

# Nash Convergence of Mean-Based Learning Algorithms in First Price Auctions

Xiaotie Deng  
Peking University  
Beijing, China  
xiaotie@pku.edu.cn

Tao Lin  
Harvard University  
Cambridge, MA, USA  
tlin@g.harvard.edu

Xinyan Hu  
Peking University  
Beijing, China  
xy.hu@pku.edu.cn

Weiqiang Zheng  
Yale University  
New Haven, CT, USA  
weiqiang.zheng@yale.edu

## ABSTRACT

Understanding the convergence properties of learning dynamics in repeated auctions is a timely and important question in the area of learning in auctions, with numerous applications in, e.g., online advertising markets. This work focuses on repeated first price auctions where bidders with fixed values for the item learn to bid using mean-based algorithms – a large class of online learning algorithms that include popular no-regret algorithms such as Multiplicative Weights Update and Follow the Perturbed Leader. We completely characterize the learning dynamics of mean-based algorithms, in terms of convergence to a Nash equilibrium of the auction, in two senses: (1) *time-average*: the fraction of rounds where bidders play a Nash equilibrium approaches 1 in the limit; (2) *last-iterate*: the mixed strategy profile of bidders approaches a Nash equilibrium in the limit. Specifically, the results depend on the number of bidders with the highest value:

- If the number is at least three, the bidding dynamics almost surely converges to a Nash equilibrium of the auction, both in time-average and in last-iterate.
- If the number is two, the bidding dynamics almost surely converges to a Nash equilibrium in time-average but not necessarily in last-iterate.
- If the number is one, the bidding dynamics may not converge to a Nash equilibrium in time-average nor in last-iterate.

Our discovery opens up new possibilities in the study of convergence dynamics of learning algorithms.

## KEYWORDS

first price auctions; online learning; mean-based algorithms

## 1 INTRODUCTION

First price auctions are the current trend in online advertising auctions. A major example is Google Ad Exchange’s switch from second price auctions to first price auctions in 2019 [23, 41].

Compared to second price auctions, first price auctions are non-truthful: bidders need to reason about other bidders’ private values and bidding strategies and choose their own bids accordingly to maximize their utilities. Finding a good bidding strategy used to

be a difficult task due to each bidder’s lack of information of other bidders. But given the repeated nature of online advertising auctions and with the advance of computing technology, nowadays’ bidders are able to learn to bid using automated bidding algorithms. As one bidder adjusts bidding strategies using a learning algorithm, other bidders’ utilities are affected and thus they will adjust their strategies as well. Then, a natural question follows: *if all bidders in a repeated first price auction use some learning algorithms to adjust bidding strategies at the same time, will they converge to a Nash equilibrium of the auction?*

A partial answer to this question is given by [27] which shows that, in a repeated first price auction where bidders have fixed values for the item, a Nash equilibrium may or may not be learned by the *Fictitious Play* algorithm, where in each round of auctions every bidder best responds to the empirical distributions of other bidders’ bids in history. Fictitious Play, however, is a deterministic algorithm that does not have the *no-regret* property — a desideratum for learning algorithms in adversarial environments. The no-regret property can only be obtained by randomized algorithms [42]. As observed by [39] that bidders’ behavior on Bing’s advertising system is consistent with no-regret learning, it is hence important, from both theoretical and practical points of view, to understand the convergence property of no-regret algorithms in repeated first price auctions. This motivates our work.

*Our contributions.* Focusing on repeated first price auctions where bidders have fixed values, we completely characterize the Nash convergence property of a wide class of randomized online learning algorithms called “mean-based algorithms” [8]. This class contains most of popular no-regret algorithms, including Multiplicative Weights Update (MWU), Follow the Perturbed Leader (FTPL), etc..

We systematically analyze two notions of Nash convergence: (1) *time-average*: the fraction of rounds where bidders play a Nash equilibrium approaches 1 in the limit; (2) *last-iterate*: the mixed strategy profile of bidders approaches a Nash equilibrium in the limit. Specifically, the results depend on the number of bidders with the highest value:

- If the number is at least three, the bidding dynamics of mean-based algorithms almost surely converges to Nash equilibrium, both in time-average and in last-iterate.
- If the number is two, the bidding dynamics almost surely converges to Nash equilibrium in time-average but not necessarily in last-iterate.

Appears at the 1st Workshop on Learning with Strategic Agents (LSA 2022). Held as part of the Workshops at the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), N. Bishop, M. Han, L. Tran-Thanh, H. Xu, H. Zhang (chairs), May 9–10, 2022, Online.

- If the number is one, the bidding dynamics may not converge to Nash equilibrium in time-average nor in last-iterate.

For the last case, the above non-convergence result is proved for the Follow the Leader algorithm, which is a mean-based algorithm that is not necessarily no-regret. We also show by experiments that no-regret mean-based algorithms such as MWU and  $\varepsilon_t$ -Greedy may not last-iterate converge to a Nash equilibrium.

*Intuitions and techniques.* The intuition behind our convergence results (the first two cases above) relates to the notion of “iterated elimination of dominated strategies” in game theory. Suppose there are three bidders all having a same integer value  $v$  for the item and choosing bids from the set  $\{0, 1, \dots, v-1\}$ . The unique Nash equilibrium is all bidders bidding  $v-1$ . The elimination of dominated bids is as follows: firstly, bidding 0 is dominated by bidding 1 for each of the three bidders no matter what other bidders bid, so bidders will learn to bid 1 or higher instead of bidding 0 at the beginning; then, given that no bidders bid 0, bidding 1 is dominated by bidding 2, so all bidders learn to bid at least 2; ...; in this way all bidders learn to bid  $v-1$ .<sup>1</sup>

The above intuition is only high-level. In particular, since bidders use mean-based algorithms which may pick a dominated bid with a small but positive probability, additional argument is needed to show that bidders will finally converge to  $v-1$  with high probability. To do this we borrow a technique (which is a combination of time-partitioning and Azuma’s inequality) from [16] which shows that bidders in a second price auction with multiple Nash equilibria converge to the truthful equilibrium if they use mean-based algorithms with an initial uniform exploration stage. Their argument relies on the fact that, in a second price auction, all bidders learn the truthful Nash equilibrium with high probability during the uniform exploration stage. In contrast, we allow any mean-based algorithms without an initial uniform exploration stage.

## 1.1 Discussion

An assumption made in our work is that each bidder has a fixed value for the item sold throughout the repeated auction. Seemingly restrictive, this assumption can be justified in several aspects. First, fixed value is in fact a quite common assumption in the literature on repeated auctions, in various contexts including value inference [39], dynamic pricing [2, 14, 29], as well as the study of bidding equilibrium [5, 27, 30, 33]. An exception is [16] which studies repeated first price auctions under the Bayesian assumption that bidders’ values are i.i.d. samples from a distribution at every round. However, their result is restricted to a 2-symmetric-bidder setting with the Uniform[0, 1] distribution where the Bayesian Nash equilibrium (BNE) is simply every bidder bidding half of their values. For general asymmetric distributions there is no explicit characterization of the BNE [34–36] despite the existence of (inefficient) numerical approximations [15, 18, 21, 43]. No known algorithms can compute BNE efficiently for all asymmetric distributions, let alone a simple, generic learning algorithm.<sup>2</sup>

<sup>1</sup>This logic has been implicitly spelled out by [27]. But their formal argument only works for deterministic algorithms like Fictitious Play.

<sup>2</sup>Recent work [19] even shows that computing a BNE in a first price auction where bidders have *subjective* priors over others’ types is PPAD-complete.

Moreover, in real-life auctions, fixed values do occur if a same item is sold repeatedly, bidders have stable values for that item, and the set of bidders is fixed. An example is a few large online travel agencies (Agoda, Airbnb, and Booking.com) competing for an ad slot about “hotel booking”. In such Internet advertising auction scenarios, auctions sometimes happen frequently – a large number of auctions happen during a short amount of time. Even if the value of a bidder changes it would not change a lot in this short time, during which bidders may be able to converge to the Nash equilibrium before the value changes dramatically.

Finally, as we show, even with the seemingly innocuous assumption of fixed values, the learning dynamics of mean-based algorithms already exhibits complicated behaviors: it may converge to different equilibria in different runs or not converge at all. One can envision more unpredictable behaviors when values are not fixed.

*Learning in general games.* Our work is related to a fundamental question in the field of Learning in Games [10, 22, 40]: if players in a repeated game employ online learning algorithms to adjust strategies, will they converge to an equilibrium? And what kinds of equilibrium? Classical results include the convergence of no-regret learning algorithms to a *Coarse Correlated Equilibrium (CCE)* and no-internal-regret algorithms to a *Correlated Equilibrium* in any game [20, 26]. But given that (coarse) correlated equilibria are much weaker than the archetypical solution concept of a Nash equilibrium, a more appealing and challenging question is the convergence towards Nash equilibrium. Positive answers to this question are only known for some special cases of algorithms and games: e.g., no-regret algorithms converge to Nash equilibria in zero-sum games,  $2 \times 2$  games, and routing games [10, 22, 40]. In contrast, several works give non-convergence examples: e.g., the non-convergence of MWU in a  $3 \times 3$  game [11] and Regularized Learning Dynamics in zero-sum games [37]. In this work we study the Nash equilibrium convergence property in first price auctions for a large class of learning algorithms, namely the mean-based algorithms, and provide both positive and negative results.

*Last v.s. average iterate convergence.* We emphasize that previous results on convergence of learning dynamics to Nash equilibria in games are mostly attained in an average sense, i.e., the empirical distributions of players’ actions converge. Our notion of time-average convergence, which requires players play a Nash equilibrium in almost every round, is different from the convergence of empirical distributions; in fact, ours is stronger if the Nash equilibrium is unique. Nevertheless, time-average convergence fails to capture the full picture of the dynamics since players’ last-iterate (mixed) strategy profile may not converge. Existing results about last-iterate convergence show that most of learning dynamics actually diverge or enter a limit cycle even in a simple  $3 \times 3$  game [11] or zero-sum games [37], except for a few convergence examples like optimistic gradient descent/ascent in two-player zero-sum games [12, 45]. Our results and techniques, regarding the convergence of any mean-based algorithm in first price auctions, shed light on further study of last-iterate convergence in more general settings.

## 1.2 Additional Related Works

We review additional related works about online learning in repeated auctions.<sup>3</sup> A large fraction of such works are from the *seller's* perspective, i.e., studying how a seller can maximize revenue by adaptively changing the rules of the auction (e.g., reservation price) over time (e.g., [1, 2, 7–9, 13, 24, 28, 31, 38]). We focus on the *bidders'* learning problem.

Existing works from bidders' perspective are mostly about “learning to bid”, studying on how to design no-regret algorithms for a bidder to bid in various formats of repeated auctions, including first price auctions [3, 4, 25], second price auctions [30, 44], and more general auctions [17, 32]. Those works take the perspective of a *single* bidder, without considering the interaction among *multiple* bidders all learning to bid at the same time. We instead study the consequence of such interaction, showing that the learning dynamics of multiple bidders may or may not converge to the Nash equilibrium of the auction.

In addition to the aforementioned works [16] and [27], other works on Nash equilibrium convergence of online learning algorithms in first price auctions include, e.g., empirical works [5, 6, 23], and a theoretical work [33]. In particular, [33] shows, independently of our work, that in repeated first price auctions with two mean-based learning bidders, *if* the dynamics converge to some limit, then this limit must be a CCE in which the bidder with the higher value submits bids that are close to the lower value. However, they do not give conditions under which the dynamics converge. We prove that the dynamics converge if the two bidders have the same value and in fact converge to the stronger notion of a Nash equilibrium.

Finally, a recent work [46] observes that, although no-regret learning algorithms are able to learn Nash equilibria in iterative-dominance-solvable games, some such games require an exponential convergence time for many no-regret algorithms. Our fixed-value first price auction game is not iterative-dominance-solvable in all cases. And we do not know the exact convergence rate of no-regret (or mean-based) algorithms in our game.

*Organization of the paper.* Section 2 is model and preliminaries. Section 3 presents main results. Section 4 presents the proof of Theorem 4, which covers the main proof techniques of all our convergence results. Section 5 includes experiments. Section 6 concludes and discusses future directions. Missing proofs are in Appendix.

## 2 MODEL AND PRELIMINARIES

*Repeated first price auctions.* We consider repeated first-price sealed-bid auctions where a single seller sells a good to a set of  $N \geq 2$  players (bidders)  $\mathcal{N} = \{1, 2, \dots, N\}$  for infinite rounds. Each player  $i \in \mathcal{N}$  has a fixed private value  $v^i$  for the good throughout. See Section 1.1 for a discussion on this assumption. We assume that  $v^i$  is a positive integer in some range  $\{1, \dots, V\}$  where  $V$  is an upper bound on  $v^i$ . Suppose  $V \geq 3$ . No player knows other players' values. Without loss of generality, assume  $v^1 \geq v^2 \geq \dots \geq v^N$ .

At each round  $t \geq 1$  of the repeated auctions, each bidder  $i$  submits a bid  $b_t^i \in \{0, 1, \dots, V\}$  to compete for the good. A discrete set of bids captures the reality that the minimum unit of money

<sup>3</sup>We do not review works about the *batch learning* setting, e.g., sample complexity.

is a cent. The bidder with the highest bid wins the good. If there are more than one highest bidders, the good is allocated to one of them uniformly at random. The bidder who wins the good pays her bid  $b_t^i$ , obtaining utility  $v^i - b_t^i$ ; other bidders obtain utility 0. Let  $u^i(b_t^i, \mathbf{b}_t^{-i})$  denote bidder  $i$ 's (expected) utility when  $i$  bids  $b_t^i$  while other bidders bid  $\mathbf{b}_t^{-i} = (b_1^1, \dots, b_t^{i-1}, b_t^{i+1}, \dots, b_t^N)$ , i.e.,  $u^i(b_t^i, \mathbf{b}_t^{-i}) = (v^i - b_t^i)\mathbb{I}[b_t^i = \max_{j \in \mathcal{N}} b_t^j] \frac{1}{|\arg\max_{j \in \mathcal{N}} b_t^j|}$ .

