\sqrt{N\log\frac{n}{\delta}}\right)\right)$-dispersed.



(a) First-price auction with 10 bidders.



(b) Second-price auction with 10 spiteful bidders.

Figure 5: Plots illustrating our estimates converging to the true incentive compatibility approximation factors.

Proof. This theorem follows by the exact same reasoning as Theorem 3.16. □

Theorem D.6. *For all agents $i \in [n]$, all $\hat{\boldsymbol{\theta}}_i \in [0,1]^m$, and all $\boldsymbol{\theta}_{-i} \in [0,1]^{(n-1)m}$, the function $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz.*

Proof. So long as all bids are fixed, the allocation is fixed, so $u_{i,M}\left(\cdot, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right)$ is 1-Lipschitz function as a function of the true type $\boldsymbol{\theta}_i \in [0,1]^m$. □

Next, we prove the following pseudo-dimension bound.

Theorem D.7. *For any agent $i \in [n]$, the pseudo-dimension of the class $\mathcal{F}_{i,M}$ is $O(m \log(nm))$.*

Proof. This theorem follows by the exact same reasoning as Theorem 3.18. □

E More details about experiments

We analyze the first-price single-item auction with $n = 10$ and the second-price auction with $n = 10$, $\alpha_1 = 1/2$, and $\alpha_2 = \alpha_3 = 0$. In both settings, the values are drawn i.i.d. from Uniform(0,1). Under the the first-price auction with $n = 10$, Equation (12) equals 0.04. Under the the second-price auction with $n = 10$, Equation (12) equals 0.01. These values correspond to the horizontal lines in Figure 5. In all four cases, we range the sample size N up to 250,000 and follow the estimation procedure from Section 4.

F Approximate *ex-ante* incentive compatibility

Definition F.1 (*Ex-ante* approximate incentive compatibility (AIC))**.** A mechanism M is *ex-ante* γ-AIC if for all agents i and all types $\hat{\theta}_i \in \Theta_i$,

$$\mathbb{E}_{\boldsymbol{\theta} \sim \mathcal{D}}\left[u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \geq E_{\boldsymbol{\theta} \sim \mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \gamma.$$

In this section, we show how to estimate an *ex-ante* incentive compatibility approximation guarantee using data. Unlike Section 3 we do not assume that the agents' types are independently distributed. For each agent $i \in [n]$, we receive a set $\mathcal{S}_i$ of samples independently drawn from $\mathcal{D}$. For each mechanism $M \in \mathcal{M}$, we show how to use the samples to estimate a value $\hat{\gamma}_M$ such that:

> *With probability* $1-\delta$ *over the draw of the* n *sets of samples* $\mathcal{S}_1, \dots, \mathcal{S}_n$*, for any agent* $i \in [n]$ *and all types* $\hat{\theta}_i \in \Theta_i$*,* $\mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \leq \hat{\gamma}_M$.

We begin with a parallel to Theorem 3.2. For any mechanism M, any agent $i \in [n]$, and any type $\hat{\theta}_i \in \Theta_i$, let $u_{i,M,\hat{\theta}_i} : \times_{i=1}^n \Theta_j \to [-1,1]$ be a function that maps the agents' types $\boldsymbol{\theta}$ to the utility of agent i with type θ_i and reported type $\hat{\theta}_i$ when the other agents report their types truthfully. In other words, $u_{i,M,\hat{\theta}_i}(\boldsymbol{\theta}) = u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)$. Let $\mathcal{F}'_{i,\mathcal{M}}$ be the set of all such functions defined by mechanisms M from the class $\mathcal{M}$ $\left(\mathcal{F}'_{i,\mathcal{M}} = \left\{u_{i,M,\hat{\theta}_i} \,\middle|\, \hat{\theta}_i \in \Theta_i, M \in \mathcal{M}\right\}\right)$.