We assume that bidders never bid above or equal to their values since that brings them negative or zero utility, which is clearly dominated by bidding 0. We denote the set of possible bids of each bidder  $i$  by  $\mathcal{B}^i = \{0, 1, \dots, v^i - 1\}$ .<sup>4</sup>

*Online learning.* We assume that each bidder  $i \in \mathcal{N}$  chooses her bids using an online learning algorithm. Specifically, we regard the set of possible bids  $\mathcal{B}^i$  as a set of actions (or arms). At each round  $t$ , the algorithm picks (possibly in a random way) an action  $b_t^i \in \mathcal{B}^i$  to play, and then receives some feedback. The feedback may include the rewards (i.e., utilities) of all possible actions in  $\mathcal{B}^i$  (in the *experts* setting) or only the reward of the chosen action  $b_t^i$  (in the *multi-arm bandit* setting). With feedback, the algorithm updates its choice of actions in future rounds. We do not assume a specific feedback model in this work. Our analysis will apply to all online learning algorithms that satisfy the following property, called “mean-based” [8, 16], which roughly says that the algorithm picks actions with low average historical rewards with low probabilities.

**DEFINITION 1 (MEAN-BASED ALGORITHM).** Let  $\alpha_t^i(b)$  be the average reward of action  $b$  in the first  $t$  rounds:  $\alpha_t^i(b) = \frac{1}{t} \sum_{s=1}^t u^i(b, \mathbf{b}_s^{-i})$ . An algorithm is  $\gamma_t$ -mean-based if, for any  $b \in \mathcal{B}^i$ , whenever there exists  $b' \in \mathcal{B}^i$  such that  $\alpha_{t-1}^i(b') - \alpha_{t-1}^i(b) > V\gamma_t$ , the probability that the algorithm picks  $b$  at round  $t$  is at most  $\gamma_t$ . An algorithm is mean-based if it is  $\gamma_t$ -mean-based for some decreasing sequence  $(\gamma_t)_{t=1}^\infty$  such that  $\gamma_t \rightarrow 0$  as  $t \rightarrow \infty$ .

In this work, we assume that the online learning algorithm may run for an infinite number of rounds. This captures the scenario where bidders do not know how long they will be in the auction and hence use learning algorithms that work for an arbitrary unknown number of rounds. Infinite-round mean-based algorithms can be obtained by modifying classical finite-round online learning algorithms (e.g., MWU) with constant learning rates to have decreasing learning rates, as shown below:

**EXAMPLE 2.** Let  $(\epsilon_t)_{t=1}^\infty$  be a decreasing sequence approaching 0. The following algorithms are mean-based:

- *Follow the Leader (or Greedy):* at each round  $t \geq 1$ , each player  $i \in \mathcal{N}$  chooses an action  $b \in \arg\max_{b \in \mathcal{B}^i} \{\alpha_{t-1}^i(b)\}$  (with a specific tie-breaking rule).
- $\epsilon_t$ -*Greedy:* at each round  $t \geq 1$ , each player  $i \in \mathcal{N}$  with probability  $1 - \epsilon_t$  chooses  $b \in \arg\max_{b \in \mathcal{B}^i} \{\alpha_{t-1}^i(b)\}$ , with probability  $\epsilon_t$  chooses an action in  $\mathcal{B}^i$  uniformly at random.
- *Multiplicative Weights Update (MWU):* at each round  $t \geq 1$ , each player  $i \in \mathcal{N}$  chooses each action  $b \in \mathcal{B}^i$  with probability  $\frac{w_{t-1}(b)}{\sum_{b' \in \mathcal{B}^i} w_{t-1}(b')}$ , where  $w_t(b) = \exp(\epsilon_t \sum_{s=1}^t u^i(b, \mathbf{b}_s^{-i}))$ .<sup>5</sup>

<sup>4</sup>We could allow a bidder to bid above  $v^i - 1$ . But a rational bidder will quickly learn to not place such bids.

<sup>5</sup>We note that the MWU defined here is different from the standard MWU algorithm with decreasing parameter where the weight of each action  $w_t(b)$  is updated by

Clearly, Follow the Leader is ( $\gamma_t = 0$ )-mean-based and  $\varepsilon_t$ -Greedy is  $\varepsilon_t$ -mean-based. One can see [8] for why MWU is mean-based. Additionally, MWU is no-regret when the sequence  $(\varepsilon_t)_{t=1}^\infty$  is set to  $\varepsilon_t = O(1/\sqrt{t})$  (see, e.g., Theorem 2.3 in [10]).

*Equilibria in first price auctions.* Before presenting our main results, we characterize the set of all Nash equilibria in the first price auction where bidders have fixed values  $v^1 \geq v^2 \geq \dots \geq v^N$ . We only consider pure-strategy Nash equilibria in this work.<sup>6</sup> Reusing the notation  $u^i(\cdot)$ , we denote by  $u^i(b^i, b^{-i})$  the utility of bidder  $i$  when she bids  $b^i$  while others bid  $b^{-i} = (b^1, \dots, b^{i-1}, b^{i+1}, \dots, b^N)$ . A bidding profile  $b = (b^1, \dots, b^N) = (b^i, b^{-i})$  is called a *Nash equilibrium* if  $u^i(b) \geq u^i(b', b^{-i})$  for any  $b' \in \mathcal{B}^i$  and any  $i \in N$ . Let  $M^1$  denote the set of bidders who have the same value as bidder  $i$ ,  $M^1 = \{j \in N : v^j = v^i\}$ .  $M^1$  is the set of bidders with the highest value.

**PROPOSITION 3.** *The set of (pure-strategy) Nash equilibria in the first price auctions with fixed values  $v^1 \geq v^2 \geq \dots \geq v^N$  are bidding profiles  $b = (b^1, \dots, b^N)$  that satisfy the following:*

- *The case of  $|M^1| \geq 3$ :  $b^i = v^1 - 1$  for  $i \in M^1$  and  $b^j \leq v^1 - 2$  for  $j \notin M^1$ .*
- *The case of  $|M^1| = 2$ :*
  - *If  $N = 2$  or  $v^1 = v^2 > v^3 + 1$ : there are two types of Nash equilibria: (1)  $b^1 = b^2 = v^1 - 1$ , with  $b^j \leq v^1 - 3$  for  $j \notin M^1$ ; (2)  $b^1 = b^2 = v^1 - 2$ , with  $b^j \leq v^1 - 3$  for  $j \notin M^1$ .*
  - *If  $N > 2$  and  $v^1 = v^2 = v^3 + 1$ :  $b^1 = b^2 = v^1 - 1$  and  $b^j \leq v^1 - 2$  for  $j \notin M^1$ .*
- *The case of  $|M^1| = 1$ :*
  - *Bidding profiles that satisfy the following are Nash equilibria:  $b^1 = v^2$ , at least one bidder in  $M^2$  bids  $v^2 - 1$ , all other bidders bid  $b^j \leq v^2 - 1$ .*
  - *If  $v^1 = v^2 + 1$  and  $|M^2| = 1$ , then there is another type of Nash equilibria:  $b^1 = b^2 = v^2 - 1$ ,  $b^j \leq v^2 - 2$  for  $j \notin \{1, 2\}$ .*

*There are no other (pure-strategy) Nash equilibria.*

The proof of this proposition is straightforward and omitted. Intuitively, whenever more than one bidder has the highest value ( $|M^1| \geq 2$ ), they should compete with each other by bidding  $v^1 - 1$  (or  $v^1 - 2$  if  $|M^1| = 2$  and no other bidders are able to compete with them). When  $|M^1| = 1$ , the unique highest-value bidder (bidder 1) competes with the second-highest bidders ( $M^2$ ).

### 3 MAIN RESULTS: CONVERGENCE OF MEAN-BASED ALGORITHMS

We introduce some additional notations. Let  $\mathbf{x}_t^i \in \mathbb{R}^{v^i}$  be the mixed strategy of player  $i$  in round  $t$ , where the  $b$ -th component of  $\mathbf{x}_t^i$  is the probability that player  $i$  bids  $b \in \mathcal{B}^i$  in round  $t$ . The sequence  $(\mathbf{x}_t^i)_{t=1}^\infty$  is a stochastic process, where the randomness for each  $\mathbf{x}_t^i$  includes the randomness of bidding by all players in previous rounds. Let  $\mathbf{1}_b = (0, \dots, 0, 1, 0, \dots, 0)$  where 1 is in the  $b$ -th position.

<sup>6</sup> $w_t(b) = w_{t-1}(b) \cdot \exp(\varepsilon_t u^i(b, b_t^{-i})) = \exp(\sum_{s=1}^t \varepsilon_s u^i(b, b_s^{-i}))$ . The standard algorithm is not mean-based because rewards  $u^i(b, b_s^{-i})$  in earlier rounds matter more than rewards in later rounds given that  $\varepsilon_s$  is decreasing. The algorithm we define here treat rewards at different rounds equally and is hence mean-based.

<sup>6</sup>Whether and how our results extend to mixed-strategy Nash equilibria is open.

Our main results about the convergence of mean-based algorithms in repeated first price auctions depend on how many bidders have the highest value,  $|M^1|$ .

*The case of  $|M^1| \geq 3$ .*

**THEOREM 4.** *If  $|M^1| \geq 3$  and every bidder follows a mean-based algorithm, then, with probability 1, both of the following events happen:*

- *Time-average convergence of bid sequence:*

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 1] = 1. \quad (1)$$

- *Last-iterate convergence of mixed strategy profile:*

$$\forall i \in M^1, \lim_{t \rightarrow \infty} \mathbf{x}_t^i = \mathbf{1}_{v^1-1}. \quad (2)$$

Theorem 4 can be interpreted as follows. According to Proposition 3, the bidding profile  $b_s$  at round  $s$  is a Nash equilibrium if and only if  $\forall i \in M^1, b_s^i = v^1 - 1$  (bidders not in  $M^1$  bid  $\leq v^1 - 2$  by assumption<sup>7</sup>). Hence, the first result of Theorem 4 implies that the fraction of rounds where bidders play a Nash equilibrium approaches 1 in the limit. The second result shows that all bidders  $i \in M^1$  bid  $v^1 - 1$  with certainty eventually, achieving a Nash equilibrium.

*The case of  $|M^1| = 2$ .*

**THEOREM 5.** *If  $|M^1| = 2$  and every bidder follows a mean-based algorithm, then, with probability 1, one of the following two events happens:*

- $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 2] = 1$ ;
- $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 1] = 1$  and  $\forall i \in M^1, \lim_{t \rightarrow \infty} \mathbf{x}_t^i = \mathbf{1}_{v^1-1}$ .

Moreover, if  $N > 2$  and  $v^3 = v^1 - 1$  then only the second event happens.

For the case  $v^3 < v^1 - 1$ , according to Proposition 3,  $b_s$  is a Nash equilibrium if and only if both bidders in  $M^1$  play  $v^1 - 1$  or  $v^1 - 2$  at the same time (with other bidders bidding  $\leq v^1 - 3$  by assumption<sup>7</sup>). Hence, the theorem shows that the bidders eventually converge to one of the two possible types of equilibria. Interestingly, experiments show that some mean-based algorithms lead to the equilibrium of  $v^1 - 1$  while some lead to  $v^1 - 2$ . Also, a same algorithm may converge to different equilibria in different runs. See Section 5 for details.

In the case of time-average convergence to the equilibrium of  $v^1 - 2$ , the last-iterate convergence result does not always hold. Consider an example with 2 bidders, with  $v^1 = v^2 = 3$ . We can construct a  $\gamma_t$ -mean-based algorithm with  $\gamma_t = O(\frac{1}{t^{1/4}})$  such that, with constant probability, it holds  $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 2] = 1$  but in infinitely many rounds we have  $\mathbf{x}_t^i = \mathbf{1}_2 = \mathbf{1}_{v^1-1}$ . The key idea is that, when  $\alpha_t^i(1) - \alpha_t^i(2)$  is positive but lower than  $V\gamma_t$  in some round  $t$  (which happens infinitely often), we can let the algorithm bid 2 with certainty in round  $t + 1$ . This does not violate the  $\gamma_t$ -mean-based property.

<sup>7</sup>We note that the bidders not in  $M^1$  can follow a mixed strategy and need not converge to a deterministic bid.

**PROPOSITION 6.** *If  $|M^1| = 2$ , then there exists a mean-based algorithm such that, when players follow this algorithm, with constant probability their mixed strategy profiles do not converge to a Nash equilibrium in last-iterate.*

*The case of  $|M^1| = 1$ .* The dynamics may not converge to a Nash equilibrium of the auction in time-average nor in last-iterate, as shown in the following example (see Appendix A for a proof).

**EXAMPLE 7.** *Let  $v^1 = 10$ ,  $v^2 = v^3 = 7$ . Assume that players use the Follow the Leader algorithm (which is 0-mean-based) with a specific tie-breaking rule. They may generate the following bidding path  $(b_t^1, b_t^2, b_t^3)_{t \geq 1}$ :*

$$(7, 6, 1), (7, 1, 6), (7, 1, 1), (7, 6, 1), (7, 1, 6), (7, 1, 1), \dots$$

*Note that  $(7, 1, 1)$  is not a Nash equilibrium according to Proposition 3 but it appears in  $\frac{1}{3}$  fraction of rounds, which means that the dynamics neither converges in the time-average sense nor in the last-iterate sense to a Nash equilibrium.*

Example 7 also shows that, in the case of  $|M^1| = 1$ , the bidding dynamics generated by a mean-based algorithm may not converge to Nash equilibrium in the classical sense of “convergence of empirical distribution”: i.e., letting  $p_t^i = \frac{1}{t} \sum_{s=1}^t \mathbb{1}_{b_s^i} \in \Delta(\mathcal{B}^i)$  denote the empirical distribution of player  $i$ ’s bids up to round  $t$ , the vector of individual empirical distributions  $(p_t^1, p_t^2, p_t^3)_{t \geq 1}$  approaches a (mixed-strategy) Nash equilibrium in the limit. To see this, note that the vector of individual empirical distributions converges to  $(p^1, p^2, p^3)$  where  $p^1(7) = 1$  and for  $i = 2, 3$ ,  $p^i(6) = \frac{1}{3}$  and  $p^i(1) = \frac{2}{3}$ . It is easy to verify that for bidder 1, bidding 2 has utility  $(10 - 2)(\frac{2}{3})^2 = \frac{32}{9}$ , which is larger than the utility of bidding 7, which is  $10 - 7 = 3$ . Thus,  $(p^1, p^2, p^3)$  is not a Nash equilibrium.

The mean-based algorithm in Example 7 is not no-regret. In Section 5 we show by experiments that such non-convergence results also hold for no-regret mean-based algorithms, e.g., MWU.

## 4 PROOF OF THEOREM 4

The proof of Theorem 4 covers the main ideas and proof techniques of our convergence results, so we present it here. We first provide a proof sketch. Then in Section 4.1 we provide properties of mean-based algorithms that will be used in the formal proof. Section 4.2 and Section 4.3 prove Theorem 4.

*Proof sketch.* The idea of the proof resembles the notion of iterative elimination of dominated strategies in game theory. We first use a step-by-step argument to show that bidders with the highest value (i.e., those in  $M^1$ ) will gradually learn to avoid bidding  $0, 1, \dots, v^1 - 3$ . Then we further prove that: if  $|M^1| = 3$ , they will avoid  $v^1 - 2$  and hence converge to  $v^1 - 1$ ; if  $|M^1| = 2$ , the two bidders may end up playing  $v^1 - 1$  or  $v^1 - 2$ .

To see why bidders in  $M^1$  will learn to avoid 0, suppose that there are two bidders in total and one of them (say, bidder  $i$ ) bids  $b$  with probability  $P(b)$  in history. For the other bidder (say, bidder  $j$ ), if bidder  $j$  bids 0, she obtains utility  $\alpha(0) = (v^1 - 0)\frac{P(0)}{2}$ ; if she bids 1, she obtains utility  $\alpha(1) = (v^1 - 1)(P(0) + \frac{P(1)}{2})$ . Since  $\alpha(1) - \alpha(0) = \frac{v^1 - 2}{2}P(0) + (v^1 - 1)\frac{P(1)}{2} > 0$  (assuming  $v^1 \geq 3$ ), bidding 1 is better than bidding 0 for bidder  $j$ . Given that bidder  $j$  is using a mean-based algorithm, she will play 0 with small probability

(say, zero probability). The same argument applies to bidder  $i$ . Hence, both bidders learn to not play 0. Then we take an inductive step: assuming that no bidders play  $0, \dots, k - 1$ , we note that  $\alpha(k+1) - \alpha(k) = \frac{v^1 - k - 2}{2}P(k) + \frac{v^1 - k - 1}{2}P(k+1) > 0$  for  $k \leq v^1 - 3$ , therefore  $k+1$  is a better response than  $k$  and both players will avoid bidding  $k$ . An induction shows that they will finally learn to avoid  $0, 1, \dots, v^1 - 3$ . Then, for the case of  $|M^1| \geq 3$ , we will use an additional claim (Claim 9) to show that, if bidders bid  $0, 1, \dots, v^1 - 3$  rarely in history, they will also avoid bidding  $v^1 - 2$  in the future.

The formal proof uses a time-partitioning technique proposed by [16]. Roughly speaking, we partition the time horizon into some periods  $1 < T_0 < T_1 < T_2 < \dots$ . If bidders bid  $0, 1, \dots, k - 1$  with low frequency from round 1 to  $T_{k-1}$ , then using the mean-based properties in Claim 8 and Claim 9, we show that they will bid  $k$  with probability at most  $\gamma_t$  in each round from  $T_{k-1} + 1, T_{k-1} + 2, \dots$ , to  $T_k$ . A use of Azuma’s inequality shows that the frequency of bid  $k$  in period  $(T_{k-1}, T_k]$  is also low with high probability, which concludes the induction. Constructing an appropriate partition allows us to argue that the frequency of bids less than  $v^1 - 1$  converges to 0 with high probability.

## 4.1 Properties of Mean-Based Algorithms in First Price Auctions

We use the following notations intensively in the proofs. We denote by  $P_t^i(k)$  the frequency of the highest bid submitted by bidders other than  $i$  during the first  $t$  rounds:  $P_t^i(k) = \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\max_{j \neq i} b_s^j = k]$ . By  $P_t^i(0 : k)$  we mean  $\sum_{\ell=0}^k P_t^i(\ell)$ . Let  $P_t^i(0 : -1)$  be 0. Additionally, let  $Q_t^i(k)$  be the probability of bidder  $i$  winning the item with ties if she bids  $k$  in history:  $Q_t^i(k) = \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\max_{j \neq i} b_s^j = k] \frac{1}{|\arg\max_{j \neq i} b_s^j| + 1}$ . Clearly,

$$0 \leq \frac{1}{N} P_t^i(k) \leq Q_t^i(k) \leq \frac{1}{2} P_t^i(k) \leq \frac{1}{2}. \quad (3)$$

We can use  $P_t^i(0 : k - 1)$  and  $Q_t^i(k)$  to express  $\alpha_t^i(k)$ :

$$\alpha_t^i(k) = (v^i - k)(P_t^i(0 : k - 1) + Q_t^i(k)). \quad (4)$$

We use  $H_t$  to denote the history of the first  $t$  rounds, which includes the realization of all randomness in the first  $t$  rounds. Bidders themselves may not observe the full history  $H_t$ . Given  $H_{t-1}$ , each bidder’s mixed strategy  $x_t^i$  at round  $t$  is determined, and the  $k$ -th component of  $x_t^i$  is exactly  $\Pr[b_t^i = k | H_{t-1}]$ . The following claim says that, if other bidders rarely bid 0 to  $k - 1$  in history, then bidder  $i$  will not bid  $k$  with large probability in round  $t$ , for  $i \in M^1$ .

**CLAIM 8.** *Assume  $v^1 \geq 3$ . For any  $i \in M^1$ , any  $k \in \{0, 1, \dots, v^1 - 3\}$ , any  $t \geq 1$ , if the history of the first  $t - 1$  rounds  $H_{t-1}$  satisfies  $P_{t-1}^i(0 : k - 1) < \frac{1}{2\sqrt{N}} - 2\gamma_t$ , then  $\Pr[b_t^i = k | H_{t-1}] \leq \gamma_t$ .*

**PROOF.** Suppose  $P_{t-1}^i(0 : k - 1) \leq \frac{1}{2\sqrt{N}} - 2\gamma_t$  holds. If  $\alpha_{t-1}^i(k+1) - \alpha_{t-1}^i(k) > V\gamma_t$ , then by the mean-based property, the conditional probability  $\Pr[b_t^i = k | \alpha_{t-1}^i(k+1) - \alpha_{t-1}^i(k) > V\gamma_t, H_{t-1}]$  is at most  $\gamma_t$ . Otherwise, using (4) and (3), we have

$$\begin{aligned} V\gamma_t &\geq \alpha_{t-1}^i(k+1) - \alpha_{t-1}^i(k) \\ &\geq (v^1 - k - 1)P_{t-1}^i(k) - P_{t-1}^i(0 : k - 1) - (v^1 - k) \frac{P_{t-1}^i(k)}{2}, \end{aligned}$$

which implies

$$P_{t-1}^i(k) \leq \frac{2}{v^1-k-2} (V\gamma_t + P_{t-1}^i(0 : k-1)). \quad (5)$$

And hence,

$$\begin{aligned} a_{t-1}^i(k) &\leq (v^1 - k)(P_{t-1}^i(0 : k-1) + \frac{1}{2}P_{t-1}^i(k)) \\ \text{by (5)} &\leq (v^1 - k)P_{t-1}^i(0 : k-1) + \frac{v^1-k}{v^1-k-2}(V\gamma_t + P_{t-1}^i(0 : k-1)) \\ &= \frac{v^1-k}{v^1-k-2}V\gamma_t + (v^1 - k + \frac{v^1-k}{v^1-k-2})P_{t-1}^i(0 : k-1) \\ (\text{since } \frac{v^1-k}{v^1-k-2} &\leq 3) \leq 3V\gamma_t + (v^1 - k + 3)P_{t-1}^i(0 : k-1) \\ &\leq 3V\gamma_t + 2VP_{t-1}^i(0 : k-1). \end{aligned}$$

By the assumption that  $P_{t-1}^i(0 : k-1) < \frac{1}{2VN} - 2\gamma_t$ , we get

$$a_{t-1}^i(k) < 3V\gamma_t + 2V(\frac{1}{2VN} - 2\gamma_t) = \frac{1}{N} - V\gamma_t.$$

Then, we note that  $a_{t-1}^i(v^1-1) = P_{t-1}^i(0 : v^1-2) + Q_{t-1}^i(v^1-1) \geq \frac{1}{N}P_{t-1}^i(0 : v^1-1) = \frac{1}{N} \cdot 1$  where the last equality holds because no bidder bids above  $v^1 - 1$  by assumption. Therefore,

$$a_{t-1}^i(v^1 - 1) - a_{t-1}^i(k) > \frac{1}{N} - (\frac{1}{N} - V\gamma_t) = V\gamma_t.$$

From the mean-based property,  $\Pr[b_t^i = k \mid a_{t-1}^i(k+1) - a_{t-1}^i(k) \leq V\gamma_t, H_{t-1}] \leq \gamma_t$ , implying  $\Pr[b_t^i = k \mid H_{t-1}] \leq \gamma_t$ .  $\square$

The following claim is for the case of  $k = v^1 - 2$ : if bidders rarely bid 0 to  $v^1 - 3$  in history, then bidder  $i$  will not bid  $v^1 - 2$  with large probability in round  $t$ , for  $i \in M^1$ , given  $|M^1| \geq 3$ .

**CLAIM 9.** Suppose  $|M^1| \geq 3$  and  $v^1 \geq 2$ . For any  $t \geq 1$  such that  $\gamma_t < \frac{1}{12N^2V^2}$ , if the history  $H_{t-1}$  of the first  $t-1$  rounds satisfies  $\frac{1}{T-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq \frac{1}{3NV}$ , then,  $\forall i \in M^1$ ,  $\Pr[b_t^i = v^1 - 2 \mid H_{t-1}] \leq \gamma_t$ .

## 4.2 Iteratively Eliminating Bids $0, 1, \dots, v^1 - 3$

In this subsection we show that, after a sufficiently long time, bidders in  $M^1$  will rarely bid  $0, 1, \dots, v^1 - 3$ , with high probability (Corollary 13). We show this by partitioning the time horizon into  $v^1 - 3$  periods and using an induction from 0 to  $v^1 - 3$ . Let constants  $c = 1 + \frac{1}{12NV}$  and  $d = \lceil \log_c(8NV) \rceil$ . Let  $T_b$  be any (constant) integer such that  $\gamma_{T_b} < \frac{1}{12N^2V^2}$  and  $\exp(-\frac{(c-1)T_b}{1152N^2V^2}) \leq \frac{1}{2}$ . Let  $T_0 = 12NVT_b$  and  $T_k = c^d T_{k-1} = c^{dk} T_0 \geq (8NV)^k T_0$  for  $k \in \{1, 2, \dots, v^1 - 3\}$ . Let  $A_k$  be event

$$A_k = \left[ \frac{1}{T_k} \sum_{t=1}^{T_k} \mathbb{I}[\exists i \in M^1, b_t^i \leq k] \leq \frac{1}{4NV} \right],$$

which says that bidders in  $M^1$  bid  $0, 1, \dots, k$  not too often in the first  $T_k$  rounds. Our goal is to show that  $\Pr[A_{v^1-3}]$  is high.