Theorem F.1. *With probability* $1-\delta$ *over the draw of the* n *sets* $\mathcal{S}_i = \left\{\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}\right\} \sim \mathcal{D}^N$*, for every mechanism* $M \in \mathcal{M}$ *and every agent* $i \in [n]$*,*

$$\begin{aligned}
&\max_{\hat{\theta}_i \in \Theta_i}\left\{\mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right]\right\}\\
&\leq \max_{\hat{\theta}_i \in \Theta_i}\left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon_{\mathcal{M}}(N,\delta),
\end{aligned}$$

where $\epsilon_{\mathcal{M}}(N,\delta) = 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ *and* $d_i = \operatorname{Pdim}\left(\mathcal{F}'_{i,\mathcal{M}}\right)$*.*

Proof. Fix an arbitrary agent $i \in [n]$. By Theorem 3.1, we know that with probability at least $1-\delta/n$ over the draw of $\mathcal{S}_i$, for all mechanisms $M \in \mathcal{M}$ and all $\hat{\theta}_i \in \theta_i$,

$$\mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) \leq \sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + \sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$$

and

$$\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) - \mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \leq \sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + \sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}.$$

Therefore,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right]\\
=\;& \mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right] - \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\\
&+ \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\\
&+ \frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) - \mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right]\\
\leq\;& \max_{\hat{\theta}_i \in \Theta_i}\left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon_{\mathcal{M}}(N,\delta).
\end{aligned}$$

Algorithm 2 Greedy cover construction for *ex-ante* incentive compatibility

Input: Mechanism $M \in \mathcal{M}$, set of samples $\mathcal{S}_i = \left\{\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}\right\}$, accuracy parameter $\epsilon > 0$.

Let U be the set of vectors

$$U \leftarrow \left\{ \frac{1}{N} \begin{pmatrix} u_{i,M}\left(\theta_i^{(1)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) - u_{i,M}\left(\theta_i^{(1)}, \theta_i^{(1)}, \boldsymbol{\theta}_{-i}^{(1)}\right) \\ \vdots \\ u_{i,M}\left(\theta_i^{(N)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(N)}\right) - u_{i,M}\left(\theta_i^{(N)}, \theta_i^{(N)}, \boldsymbol{\theta}_{-i}^{(N)}\right) \end{pmatrix} : \hat{\theta}_i \in \Theta_i \right\}.$$

Let $V \leftarrow \emptyset$ and $\mathcal{G} \leftarrow \emptyset$.

while $U \setminus \left(\bigcup_{\boldsymbol{v} \in V} B_1(\boldsymbol{v}, \epsilon)\right) \neq \emptyset$ **do**

  Select an arbitrary vector $\boldsymbol{v}' \in U \setminus \left(\bigcup_{\boldsymbol{v} \in V} B_1(\boldsymbol{v}, \epsilon)\right)$.

  Let $\hat{\theta}_i \in \Theta_i$ be the type such that

$$\boldsymbol{v}' = \frac{1}{N} \begin{pmatrix} u_{i,M}\left(\theta_i^{(1)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(1)}\right) - u_{i,M}\left(\theta_i^{(1)}, \theta_i^{(1)}, \boldsymbol{\theta}_{-i}^{(1)}\right) \\ \vdots \\ u_{i,M}\left(\theta_i^{(N)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(N)}\right) - u_{i,M}\left(\theta_i^{(N)}, \theta_i^{(N)}, \boldsymbol{\theta}_{-i}^{(N)}\right) \end{pmatrix}.$$

  Add $\boldsymbol{v}'$ to V and $\hat{\theta}_i$ to $\mathcal{G}$.

Output: The cover $\mathcal{G} \subseteq \Theta_i$.

Since the above inequality holds for all $i \in [n]$ with probability $1 - \delta/n$, a union bound implies that with probability $1 - \delta$, for all $i \in [n]$ and all mechanism $M \in \mathcal{M}$,

$$\begin{aligned} &\max_{\hat{\theta}_i \in \Theta_i} \left\{ \mathbb{E}_{\boldsymbol{\theta} \sim \mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \right\} \\ &\leq \max_{\hat{\theta}_i \in \Theta_i} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + \epsilon_{\mathcal{M}}(N, \delta), \end{aligned}$$

where $\epsilon_{\mathcal{M}}(N, \delta) = 2\sqrt{\frac{2d_i}{N} \ln \frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N} \ln \frac{2n}{\delta}}$ and $d_i = \operatorname{Pdim}\left(\mathcal{F}'_{i,\mathcal{M}}\right)$. □

F.1 *Ex-ante* incentive compatibility guarantees via finite covers

The following is a variation on Theorem 3.3.