**LEMMA 10.**  $\Pr[A_0] \geq 1 - \exp\left(-\frac{T_b}{24NV}\right)$ .

**PROOF.** Consider any round  $t \geq T_b$ . For any  $i \in M^1$ , given any history  $H_{t-1}$  of the first  $t-1$  rounds, it holds that  $P_{t-1}^i(0 : -1) = 0 \leq \frac{1}{2VN} - 2\gamma_t$ . Hence, by Claim 8,  $\Pr[b_t^i = 0 \mid H_{t-1}] \leq \gamma_t$ . Using a union bound over  $i \in M^1$ ,

$$\Pr[\exists i \in M^1, b_t^i = 0 \mid H_{t-1}] \leq |M^1|\gamma_t.$$

Let  $Z_t = \mathbb{I}[\exists i \in M^1, b_t^i = 0] - |M^1|\gamma_t$  and let  $X_t = \sum_{s=T_b+1}^t Z_s$ . We have  $\mathbb{E}[Z_t \mid H_{t-1}] \leq 0$ . Therefore, the sequence  $X_{T_b+1}, X_{T_b+2}, \dots, X_{T_b+1}$  is a supermartingale (with respect to the sequence of history  $H_{T_b}, H_{T_b+1}, \dots, H_{T_b+1}$ ). By Azuma's inequality, we have

$$\Pr\left[\sum_{t=T_b+1}^{T_0} Z_t \geq \Delta\right] \leq \exp\left(-\frac{\Delta^2}{2(T_0-T_b)}\right).$$

Let  $\Delta = T_b$ , so with probability at least  $1 - \exp\left(-\frac{\Delta^2}{2(T_0-T_b)}\right) \geq 1 - \exp\left(-\frac{T_b}{24NV}\right)$ , we have  $\sum_{t=T_b+1}^{T_0} Z_t < T_b$ , or  $\sum_{t=T_b+1}^{T_0} \mathbb{I}[\exists i \in M^1, b_t^i = 0] < T_b + \sum_{t=T_b+1}^{T_0} |M^1|\gamma_t$ , which implies

$$\begin{aligned} \frac{1}{T_0} \sum_{t=1}^{T_0} \mathbb{I}[\exists i \in M^1, b_t^i = 0] &\leq \frac{1}{T_0} \left( T_b + \sum_{t=T_b+1}^{T_0} \mathbb{I}[\exists i \in M^1, b_t^i = 0] \right) \\ &< \frac{1}{T_0} \left( 2T_b + \sum_{t=T_b+1}^{T_0} |M^1|\gamma_t \right) \leq \frac{1}{4NV}, \end{aligned}$$

where the last inequality holds due to  $\frac{T_b}{T_0} = \frac{1}{12NV}$  and

$$\frac{1}{T_0} \sum_{t=T_b+1}^{T_0} |M^1|\gamma_t \leq |M^1|\gamma_{T_b} \leq \frac{1}{12NV}.$$

$\square$

The following lemma says that, if bidders in  $M^1$  seldom bid  $0, 1, \dots, k$  in the first  $T_k$  rounds, then they will also seldom bid  $0, 1, \dots, k+1$  in the first  $T_{k+1}$  rounds, with high probability.

**LEMMA 11.** If  $|M^1| \geq 2$ , then for  $k = 0, 1, \dots, v^1 - 4$ , we have  $\Pr[A_{k+1} \mid A_k] \geq 1 - \sum_{j=1}^d \exp\left(-\frac{|\Gamma_k^j|}{1152N^2V^2}\right)$ .

**PROOF.** Suppose  $A_k$  holds. Consider  $A_{k+1}$ . We divide the rounds in  $[T_k, T_{k+1}]$  to  $d = \lceil \log_c(8NV) \rceil$  episodes such that  $T_k = T_k^0 < T_k^1 < \dots < T_k^d = T_{k+1}$  where  $T_k^j = cT_k^{j-1}$  for  $j \in [1, d]$ . Let  $\Gamma_k^j = [T_k^{j-1} + 1, T_k^j]$ , with  $|\Gamma_k^j| = T_k^j - T_k^{j-1}$ .

We define a series of events  $B_k^j$  for  $j \in [0, d]$ .  $B_k^0$  is the same as  $A_k$ . For  $j \in [1, d]$ ,  $B_k^j$  is the event  $\sum_{t \in \Gamma_k^j} \mathbb{I}[\exists i \in M^1, b_t^i \leq k+1] \leq \frac{|\Gamma_k^j|}{8NV}$ .

**CLAIM 12.**  $\Pr[B_k^{j+1} \mid A_k, B_k^1, \dots, B_k^j] \geq 1 - \exp\left(-\frac{|\Gamma_k^{j+1}|}{1152N^2V^2}\right)$ .

**PROOF.** Suppose  $A_k, B_k^1, \dots, B_k^j$  happen. For simplicity, we write  $A_k^j = [A_k, B_k^1, \dots, B_k^j]$ . Fix an  $i \in M^1$ , consider  $P_{T_k^j}^i(0 : k)$ . Recall that  $P_{T_k^j}^i(0 : k) = \frac{1}{T_k^j} \sum_{t=1}^{T_k^j} \mathbb{I}[\max_{i' \neq i} b_t^{i'} \leq k]$ . Because  $|M^1| \geq 2$ , the event  $[\max_{i' \neq i} b_t^{i'} \leq k]$  implies that there exists  $i^* \in M^1, i^* \neq i$ , such that  $b_t^{i^*} \leq k$ . Therefore  $P_{T_k^j}^i(0 : k) \leq \frac{1}{T_k^j} \sum_{t=1}^{T_k^j} \mathbb{I}[\exists i \in M^1, b_t^i \leq k]$ . Given  $A_k^j$ , we have

$$P_{T_k^j}^i(0 : k) \leq \frac{1}{T_k^j} \left( T_k \frac{1}{4NV} + (T_k^j - T_k) \frac{1}{8NV} \right) \leq \frac{1}{4NV}.$$

Then, for any round  $t \in \Gamma_k^{j+1} = [T_k^j + 1, T_k^{j+1}]$ , we have

$$\begin{aligned} P_{t-1}^i(0 : k) &= \frac{1}{t-1} \left( T_k^j P_{T_k^j}^i(0 : k) + \sum_{s=T_k^j+1}^{t-1} \mathbb{I}[\max_{i' \neq i} b_s^{i'} \leq k] \right) \\ &\leq \frac{1}{t-1} \left( T_k^j \frac{1}{4NV} + (t-1 - T_k^j) \right) \\ (\text{since } T_k^j &\leq t-1 \leq T_k^{j+1}) \leq \frac{1}{4NV} + \frac{T_k^{j+1} - T_k^j}{T_k^{j+1}} \\ (\text{since } T_k^{j+1} &= cT_k^j) \leq \frac{1}{3NV} \\ (\text{since } \gamma_t \leq \gamma_{T_b} < \frac{1}{12NV}) &< \frac{1}{2NV} - 2\gamma_t. \end{aligned}$$

Therefore, according to Claim 8, for any history  $H_{t-1}$  that satisfies  $A_k^j$  we have  $\Pr[b_t^i = b \mid H_{t-1}, A_k^j] \leq \gamma_t$  for any  $b \in [0, k+1]$ . Consider the event  $[\exists i \in M^1, b_t^i \leq k+1]$ . Using union bounds over  $i \in M^1$  and  $b \in \{0, \dots, k+1\}$ ,

$$\begin{aligned} & \Pr[\exists i \in M^1, b_t^i \leq k+1 \mid H_{t-1}, A_k^j] \\ & \leq |M^1| \cdot \Pr[b_t^i \leq k+1 \mid H_{t-1}, A_k^j] \leq |M^1|(k+2)\gamma_t \leq |M^1|V\gamma_t. \end{aligned}$$

Let  $Z_t = \mathbb{I}[\exists i \in M^1, b_t^i \leq k+1] - |M^1|V\gamma_t$  and let  $X_t = \sum_{s=T_k^{j+1}}^t Z_s$ .

We have  $\mathbb{E}[Z_t \mid A_k^j, H_{t-1}] \leq 0$ . Therefore, the sequence  $X_{T_k^{j+1}}, X_{T_k^{j+2}}, \dots, X_{T_k^{j+1}}$  is a supermartingale (with respect to the sequence of history  $H_{T_k^j}, H_{T_k^{j+1}}, \dots, H_{T_k^{j+1}-1}$ ). By Azuma's inequality, for any  $\Delta > 0$ , we have

$$\Pr\left[\sum_{t=T_k^{j+1}}^{T_k^{j+1}} Z_t \geq \Delta \mid A_k^j\right] \leq \exp\left(-\frac{\Delta^2}{2|\Gamma_k^{j+1}|}\right).$$

Let  $\Delta = \frac{|\Gamma_k^{j+1}|}{24NV}$ . Then with probability at least  $1 - \exp\left(-\frac{|\Gamma_k^{j+1}|}{1152N^2V^2}\right)$

we have  $\sum_{t \in \Gamma_k^{j+1}} Z_t < \frac{|\Gamma_k^{j+1}|}{24NV}$ , which implies

$$\begin{aligned} \sum_{t \in \Gamma_k^{j+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq k+1] & < \frac{|\Gamma_k^{j+1}|}{24NV} + \sum_{t \in \Gamma_k^{j+1}} |M^1|V\gamma_t \\ & \leq \frac{|\Gamma_k^{j+1}|}{24NV} + |M^1|V \frac{|\Gamma_k^{j+1}|}{12N^2V^2} \leq \frac{|\Gamma_k^{j+1}|}{8NV}. \end{aligned} \quad \square$$

Suppose  $A_k$  happens. Using Claim 12 with  $j = 0, 1, \dots, d-1$ , we have, with probability at least  $1 - \sum_{j=1}^d \exp\left(-\frac{|\Gamma_k^j|}{1152N^2V^2}\right)$ , all the events  $B_1^1, \dots, B_k^d$  hold, which implies

$$\begin{aligned} & \frac{1}{T_{k+1}} \sum_{t=1}^{T_{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq k+1] \\ & \leq \frac{1}{T_{k+1}} \left( T_k \cdot 1 + \sum_{t \in \Gamma_k^1 \cup \dots \cup \Gamma_k^d} \mathbb{I}[\exists i \in M^1, b_t^i \leq k+1] \right) \\ & \leq \frac{1}{T_{k+1}} \left( T_k \cdot 1 + (T_{k+1} - T_k) \cdot \frac{1}{8NV} \right) \\ & (\text{since } T_{k+1} \geq (8NV)T_k) \leq \frac{1}{8NV} + \left(1 - \frac{T_k}{T_{k+1}}\right) \frac{1}{8NV} \leq \frac{1}{4NV}. \end{aligned}$$

Thus  $A_{k+1}$  holds.  $\square$

Using an induction from  $k = 0, 1, \dots$  to  $v^1 - 4$ , we get, with probability at least  $1 - \exp\left(-\frac{T_b}{24NV}\right) - \sum_{k=0}^{v^1-4} \sum_{j=1}^d \exp\left(-\frac{|\Gamma_k^j|}{1152N^2V^2}\right)$ , all events  $A_0, A_1, \dots, A_{v^1-3}$  hold. By bounding the summation  $\sum_{k=0}^{v^1-4} \sum_{j=1}^d \exp\left(-\frac{|\Gamma_k^j|}{1152N^2V^2}\right)$  using the definition of  $\Gamma_k^j$ , we obtain:

$$\text{COROLLARY 13. } \Pr[A_{v^1-3}] \geq 1 - \exp\left(-\frac{T_b}{24NV}\right) - 2 \exp\left(-\frac{T_b}{1152N^2V^2}\right).$$

### 4.3 Eliminating $v^1 - 2$

In this subsection, we continue partitioning the time horizon after  $T_{v^1-3}$ , all the way to infinity, to show two points: (1) the frequency of bids in  $\{0, 1, \dots, v^1 - 3\}$  from bidders in  $M^1$  approaches 0; (2) the frequency of  $v^1 - 2$  also approaches 0. Again let  $c = 1 + \frac{1}{12NV}$ . Let  $T_a^0 = T_{v^1-3}$ ,  $T_a^{k+1} = cT_a^k$ ,  $\Gamma_a^{k+1} = [T_a^k + 1, T_a^{k+1}]$ ,  $k \geq 0$ . Let  $\delta_t = (\frac{1}{t})^{\frac{1}{3}}$ ,  $t \geq 0$ . For each  $k \geq 0$ , define

$$F_{T_a^k} = \frac{1}{c^k} \frac{1}{4NV} + \sum_{s=0}^{k-1} \frac{c-1}{c^{k-s}} \delta_{T_a^s} + \sum_{s=0}^{k-1} |M^1|V \frac{c-1}{c^{k-s}} \gamma_{T_a^s},$$

and

$$\widetilde{F}_{T_a^k} = \frac{1}{c^k} + \sum_{s=0}^{k-1} \frac{c-1}{c^{k-s}} \delta_{T_a^s} + \sum_{s=0}^{k-1} |M^1|V \frac{c-1}{c^{k-s}} \gamma_{T_a^s}.$$

CLAIM 14.  $F_{T_a^{k+1}} \leq F_{T_a^k} \leq \frac{1}{4NV}$  for every  $k \geq 0$  and  $\lim_{k \rightarrow \infty} F_{T_a^k} = \lim_{k \rightarrow \infty} \widetilde{F}_{T_a^k} = 0$ .

PROOF SKETCH. The claim is a direct corollary from the fact that  $\delta_t \rightarrow 0$  and  $\gamma_t \rightarrow 0$  as  $t \rightarrow \infty$ , and that  $T_a^k \rightarrow \infty$  as  $k \rightarrow 0$ .  $\square$

LEMMA 15. Suppose  $|M^1| \geq 2$ . Let  $T_b$  be any sufficiently large constant. Let  $A_a^k$  be the event that for all  $s \leq k$ ,  $\frac{1}{T_a^s} \sum_{t=1}^{T_a^s} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] \leq F_{T_a^s}$ . Then,  $\Pr[A_a^k] \geq 1 - \exp\left(-\frac{T_b}{24NV}\right) - 2 \exp\left(-\frac{T_b}{1152N^2V^2}\right) - 2 \exp\left(-\left(\frac{T_b}{1152N^2V^2}\right)^{\frac{1}{3}}\right)$ .

Moreover, if  $|M^1| \geq 3$ , we can add the following event to  $A_a^k$ : for all  $s \leq k$ ,  $\frac{1}{T_a^s} \sum_{t=1}^{T_a^s} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 2] \leq \widetilde{F}_{T_a^s}$ .

The proof is similar to that of Lemma 11 except that we use Claim 9 to argue that bidders bid  $v^1 - 2$  with low frequency.

**Proof of Theorem 4.** Suppose  $|M^1| \geq 3$ . We note that the event  $A_a^k$  implies that for any time  $t \in \Gamma_a^k = [T_a^{k-1} + 1, T_a^k]$ ,

$$\begin{aligned} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 2] & \leq \frac{1}{t} \sum_{s=1}^{T_a^k} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 2] \\ & \leq \frac{1}{t} T_a^k \widetilde{F}_{T_a^k} \quad (\text{since } t \geq \frac{1}{c} T_a^k) \\ & \leq c \widetilde{F}_{T_a^k}. \end{aligned} \quad (6)$$

We note that  $A_a^{k-1} \supseteq A_a^k$ , so by Lemma 15 with probability at least  $\Pr[\bigcap_{k=0}^{\infty} A_a^k] = \lim_{k \rightarrow \infty} \Pr[A_a^k] \geq 1 - \exp\left(-\frac{T_b}{24NV}\right) - 2 \exp\left(-\frac{T_b}{1152N^2V^2}\right) - 2 \exp\left(-\left(\frac{T_b}{1152N^2V^2}\right)^{\frac{1}{3}}\right)$  all events  $A_a^0, A_a^1, \dots, A_a^k, \dots$  happen. Then, according to (6) and Claim 14, we have

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 2] \leq \lim_{k \rightarrow \infty} c \widetilde{F}_{T_a^k} = 0.$$

Letting  $T_b \rightarrow \infty$  proves the first result of the theorem. The second result follows from the observation that, when  $\frac{1}{t} \sum_{s=1}^t \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 2] \leq \frac{1}{3NV}$ , all bidders in  $M^1$  will choose bids in  $\{0, 1, \dots, v^1 - 2\}$  with probability at most  $(v^1 - 1)\gamma_{t+1}$  in round  $t+1$  according to Claim 8 and 9, and that  $(v^1 - 1)\gamma_{t+1} \rightarrow 0$  as  $t \rightarrow \infty$ .  $\square$

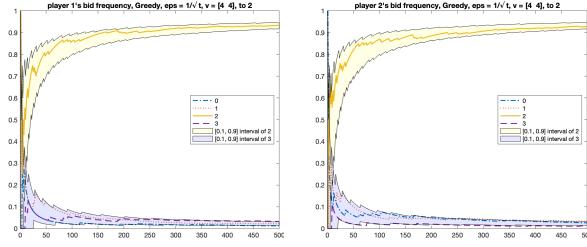
## 5 EXPERIMENTAL RESULTS

Code for the experiments can be found at <https://github.com/tao-l/FPA-mean-based>.

### 5.1 $|M^1| = 2$ : Convergence to Two Equilibria

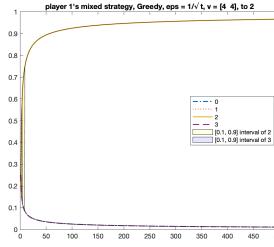
For the case of  $|M^1| = 2$ , we showed in Theorem 5 that any mean-based algorithm must converge to one of the two equilibria where the two players in  $M^1$  bid  $v^1 - 1$  or  $v^1 - 2$ . One may wonder whether there is a theoretical guarantee of which equilibrium will be obtained. We give experimental results to show that, in fact, both equilibria can be obtained under a same randomized mean-based algorithm in different runs. We demonstrate this by the  $\varepsilon_t$ -Greedy algorithm (defined in Example 2). Interestingly, under the same setting, the MWU algorithm always converges to the equilibrium of  $v^1 - 1$ .

In the experiment, we let  $n = |M^1| = 2$ ,  $v^1 = v^2 = V = 4$ .



(a) Player 1's bid frequency

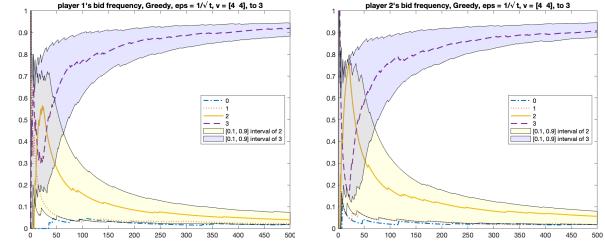
(b) Player 2's bid frequency



(c) Player 1's mixed strategy

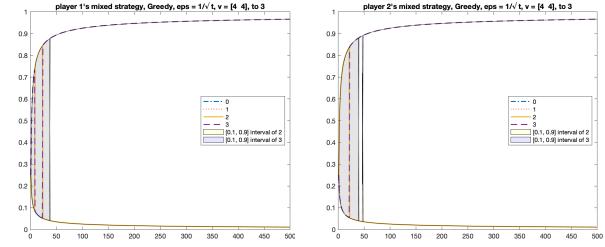
(d) Player 2's mixed strategy

**Figure 1:**  $|M^1| = 2$ ,  $\varepsilon_t$ -Greedy,  $v^1 = v^1 = 4$ , converging to  $v^1 - 2 = 2$ . The four curves in each plot show (a)(b) the changes of frequencies of bids 0, 1, 2, 3 and (c)(d) the changes of mixed strategies, in one simulation. The frequency of 2 approaches 1. The two regions show the [10%, 90%]-confidence intervals of the corresponding curves (the upper is for bid 2, the lower is for bid 3), among all simulations that converge to  $v^1 - 2$ .



(a) Player 1's bid frequency

(b) Player 2's bid frequency



(c) Player 1's mixed strategy

(d) Player 2's mixed strategy

**Figure 2:**  $|M^1| = 2$ ,  $\varepsilon_t$ -Greedy,  $v^1 = v^1 = 4$ , converging to  $v^1 - 1 = 3$ . The four curves in each plot show (a)(b) the changes of frequencies of bids 0, 1, 2, 3 and (c)(d) the changes of mixed strategies, in one simulation. The frequency of 3 approaches 1. The two regions show the [10%, 90%]-confidence intervals of the corresponding curves (the upper is for bid 3, the lower is for bid 2), among all simulations that converge to  $v^1 - 1$ .

## 5.2 $|M^1| = 1$ : Non-Convergence

For the case of  $|M^1| = 1$  we showed that not all mean-based algorithms can converge to equilibrium, using the example of Follow the Leader (Example 7). Here we experimentally demonstrate that such non-convergence phenomena can also happen with more natural (and even no-regret) mean-based algorithms like  $\varepsilon$ -Greedy and MWU.

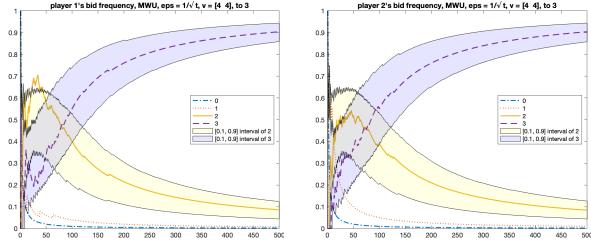
In the experiment we let  $n = 2$ ,  $v^1 = 8$ ,  $v^2 = 6$ . We run  $\varepsilon_t$ -Greedy and MWU both with  $\varepsilon_t = 1/\sqrt{t}$  for  $T = 20000$  rounds.

For  $\varepsilon_t$ -Greedy, Figure 4 shows that the two bidders do not converge to a pure-strategy equilibrium, either in time-average or last-iterate. According to Proposition 3, a pure-strategy equilibrium must have bidder 1 bidding  $v^2 = 6$  and bidder 2 bidding  $v^2 - 1 = 5$ . But figure (b) shows that bidder 2's frequency of bidding 5 does not converge to 1. The frequency oscillates and we do not know whether it will stabilize at some limit less than 1. Looking closer, we see that bidder 2 constantly switches between bids 5 and 3, and bidder 1 switches between 5 and 6. Intuitively, this is because: in the  $\varepsilon_t$ -Greedy algorithm, when bidder 1 bids  $v^2 = 6$  with high probability, she also sometimes (with probability  $\varepsilon_t$ ) chooses bids uniformly at random, in which case the best response for bidder 2 is to bid  $v^2/2 = 3$ ; but after bidder 2 switches to 3, bidder 1 will find it beneficial to lower her bid from 6 to 5; then, bidder 2 will switch to 5 to compete with bidder 1, winning the item with probability 1/2; but then bidder 1 will increase to 6 to outbid bidder 2; ... They hence enter a cycle.

$\varepsilon_t$ -Greedy converges to two equilibria. We run  $\varepsilon_t$ -Greedy with  $\varepsilon_t = \sqrt{1/t}$  for 1000 times. In each time, we run it for  $T = 2000$  rounds. After it finishes, we use the frequency of bids from bidder 1 to determine which equilibrium the algorithm will converge to: if the frequency of bid 2 is above 0.9, we consider it converging to the equilibrium of  $v^1 - 2$ ; if the frequency of bid 3 is above 0.9, we consider it converging to the equilibrium of  $v^1 - 1$ ; if neither happens, we consider it as "not converged yet". Among the 1000 times we found 868 times of  $v^1 - 2$ , 132 times of  $v^1 - 1$ , and 0 times of "not converged yet"; namely, the probability of converging to  $v^1 - 2$  is roughly 87%.

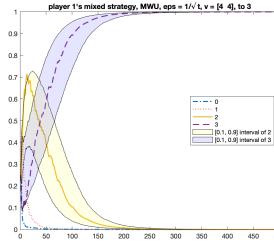
We give two figures of the changes of bid frequencies and mixed strategies of player 1 and 2: Figure 1 is for the case of converging to  $v^1 - 2$ ; Figure 2 is for  $v^1 - 1$ . The x-axis is round number  $t$  and the y-axis is the frequency  $\frac{1}{t} \sum_{s=1}^t \mathbb{I}[b_s^i = b]$  of each bid  $b \in \{0, 1, 2, 3\}$  or the mixed strategy  $x_t^i = (x_t^i(0), x_t^i(1), x_t^i(2), x_t^i(3))$ . For clarity, we only show the first 500 rounds.

MWU always converges to  $v^1 - 1$ . We run MWU with  $\varepsilon_t = \sqrt{1/t}$ . Same as the previous experiment, we run the algorithm for 1000 times and count how many times the algorithm converges to the equilibrium of  $v^1 - 2$  and to  $v^1 - 1$ . We found that, in all 1000 times, MWU converged to  $v^1 - 1$ . Figure 3 shows the changes of bid frequencies and mixed strategies of both players.

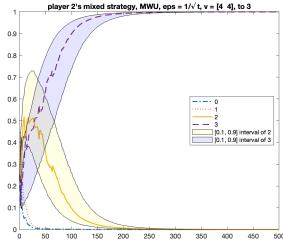


(a) Player 1's bid frequency

(b) Player 2's bid frequency

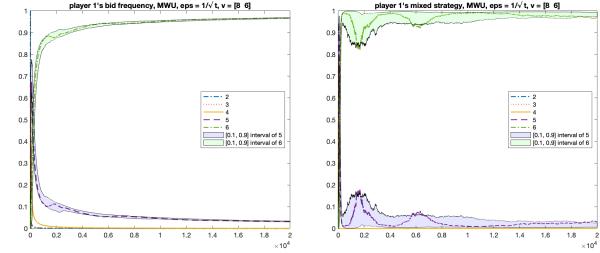


(c) Player 1's mixed strategy



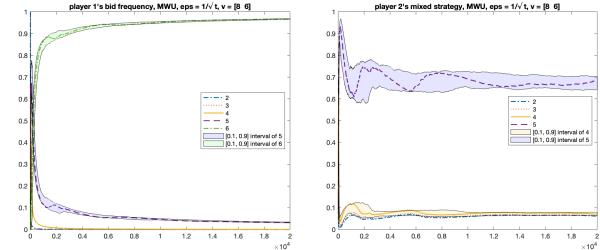
(d) Player 2's mixed strategy

**Figure 3:**  $|M^1| = 2$ , MWU,  $v^1 = v^2 = 4$ , converging to  $v^1 - 1 = 3$ . The four curves in each plot show (a) (b) the changes of frequencies of bids 0, 1, 2, 3 and (c) (d) the changes of mixed strategies, in one simulation. The frequency of 3 approaches 1. The two regions show the [10%, 90%]-confidence intervals of the corresponding curves (the upper is for bid 3, the lower is for bid 2), among 1000 simulations.

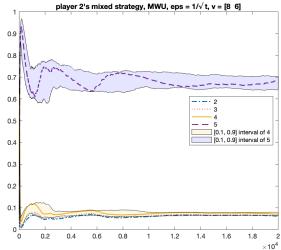


(a) Player 1's bid frequency

(b) Player 2's bid frequency



(c) Player 1's mixed strategy



(d) Player 2's mixed strategy

**Figure 5:**  $|M^1| = 1$ , MWU,  $v^1 = 8, v^2 = 6$ . The curves shown are results from one simulation. The region around the curves are [10%, 90%]-confidence intervals of the curves among 100 simulations.

For MWU, Figure 5 shows that: bidder 1's bid frequency (a) and mixed strategy (c) seem to converge to bidding  $v^2 = 6$ ; but bidder 2's bid frequency (b) and mixed strategy (d) do not seem to converge.