Theorem F.2. *Given a set of samples $\mathcal{S}_i = \left\{\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}\right\}$, a mechanism $M \in \mathcal{M}$, and accuracy parameter $\epsilon > 0$, let $\mathcal{G}(\mathcal{S}_i, M, \epsilon)$ be the cover returned by Algorithm 2. With probability $1 - \delta$ over the draw of the n sets $\mathcal{S}_i = \left\{\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}\right\} \sim \mathcal{D}^N$, for every mechanism $M \in \mathcal{M}$, every agent $i \in [n]$, and every $\hat{\theta}_i \in \Theta_i$,*

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{\theta} \sim \mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right] \\ &\leq \max_{\hat{\theta}_i \in \mathcal{G}(\mathcal{S}_i, M, \epsilon)} \left\{ \frac{1}{N} \sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) \right\} + \epsilon + \tilde{O}\left(\sqrt{\frac{d_i}{N}}\right), \end{aligned} \tag{19}$$

where $d_i = \operatorname{Pdim}\left(\mathcal{F}'_{i,\mathcal{M}}\right)$. Moreover, with probability 1, $|\mathcal{G}(\mathcal{S}_i, M, \epsilon)| \leq \left(8eN/(\epsilon d_i)\right)^{2d_i}$.

Proof. Let $\hat{\theta}_i^*$ be the type in Θ_i that maximizes

$$\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right).$$

By definition of the set $\mathcal{G}\left(\mathcal{S}_i, M, \epsilon\right)$, we know there exists $\hat{\theta}_i' \in \mathcal{G}\left(\mathcal{S}_i, M, \epsilon\right)$ such that

$$\begin{aligned}
&\max_{\hat{\theta}_i \in \Theta_i}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
=&\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i^*, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\\
\leq&\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i', \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) + \epsilon\\
\leq&\max_{\hat{\theta}_i \in \mathcal{G}(\mathcal{S}_i, M, \epsilon)}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon.
\end{aligned}$$

Therefore, by this inequality and Theorem F.1, we know that with probability at least $1-\delta$,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right]\\
&-\max_{\hat{\theta}_i \in \mathcal{G}(\mathcal{S}_i, M, \epsilon)}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
=&\mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{D}}\left[u_{i,M}\left(\theta_i, \hat{\theta}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\theta_i, \theta_i, \boldsymbol{\theta}_{-i}\right)\right]\\
&-\max_{\hat{\theta}_i \in \Theta_i}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
&+\max_{\hat{\theta}_i \in \Theta_i}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
&-\max_{\hat{\theta}_i \in \mathcal{G}(\mathcal{S}_i, M, \epsilon)}\left\{\frac{1}{N}\sum_{j=1}^{N} u_{i,M}\left(\theta_i^{(j)}, \hat{\theta}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\theta_i^{(j)}, \theta_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\}\\
&\leq +\epsilon + \tilde{O}\left(\sqrt{\frac{d_i}{N}}\right).
\end{aligned}$$

The bound on $\left|\mathcal{G}\left(\mathcal{S}_i, M, \epsilon\right)\right|$ follows from Lemma C.1. □

Next, the following is a variation on Lemma 3.4. For the rest of this section, we assume $\Theta_i = [0,1]^m$.