## 6 CONCLUSIONS AND FUTURE DIRECTIONS

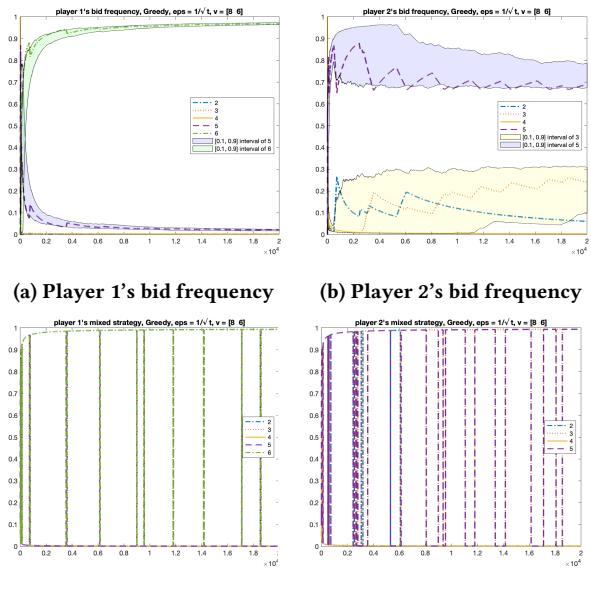
In this work we show that, in repeated fixed-value first price auctions, mean-based learning bidders converge to a Nash equilibrium in the presence of *competition*, in the sense that at least two bidders share the highest value. Without competition, we give non-convergence examples using mean-based algorithms that are not necessarily no-regret. Understanding the convergence property of no-regret algorithms in the absence of competition is a natural and interesting future direction. In fact, [33] shows that some non-mean-based no-regret algorithms do not converge. It is hence open to prove (non-)convergence for mean-based no-regret algorithms.

The convergence result we give is in the limit sense. [46] notes that many no-regret algorithms actually need an exponential time to converge to Nash equilibrium in some iterative-dominance-solvable game. Our theoretical analysis for the first price auction demonstrates a  $T = O(c^{O(v^1)})$  upper bound on the convergence time for the case of  $|M^1| = 3$ . But the convergence time in experiments is significantly shorter. The exact convergence rate remains open.

Analyzing repeated first price auctions where bidders have time-varying values is also a natural, yet possibly challenging, future direction.

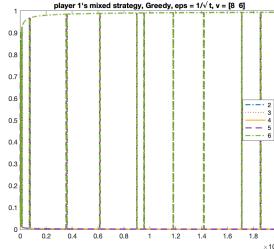
## REFERENCES

- [1] Jacob D Abernethy, Rachel Cummings, Bhuvesh Kumar, Sam Taggart, and Jamie H Morgenstern. 2019. Learning Auctions with Robust Incentive Guarantees. In

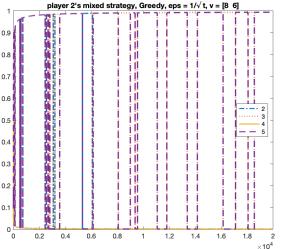


(a) Player 1's bid frequency

(b) Player 2's bid frequency



(c) Player 1's mixed strategy



**Figure 4:**  $|M^1| = 1$ ,  $\epsilon_t$ -Greedy,  $v^1 = 8, v^2 = 6$ . The curves shown are results from one simulation. The region around the curves in (a) (b) are [10%, 90%]-confidence intervals of the curves among 100 simulations.

- Proceedings of the 33rd International Conference on Neural Information Processing Systems (NIPS'19)*. 11587–11597.
- [2] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. 2013. Learning Prices for Repeated Auctions with Strategic Buyers. In *Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS'13)*. 1169–1177.
  - [3] Ashwinkumar Badanidiyuru, Zhe Feng, and Guru Guruganesh. 2021. Learning to Bid in Contextual First Price Auctions. *arXiv preprint arXiv:2109.03173* (2021).
  - [4] Santiago Balseiro, Negin Golrezaei, Mohammad Mahdian, Vahab Mirrokni, and Jon Schneider. 2019. Contextual Bandits with Cross-Learning. In *Advances in Neural Information Processing Systems*, Vol. 32.
  - [5] Martino Banchio and Andrzej Skrzypacz. 2022. Artificial Intelligence and Auction Design. *arXiv preprint arXiv:2202.05947* (2022).
  - [6] Martin Bichler, Maximilian Fichtl, Stefan Heidekrüger, Nils Kohring, and Paul Sutterer. 2021. Learning equilibria in symmetric auction games using artificial neural networks. *Nature Machine Intelligence* 3, 8 (Aug. 2021), 687–695.
  - [7] Avrim Blum and Jason D. Hartline. 2005. Near-Optimal Online Auctions. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '05)*. Society for Industrial and Applied Mathematics, 1156–1163.
  - [8] Mark Braverman, Jieming Mao, Jon Schneider, and Matt Weinberg. 2018. Selling to a No-Regret Buyer. In *Proceedings of the 2018 ACM Conference on Economics and Computation*. ACM, Ithaca NY USA, 523–538.
  - [9] Nicolo Cesa-Bianchi, Claudio Gentile, and Yishay Mansour. 2015. Regret Minimization for Reserve Prices in Second-Price Auctions. *IEEE Transactions on Information Theory* 61, 1 (Jan. 2015), 549–564.
  - [10] Nicolo Cesa-Bianchi and Gabor Lugosi. 2006. *Prediction, Learning, and Games*. Cambridge University Press, Cambridge.
  - [11] Constantinos Daskalakis, Rafael Frongillo, Christos H. Papadimitriou, George Pierrakos, and Gregory Valiant. 2010. On Learning Algorithms for Nash Equilibria. In *Algorithmic Game Theory*. Springer Berlin Heidelberg, 114–125.
  - [12] Constantinos Daskalakis and Ioannis Panageas. 2018. Last-Iterate Convergence: Zero-Sum Games and Constrained Min-Max Optimization. (2018). <http://drops.dagstuhl.de/opus/volltexte/2018/1020/>
  - [13] Xiaotie Deng, Ron Lavi, Tao Lin, Qi Qi, Wenwei WANG, and Xiang Yan. 2020. A Game-Theoretic Analysis of the Empirical Revenue Maximization Algorithm with Endogenous Sampling. In *Advances in Neural Information Processing Systems*, Vol. 33. 5215–5226.
  - [14] Nikhil R. Devanur, Yuval Peres, and Balasubramanian Sivan. 2015. Perfect Bayesian Equilibria in Repeated Sales. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '15)*. Society for Industrial and Applied Mathematics, USA, 983–1002.
  - [15] Guillaume Escamocher, Peter Bro Miltersen, and Santillan-Rodriguez Rocio. 2009. Existence and Computation of Equilibria of First-Price Auctions with Integral Valuations and Bids. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 2 (AAMAS '09)*. 1227–1228.
  - [16] Zhe Feng, Guru Guruganesh, Christopher Liaw, Aranyak Mehta, and Abhishek Sethi. 2021. Convergence Analysis of No-Regret Bidding Algorithms in Repeated Auctions. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*. <https://arxiv.org/pdf/2009.06136.pdf>
  - [17] Zhe Feng, Chara Podimata, and Vasilis Syrgkanis. 2018. Learning to Bid Without Knowing your Value. In *Proceedings of the 2018 ACM Conference on Economics and Computation*. ACM, Ithaca NY USA, 505–522.
  - [18] Gadi Fibich and Arieh Gavious. 2003. Asymmetric First-Price Auctions: A Perturbation Approach. *Mathematics of Operations Research* 28, 4 (2003), 836–852.
  - [19] Aris Filos-Ratsikas, Yiannis Giannakopoulos, Alexandros Hollender, Philip Lazos, and Diogo Poças. 2021. On the Complexity of Equilibrium Computation in First-Price Auctions. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. ACM, Budapest Hungary, 454–476.
  - [20] Dean P. Foster and Rakesh V. Vohra. 1997. Calibrated Learning and Correlated Equilibrium. *Games and Economic Behavior* 21, 1-2 (Oct. 1997), 40–55.
  - [21] Hu Fu and Tae Lin. 2020. Learning Utilities and Equilibria in Non-Truthful Auctions. In *Advances in Neural Information Processing Systems*. 14231–14242.
  - [22] Drew Fudenberg and David K. Levine. 1998. *The theory of learning in games*. Number 2 in MIT Press series on economic learning and social evolution. MIT Press, Cambridge, Mass.
  - [23] Shumei Goke, Gabriel Y Weintraub, Ralph Mastromonaco, and Sam Seljan. 2021. Learning New Auction Format by Bidders in Internet Display Ad Auctions. *arXiv preprint arXiv:2110.13814* (2021).
  - [24] Negin Golrezaei, Adel Javanmard, and Vahab Mirrokni. 2021. Dynamic Incentive-Aware Learning: Robust Pricing in Contextual Auctions. *Operations Research* 69, 1 (Jan. 2021), 297–314.
  - [25] Yanjun Han, Zhengyuan Zhou, Aaron Flores, Erik Ordentlich, and Tsachy Weissman. 2020. Learning to Bid Optimally and Efficiently in Adversarial First-price Auctions. *arXiv preprint arXiv:2007.04568* (2020).
  - [26] Sergiu Hart and Andreu Mas-Colell. 2000. A Simple Adaptive Procedure Leading to Correlated Equilibrium. *Econometrica* 68, 5 (Sept. 2000), 1127–1150.
  - [27] Shlomit Hon-Snir, Dov Monderer, and Aner Sela. 1998. A Learning Approach to Auctions. *Journal of Economic Theory* 82, 1 (Sept. 1998), 65–88.
  - [28] Zhiyi Huang, Jinyan Liu, and Xiangning Wang. 2018. Learning Optimal Reserve Price against Non-Myopic Bidders. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18)*. 2042–2052.
  - [29] Nicole Immorlica, Brendan Lucier, Emmanouil Pountourakis, and Samuel Taggart. 2017. Repeated Sales with Multiple Strategic Buyers. In *Proceedings of the 2017 ACM Conference on Economics and Computation*. ACM, 167–168.
  - [30] Krishnamurthy Iyer, Ramesh Johari, and Mukund Sundararajan. 2014. Mean Field Equilibria of Dynamic Auctions with Learning. *Management Science* 60, 12 (Dec. 2014), 2949–2970.
  - [31] Yash Kanoria and Hamid Nazerzadeh. 2019. Incentive-Compatible Learning of Reserve Prices for Repeated Auctions. In *Companion The 2019 World Wide Web Conference*.
  - [32] Orcun Karaca, Pier Giuseppe Sessa, Anna Leidi, and Maryam Kamgarpour. 2020. No-regret learning from partially observed data in repeated auctions. *IFAC-PapersOnLine* 53, 2 (2020), 14–19.
  - [33] Yoav Kolumbus and Noam Nisan. 2021. Auctions Between Regret-Minimizing Agents. *arXiv preprint arXiv:2110.11855* (2021).
  - [34] Bernard Lebrun. 1996. Existence of an Equilibrium in First Price Auctions. *Economic Theory* 7, 3 (1996), 421–443. Publisher: Springer.
  - [35] Bernard Lebrun. 1999. First Price Auctions in the Asymmetric N Bidder Case. *International Economic Review* 40, 1 (Feb. 1999), 125–142.
  - [36] Eric Maskin and John Riley. 2000. Equilibrium in Sealed High Bid Auctions. *Review of Economic Studies* 67, 3 (July 2000), 439–454.
  - [37] Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. 2017. Cycles in adversarial regularized learning. *arXiv:1709.02738 [cs.GT]*
  - [38] Mehryar Mohri and Andres Muñoz Medina. 2014. Optimal Regret Minimization in Posted-Price Auctions with Strategic Buyers. In *Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14)*.
  - [39] Denis Nekipelov, Vasilis Syrgkanis, and Eva Tardos. 2015. Econometrics for Learning Agents. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation - EC '15*. ACM Press, Portland, Oregon, USA, 1–18.
  - [40] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani (Eds.). 2007. *Algorithmic Game Theory*. Cambridge University Press, Cambridge.
  - [41] Renato Paes Leme, Balasubramanian Sivan, and Yifeng Teng. 2020. Why Do Competitive Markets Converge to First-Price Auctions?. In *Proceedings of The Web Conference 2020*. ACM, Taipei Taiwan, 596–605.
  - [42] Tim Roughgarden. 2016. Lecture #17: No-Regret Dynamics. In *Twenty lectures on algorithmic game theory*. Cambridge University Press, Cambridge ; New York, NY. <https://theory.stanford.edu/~tim/f13/l17.pdf>
  - [43] Zihe Wang, Weiran Shen, and Song Zuo. 2020. Bayesian Nash Equilibrium in First-Price Auction with Discrete Value Distributions. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '20)*. 1458–1466.
  - [44] Jonathan Weed, Vianney Perchet, and Philippe Rigollet. 2016. Online learning in repeated auctions. In *Conference on Learning Theory*. PMLR, 1562–1583.
  - [45] Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. 2021. Linear Last-iterate Convergence in Constrained Saddle-point Optimization. (March 2021). <http://arxiv.org/abs/2006.09517>
  - [46] Jibang Wu, Haifeng Xu, and Fan Yao. 2021. Multi-Agent Learning for Iterative Dominance Elimination: Formal Barriers and New Algorithms. *arXiv preprint arXiv:2111.05486* (2021).

## A MISSING PROOFS FROM SECTION 3

### A.1 Proof of Theorem 5

Suppose  $|M^1| = 2$ . We will prove that, for any sufficiently large integer  $T_b$ , with probability at least  $1 - \exp(-\frac{T_b}{24NV}) - 2 \exp(-\frac{T_b}{1152N^2V^2}) - \frac{6}{e-2}(\frac{48NV}{T_b})^{3e/4}$ , one of following two events must happen:

- $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 2] = 1$ ;
- $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 1] = 1$  and  $\lim_{t \rightarrow \infty} \Pr[b_t^i = v^1 - 1] = 1$ .

And if  $n \geq 3$  and  $v^3 = v^1 - 1$ , only the second event happens. Letting  $T_b \rightarrow \infty$  proves Theorem 5.

We reuse the argument in Section 4.2. Assume  $v^1 \geq 3$ .<sup>8</sup> Recall that we defined  $c = 1 + \frac{1}{12NV}$ ,  $d = \lceil \log_c(8NV) \rceil$ ;  $T_b$  is any integer such that  $\gamma_{T_b} < \frac{1}{12N^2V^2}$  and  $\exp\left(-\frac{(c-1)T_b}{1152N^2V^2}\right) \leq \frac{1}{2}$ ;  $T_0 = 12NV T_b$ ;  $T_{v^1-3} = c(v^1-3)d T_0$ . We defined  $A_{v^1-3}$  to be the event  $\frac{1}{T_{v^1-3}} \sum_{t=1}^{T_{v^1-3}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] \leq \frac{1}{4NV}$ . According to Corollary 13,  $A_{v^1-3}$  holds with probability at least  $1 - \exp(-\frac{T_b}{24NV}) - 2 \exp(-\frac{T_b}{1152N^2V^2})$ . Suppose  $A_{v^1-3}$  holds.

Now we partition the time horizon after  $T_{v^1-3}$  as follows: let  $T_a^0 = T_{v^1-3}$ ,  $T_a^k = C(k+24NV)^2$ ,  $\forall k \geq 0$ , where  $C = \frac{T_{v^1-3}}{(24NV)^2}$ , so that  $T_a^0 = C(0+24NV)^2$ . Denote  $\Gamma_a^{k+1} = [T_a^k + 1, T_a^{k+1}]$ , with  $|\Gamma_a^{k+1}| = T_a^{k+1} - T_a^k$ . (We note that the notations here have different meanings than those in Section 4.3.) We define  $\delta_t = (\frac{1}{t})^{1/8}$ ,  $t \geq 0$ . For each  $k \geq 0$ , we define

$$F_{T_a^k} = \frac{T_a^0}{T_a^k} \frac{1}{4NV} + \sum_{s=0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} \delta_{T_a^s} + \sum_{s=0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} |M^1| V \gamma_{T_a^s}.$$

Let  $A_a^k$  be event

$$A_a^k = \left[ \frac{1}{T_a^k} \sum_{t=1}^{T_a^k} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] \leq F_{T_a^k} \right].$$

We note that  $A_a^0 = A_{v^1-3}$  because  $F_{T_a^0} = \frac{1}{4NV}$ .

In the proof we will always let  $T_b$  to be sufficiently large. This implies that all the times  $T_0, T_{v^1-3}, T_a^0, T_a^k$ , etc., are sufficiently large.

#### A.1.1 Additional Notations, Claims, and Lemmas.

**CLAIM 16.** When  $T_b$  is sufficiently large,

- $F_{T_a^{k+1}} \leq F_{T_a^k} \leq \frac{1}{4NV}$  for every  $k \geq 0$ .
- $\lim_{k \rightarrow \infty} F_{T_a^k} = 0$ .

**PROOF.** Since  $\delta_{T_a^0} \rightarrow 0$  and  $\gamma_{T_a^0} \rightarrow 0$  as  $T_b \rightarrow \infty$ , when  $T_b$  is sufficiently large we have

$$F_{T_a^1} = \frac{T_a^0}{T_a^1} \frac{1}{4NV} + \frac{T_a^1 - T_a^0}{T_a^1} \left( \delta_{T_a^0} + |M^1| V \gamma_{T_a^0} \right) \leq \frac{T_a^0}{T_a^1} \frac{1}{4NV} + \frac{T_a^1 - T_a^0}{T_a^1} \frac{1}{4NV} = \frac{1}{4NV} = F_{T_a^0}.$$

Since  $\delta_{T_a^s}$  and  $\gamma_{T_a^s}$  are both decreasing, we have

$$\begin{aligned} F_{T_a^k} &> \sum_{s=0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} \delta_{T_a^s} + \sum_{s=0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} |M^1| V \gamma_{T_a^s} \\ &\geq \sum_{s=0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} \delta_{T_a^k} + \sum_{s=0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} |M^1| V \gamma_{T_a^k} = \delta_{T_a^k} + |M^1| V \gamma_{T_a^k}. \end{aligned}$$

Thus,

$$F_{T_a^{k+1}} \stackrel{\text{by definition}}{=} \frac{T_a^k}{T_a^{k+1}} F_{T_a^k} + \frac{T_a^{k+1} - T_a^k}{T_a^{k+1}} \left( \delta_{T_a^k} + |M^1| V \gamma_{T_a^k} \right) < \frac{T_a^k}{T_a^{k+1}} F_{T_a^k} + \frac{T_a^{k+1} - T_a^k}{T_a^{k+1}} F_{T_a^k} = F_{T_a^k}.$$

---

<sup>8</sup>If  $v^1 = 1$ , Theorem 5 trivially holds. If  $v^1 = 2$ , we let  $T_{v^1-3} = T_0 = T_b$ ;  $A_{v^1-3}$  holds with probability 1 since  $\frac{1}{T_{v^1-3}} \sum_{t=1}^{T_{v^1-3}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] = 0$ ; the argument for  $v^1 \geq 3$  will still apply.

Then we prove  $\lim_{k \rightarrow \infty} F_{T_a^k} = 0$ . For every  $0 < \varepsilon < \frac{1}{4NV}$ , we can find  $k$  sufficiently large such that  $\delta_{T_a^k} \leq \frac{\varepsilon}{6}$ , and  $\gamma_{T_a^k} \leq \frac{\varepsilon}{6|M^1|V}$ . For any  $l \geq \lceil k/\varepsilon \rceil$ , we have  $\frac{T_a^0}{T_a^l} \leq \frac{T_a^k}{T_a^l} \leq \frac{\varepsilon}{6}$ . Then

$$\begin{aligned} F_{T_a^l} &= \frac{T_a^0}{T_a^l} \frac{1}{4NV} + \sum_{s=0}^{l-1} \frac{T_a^{s+1} - T_a^s}{T_a^l} (\delta_{T_a^s} + |M^1|V\gamma_{T_a^s}) \\ &\leq \frac{\varepsilon}{3} + 2 \sum_{s=0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^l} + \sum_{s=k}^{l-1} \frac{T_a^{s+1} - T_a^s}{T_a^l} (\delta_{T_a^k} + |M^1|V\gamma_{T_a^k}) \\ &\leq \frac{\varepsilon}{3} + 2 \frac{T_a^k}{T_a^l} + \delta_{T_a^k} + |M^1|V\gamma_{T_a^k} \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Since  $F_{T_a^k}$  is non-negative, we have  $\lim_{k \rightarrow \infty} F_{T_a^k} = 0$ .  $\square$

CLAIM 17.  $\sum_{s=0}^{\infty} \exp\left(-\frac{1}{2}|\Gamma_a^{s+1}|\delta_{T_a^s}^2\right) \leq \frac{2}{e-2} \frac{1}{C^{3e/4}} \leq \frac{2}{e-2} \left(\frac{48NV}{T_b}\right)^{3e/4}$ .

PROOF. Recall that  $|\Gamma_a^{s+1}| = T_a^{s+1} - T_a^s$ ,  $\delta_{T_a^s}^2 = (\frac{1}{T_a^s})^{1/8}$ , and  $T_a^s = C(s + 24NV)^2$ . Hence,

$$\begin{aligned} \sum_{s=0}^{\infty} \exp\left(-\frac{1}{2}|\Gamma_a^{s+1}|\delta_{T_a^s}^2\right) &= \sum_{s=0}^{\infty} \exp\left(-\frac{1}{2}(T_a^{s+1} - T_a^s)\left(\frac{1}{T_a^s}\right)^{1/4}\right) \\ &= \sum_{s=0}^{\infty} \exp\left(-\frac{1}{2}C(2(s + 24NV) + 1)\left(\frac{1}{C(s + 24NV)^2}\right)^{1/4}\right) \\ &\leq \sum_{s=0}^{\infty} \exp\left(-C^{3/4}(s + 24NV)\left(\frac{1}{s + 24NV}\right)^{1/2}\right) \\ &= \sum_{s=0}^{\infty} \exp\left(-C^{3/4}\sqrt{s + 24NV}\right) \\ &\leq \sum_{x=2}^{\infty} \exp\left(-C^{3/4}\sqrt{x}\right) \\ &\leq \int_{x=1}^{\infty} \exp\left(-C^{3/4}\sqrt{x}\right) dx \\ (\text{using } e^x \geq x^e \text{ for } x \geq 0) &\leq \int_{x=1}^{\infty} \frac{1}{(C^{3/4}\sqrt{x})^e} dx = \frac{1}{C^{3e/4}} \cdot \frac{2}{e-2}. \end{aligned}$$

Substituting  $C = \frac{T_{a^1-3}}{(24NV)^2} = \frac{c^{(v^1-3)d}12NV T_b}{(24NV)^2} \geq \frac{12NV T_b}{(24NV)^2} = \frac{T_b}{48NV}$  proves the claim.  $\square$

FACT 18.  $\frac{T_a^k}{T_a^{k+1}} \geq 1 - \frac{2}{k+24NV}$ .

PROOF. By definition,

$$\frac{T_a^k}{T_a^{k+1}} = \frac{(k+24NV)^2}{(k+24NV+1)^2} = 1 - \frac{2(k+24NV)+1}{(k+24NV+1)^2} \geq 1 - \frac{2}{k+24NV+1} \geq 1 - \frac{2}{k+24NV}.$$

$\square$

CLAIM 19. When  $A_a^k$  holds, we have, for every  $t \in \Gamma_a^{k+1} = [T_a^k + 1, T_a^{k+1}]$ ,

$$\frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq F_{T_a^k} + \frac{2}{k+24NV} \leq \frac{1}{2NV} - 2\gamma_t.$$

PROOF. When  $A_a^k$  holds, for every  $t \in \Gamma_a^{k+1}$ ,

$$\begin{aligned} \frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] &\leq \frac{1}{t-1} \left( T_a^k F_{T_a^k} + (t-1-T_a^k) \right) \\ (\text{since } T_a^k \leq t-1 \leq T_a^{k+1}) &\leq F_{T_a^k} + \frac{T_a^{k+1} - T_a^k}{T_a^{k+1}} \\ (\text{by Fact 18}) &\leq F_{T_a^k} + \frac{2}{k+24NV}. \end{aligned}$$

Since  $F_{T_a^k} \leq \frac{1}{4NV}$  by Claim 16 and  $\gamma_t \leq \frac{1}{12N^2V^2}$  by assumption, the above expression is further bounded by  $\frac{1}{4NV} + \frac{2}{k+24NV} \leq \frac{1}{4NV} + \frac{2}{24NV} = \frac{1}{3NV} \leq \frac{1}{2NV} - 2\gamma_t$ .  $\square$

LEMMA 20. For every  $k \geq 0$ ,  $\Pr[A_a^{k+1} | A_a^k] \geq 1 - \exp\left(-\frac{1}{2}|\Gamma_a^{k+1}|\delta_{T_a^k}^2\right)$ .

PROOF. Given  $A_a^k$ , according to Claim 19, it holds that for every  $t \in \Gamma_a^{k+1}$ ,  $\frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq \frac{1}{2NV} - 2\gamma_t$ . Then according to Claim 8, for any history  $H_{t-1}$ ,

$$\Pr[\exists i \in M^1, b_t^i \leq v^1 - 3 | H_{t-1}, A_a^k] \leq |M^1|V\gamma_t.$$

Let  $Z_t = \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] - |M^1|V\gamma_t$  and let  $X_t = \sum_{s=T_a^k+1}^t Z_s$ . We have  $\mathbb{E}[Z_t | H_{t-1}, A_a^k] \leq 0$ . Therefore, the sequence  $X_{T_a^k+1}, X_{T_a^k+2}, \dots, X_{T_a^{k+1}}$  is a supermartingale (with respect to the sequence of history  $H_{T_a^k}, H_{T_a^k+1}, \dots, H_{T_a^{k+1}-1}$ ). By Azuma's inequality, for any  $\Delta > 0$ , we have

$$\Pr\left[\sum_{t \in \Gamma_a^{k+1}} Z_t \geq \Delta \mid A_a^k\right] \leq \exp\left(-\frac{\Delta^2}{2|\Gamma_a^{k+1}|}\right).$$

Let  $\Delta = |\Gamma_a^{k+1}|\delta_{T_a^k}$ . Then with probability at least  $1 - \exp(-\frac{1}{2}|\Gamma_a^{k+1}|\delta_{T_a^k}^2)$ , we get  $\sum_{t \in \Gamma_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] < \Delta + |M^1|V \sum_{t \in \Gamma_a^{k+1}} \gamma_t \leq |\Gamma_a^{k+1}|\delta_{T_a^k} + |M^1|V|\Gamma_a^{k+1}|\gamma_{T_a^k}$ , which implies

$$\begin{aligned} &\frac{1}{T_a^{k+1}} \sum_{t=1}^{T_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] \\ &= \frac{1}{T_a^{k+1}} \left( \sum_{t=1}^{T_a^k} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] + \sum_{t \in \Gamma_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] \right) \\ &\leq \frac{1}{T_a^{k+1}} \left( T_a^k F_{T_a^k} + |\Gamma_a^{k+1}|\delta_{T_a^k} + |M^1|V|\Gamma_a^{k+1}|\gamma_{T_a^k} \right) \\ (\text{by definition}) &= F_{T_a^{k+1}} \end{aligned}$$

and thus  $A_a^{k+1}$  holds.  $\square$

Denote by  $f_t^i(b)$  the frequency of bid  $b$  in the first  $t$  rounds for bidder  $i$ :

$$f_t^i(b) = \frac{1}{t} \sum_{s=1}^t \mathbb{I}[b_s^i = b].$$

Let  $f_t^i(0 : v^1 - 3) = \frac{1}{t} \sum_{s=1}^t \mathbb{I}[b_s^i \leq v^1 - 3]$ .