Lemma F.3. *Suppose that for each $i \in [n]$, there exist $L_i, k_i, w_i \in \mathbb{R}$ such that with probability $1-\delta$ over the n sets $\mathcal{S}_i = \left\{\boldsymbol{\theta}^{(j)}\right\}_{j \in [N]}$, for each $M \in \mathcal{M}$ and $i \in [n]$, the functions*

$$u_{i,M}\left(\boldsymbol{\theta}_i^{(1)}, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i^{(N)}, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$$

are piecewise L_i-Lipschitz and (w_i, k_i)-dispersed. Then with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_i \sim \mathcal{D}^N$, for all $M \in \mathcal{M}$ and $i \in [n]$,

$$\begin{aligned}
&\max_{\hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} \\
&\leq \max_{\hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + L_i w_i + \frac{2k_i}{N}. \qquad (20)
\end{aligned}$$

Proof. By definition of dispersion, we know that with probability $1-\delta$ over the draw of the n sets $\mathcal{S}_i = \left\{\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}\right\} \sim \mathcal{D}^N$, for all mechanisms $M \in \mathcal{M}$, all agents $i \in [n]$, and all reported types $\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_i' \in [0,1]^m$, if $\left\|\hat{\boldsymbol{\theta}}_i - \hat{\boldsymbol{\theta}}_i'\right\|_1 \leq w_i$, then

$$\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i', \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}. \qquad (21)$$

By definition of $\mathcal{G}_w$, there must be a point $\hat{\boldsymbol{p}} \in \mathcal{G}_w$ such that $\left\|\hat{\boldsymbol{p}} - \hat{\boldsymbol{\theta}}_i\right\|_1 \leq w_i$. By Equation (21),

$$\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \leq L_i w_i + \frac{2k_i}{N}.$$

Therefore, with probability $1-\delta$

$$\begin{aligned}
&\left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \\
\leq\ & \left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \\
+\ & \left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| \\
\leq\ & \left|\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{p}}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right| + L_i w_i + \frac{k_i}{N}.
\end{aligned}$$

Since $\hat{\boldsymbol{p}} \in \mathcal{G}_{w_i}$, the lemma statement holds. □

We conclude with a variation on Theorem 3.5.

Theorem F.4. *Suppose that for each* $i \in [n]$*, there exist* $L_i, k_i, w_i \in \mathbb{R}$ *such that with probability* $1-\delta$ *over the* n *sets* $\mathcal{S}_i = \left\{\boldsymbol{\theta}^{(j)}\right\}_{j \in [N]}$*, for each* $M \in \mathcal{M}$ *and* $i \in [n]$*, the functions*

$$u_{i,M}\left(\boldsymbol{\theta}_i^{(1)}, \cdot, \boldsymbol{\theta}_{-i}^{(1)}\right), \dots, u_{i,M}\left(\boldsymbol{\theta}_i^{(N)}, \cdot, \boldsymbol{\theta}_{-i}^{(N)}\right)$$

are piecewise L_i*-Lipschitz and* (w_i, k_i)*-dispersed. Then with probability* $1-2\delta$*, for every mechanism* $M \in \mathcal{M}$ *and every* $i \in [n]$*,*

$$\begin{aligned}
&\max_{\hat{\boldsymbol{\theta}}_i \in [0,1]^m} \left\{\mathbb{E}_{\boldsymbol{\theta} \sim \mathcal{D}}\left[u_{i,M}\left(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}\right) - u_{i,M}\left(\boldsymbol{\theta}_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}\right)\right]\right\} \\
&\leq \max_{\hat{\boldsymbol{\theta}}_i \in \mathcal{G}_{w_i}} \left\{\frac{1}{N}\sum_{j=1}^N u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \hat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{-i}^{(j)}\right) - u_{i,M}\left(\boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_i^{(j)}, \boldsymbol{\theta}_{-i}^{(j)}\right)\right\} + \epsilon,
\end{aligned}$$

where $\epsilon = L_i w_i + \frac{k_i}{N} + 2\sqrt{\frac{2d_i}{N}\ln\frac{eN}{d_i}} + 2\sqrt{\frac{1}{2N}\ln\frac{2n}{\delta}}$ *and* $d_i = \operatorname{Pdim}\left(\mathcal{F}'_{i,\mathcal{M}}\right)$.

Proof. This theorem follows from Theorem F.1 and Lemma F.3 in the exact same way that Theorem 3.5 follows from Theorem 3.2 and Lemma 3.4. □