CLAIM 21. If the history  $H_{t-1}$  satisfies  $f_{t-1}^i(v^1 - 1) > 2(X + V\gamma_t)$  and  $\frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq X$  for some  $X \in [0, 1]$ , then we have  $\Pr[b_t^{i'} = v^1 - 2 | H_{t-1}] \leq \gamma_t$  for the other  $i' \neq i \in M^1$ .

PROOF. Consider  $\alpha_{t-1}^{i'}(v^1 - 1)$  and  $\alpha_{t-1}^{i'}(v^1 - 2)$ . On the one hand,

$$\alpha_{t-1}^{i'}(v^1 - 1) = 1 \times (1 - f_{t-1}^i(v^1 - 1)) + \frac{1}{2} \times f_{t-1}^i(v^1 - 1) = 1 - \frac{1}{2}f_{t-1}^i(v^1 - 1). \quad (7)$$

On the other hand, since having more bidders with bids no larger than  $v^1 - 2$  only decreases the utility of a bidder who bids  $v^1 - 2$ , we can upper bound  $\alpha_{t-1}^{i'}(v^1 - 2)$  by

$$\begin{aligned}\alpha_{t-1}^{i'}(v^1 - 2) &\leq 2 \times f_{t-1}^i(0 : v^1 - 3) + 1 \times (1 - f_{t-1}^i(v^1 - 1)) - f_{t-1}^i(0 : v^1 - 3) \\ &= 1 - f_{t-1}^i(v^1 - 1) + f_{t-1}^i(0 : v^1 - 3) \\ &\leq 1 - f_{t-1}^i(v^1 - 1) + X,\end{aligned}\tag{8}$$

where the last inequality holds because  $f_{t-1}^i(0 : v^1 - 3) \leq \frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq X$ . Combining (7) and (8), we get

$$\alpha_{t-1}^{i'}(v^1 - 1) - \alpha_{t-1}^{i'}(v^1 - 2) \geq (1 - \frac{1}{2} f_{t-1}^i) - (1 - f_{t-1}^i + X) = \frac{1}{2} f_{t-1}^i(v^1 - 1) - X > V\gamma_t.$$

This implies  $\Pr[b_t^{i'} = v^1 - 2 | H_{t-1}] \leq \gamma_t$  according to the mean-based property.  $\square$

**A.1.2 Proof of the General Case.** We consider  $k = 0, 1, \dots$  to  $\infty$ . For each  $k$ , we suppose  $A_a^0, A_a^1, \dots, A_a^k$  hold, which happens with probability at least  $1 - \sum_{s=0}^{k-1} \exp(-\frac{1}{2} |\Gamma_a^{s+1}| \delta_{T_a^s}^2)$  according to Lemma 20, given that  $A_a^0 = A_{v^1-3}$  already held. The proof is divided into two cases based on  $f_{T_a^k}^i(v^1 - 1)$ .

*Case 1: For all  $k \geq 0$ ,  $f_{T_a^k}^i(v^1 - 1) \leq 16(F_{T_a^k} + \frac{2}{k+24NV} + V\gamma_{T_a^k})$  for both  $i \in M^1$ .*

We argue that the two bidders in  $M^1$  converge to playing  $v^1 - 2$  in this case.

According to Lemma 20, all events  $A_a^0, A_a^1, \dots, A_a^k, \dots$  happen with probability at least  $1 - \sum_{k=0}^{\infty} \exp(-\frac{1}{2} |\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ . Claim 19 and Claim 16 then imply that, for both  $i \in M^1$ ,

$$\lim_{t \rightarrow \infty} f_t^i(0 : v^1 - 3) \leq \lim_{k \rightarrow \infty} \left( F_{T_a^k} + \frac{2}{k+24NV} \right) = 0.$$

Because for every  $t \in \Gamma_a^{k+1} = [T_a^k + 1, T_a^{k+1}]$  we have  $f_t^i(v^1 - 1) \leq \frac{T_a^{k+1}}{t} f_{T_a^k}^i(v^1 - 1) \leq \frac{T_a^{k+1}}{T_a^k} f_{T_a^k}^i(v^1 - 1) \leq 2f_{T_a^k}^i(v^1 - 1)$  and by condition  $f_{T_a^k}^i(v^1 - 1) \rightarrow 0$  as  $k \rightarrow \infty$ , we have

$$\lim_{t \rightarrow \infty} f_t^i(v^1 - 1) = 0.$$

Therefore,

$$\lim_{t \rightarrow \infty} f_t^i(v^1 - 2) = \lim_{t \rightarrow \infty} 1 - f_t^i(0 : v^1 - 3) - f_t^i(v^1 - 1) = 1,$$

which implies

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 2] = 1.$$

*Case 2: There exists  $k \geq 0$  such that  $f_{T_a^k}^i(v^1 - 1) > 16(F_{T_a^k} + \frac{2}{k+24NV} + V\gamma_{T_a^k})$  for some  $i \in M^1$ .*

If this case happens, we argue that the two bidders in  $M^1$  converge to playing  $v^1 - 1$ .

We first prove that, after  $\ell = k + 24NV$  periods (i.e., at time  $T_a^{k+\ell}$ ), the frequency of  $v^1 - 1$  for both bidders in  $M^1$  is greater than  $4(F_{T_a^{k+\ell}} + \frac{2}{(k+\ell)+24NV} + V\gamma_{T_a^{k+\ell}})$ , with high probability.

**LEMMA 22.** Suppose that, at time  $T_a^k$ ,  $A_a^k$  holds and for some  $i \in M^1$ ,  $f_{T_a^k}^i(v^1 - 1) > 16(F_{T_a^k} + \frac{2}{k+24NV} + V\gamma_{T_a^k})$  holds. Then, with probability at least  $1 - 2 \sum_{j=k}^{k+\ell-1} \exp(-\frac{1}{2} |\Gamma_a^{j+1}| \delta_{T_a^j}^2)$ , the following events happen at time  $T_a^{k+\ell}$ , where  $\ell = k + 24NV$ :

- $A_a^{k+\ell}$ ;
- For both  $i \in M^1$ ,  $f_{T_a^{k+\ell}}^i(v^1 - 1) > 4(F_{T_a^{k+\ell}} + \frac{2}{(k+\ell)+24NV} + V\gamma_{T_a^{k+\ell}})$ .

**PROOF.** We prove by an induction from  $j = k$  to  $k + \ell - 1$ . Given  $A_a^j$ ,  $A_a^{j+1}$  happens with probability at least  $1 - \exp(-\frac{1}{2} |\Gamma_a^{j+1}| \delta_{T_a^j}^2)$  according to Lemma 20. Hence, with probability at least  $1 - \sum_{j=k}^{k+\ell-1} \exp(-\frac{1}{2} |\Gamma_a^{j+1}| \delta_{T_a^j}^2)$ , all events  $A_a^k, A_a^{k+1}, \dots, A_a^{k+\ell}$  happen.

Now we consider the second event. For all  $t \in \Gamma_a^{j+1}$ , noticing that  $\frac{T_a^k}{t-1} \geq \frac{T_a^k}{T_a^{k+\ell}} \geq \frac{T_a^k}{T_a^{k+\ell}} = \frac{(k+24NV)^2}{(2(k+24NV))^2} = \frac{1}{4}$ , we have

$$\begin{aligned} f_{t-1}^i(v^1 - 1) &\geq \frac{T_a^k}{t-1} f_{T_a^k}^i(v^1 - 1) \geq \frac{1}{4} f_{T_a^k}^i(v^1 - 1) \\ &\stackrel{\text{(by condition)}}{>} 4(F_{T_a^k} + \frac{2}{k+24NV} + V\gamma_{T_a^k}) \\ &\stackrel{\text{($F_{T_a^k}$ and $\gamma_{T_a^k}$ are decreasing in $k$)}}{\geq} 4(F_{T_a^j} + \frac{2}{j+24NV} + V\gamma_{T_a^j}). \end{aligned} \quad (9)$$

According to Claim 19, given  $A_a^j$  we have  $\frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq F_{T_a^j} + \frac{2}{j+24NV} \leq \frac{1}{2NV} - 2\gamma_t$ . Using Claim 21 with  $X = F_{T_a^j} + \frac{2}{j+24NV}$ , we have, for bidder  $i' \neq i, i' \in M^1$ ,  $\Pr[b_t^{i'} = v^1 - 2 \mid H_{t-1}] \leq \gamma_t$ . By Claim 8,  $\Pr[b_t^{i'} \leq v^1 - 3 \mid H_{t-1}] \leq (V-1)\gamma_t$ . Combining the two, we get

$$\Pr[b_t^{i'} = v^1 - 1 \mid H_{t-1}] \geq 1 - V\gamma_t.$$

Let  $\Delta = |\Gamma_a^{k+1}| \delta_{T_a^k}$ . Similar to the proof of Lemma 20, we can use Azuma's inequality to argue that, with probability at least  $1 - \exp(-\frac{1}{2}|\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ , it holds that

$$\sum_{t \in \Gamma_a^{j+1}} \mathbb{I}[b_t^{i'} = v^1 - 1] \geq \sum_{t \in \Gamma_a^{j+1}} (1 - V\gamma_t - \delta_{T_a^j}) \geq |\Gamma_a^{j+1}|(1 - V\gamma_{T_a^j} - \delta_{T_a^j}).$$

An induction shows that, with probability at least  $1 - \sum_{j=k}^{k+\ell-1} \exp\left(-\frac{1}{2}|\Gamma_a^{j+1}| \delta_{T_a^j}^2\right)$ ,  $\sum_{t \in \Gamma_a^{j+1}} \mathbb{I}[b_t^{i'} = v^1 - 1] \geq |\Gamma_a^{j+1}|(1 - V\gamma_{T_a^j} - \delta_{T_a^j})$  holds for all  $j \in \{k, \dots, k+\ell-1\}$ . Therefore,

$$\begin{aligned} f_{T_a^{k+\ell}}^{i'}(v^1 - 1) &\geq \frac{1}{T_a^{k+\ell}} \left( 0 + \sum_{t \in \Gamma_a^{k+1} \cup \dots \cup \Gamma_a^{k+\ell}} \mathbb{I}[b_t^{i'} = v^1 - 1] \right) \\ &\geq \frac{1}{T_a^{k+\ell}} \left( |\Gamma_a^{k+1}|(1 - V\gamma_{T_a^k} - \delta_{T_a^k}) + \dots + |\Gamma_a^{k+\ell}|(1 - V\gamma_{T_a^{k+\ell-1}} - \delta_{T_a^{k+\ell-1}}) \right) \\ &\geq \frac{1}{T_a^{k+\ell}} \left( (|\Gamma_a^{k+1}| + \dots + |\Gamma_a^{k+\ell}|) \cdot (1 - V\gamma_{T_a^k} - \delta_{T_a^k}) \right) \\ &= \frac{T_a^{k+\ell} - T_a^k}{T_a^{k+\ell}} (1 - V\gamma_{T_a^k} - \delta_{T_a^k}) \\ &= \frac{4(k+24NV)^2 - (k+24NV)^2}{4(k+24NV)^2} (1 - V\gamma_{T_a^k} - \delta_{T_a^k}) \\ &= \frac{3}{4} (1 - V\gamma_{T_a^k} - \delta_{T_a^k}) \\ &\stackrel{\text{(assuming $T_b$ is large enough)}}{>} 4 \left( F_{T_a^{k+\ell}} + \frac{2}{(k+\ell)+24NV} + V\gamma_{T_a^{k+\ell}} \right). \end{aligned}$$

This proves the claim for  $i' \in M^1$ . The claim for  $i \in M^1$  follows from (9) and the fact that  $F_{T_a^k}$  and  $\gamma_{T_a^k}$  are decreasing in  $k$ .  $\square$

We denote by  $k_0 = k + \ell$  the time period at which  $f_{T_a^{k_0}}^i(v^1 - 1) > 4(F_{T_a^{k_0}} + \frac{2}{k_0+24NV} + V\gamma_{T_a^{k_0}})$  for both  $i \in M^1$ . We continuing the analysis for each period  $k \geq k_0$ . Define sequence  $(G_{T_a^k})$ :

$$G_{T_a^k} = \frac{T_a^{k_0}}{T_a^k} \cdot 4 \left( F_{T_a^{k_0}} + \frac{2}{k_0+24NV} + V\gamma_{T_a^{k_0}} \right) + \sum_{s=k_0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} (1 - V\gamma_{T_a^s} - \delta_{T_a^s}), \quad \text{for } k \geq k_0,$$

where we recall that  $\delta_t = (\frac{1}{t})^{1/8}$ . We note that  $f_{T_a^{k_0}}^i(v^1 - 1) > G_{T_a^{k_0}} = 4(F_{T_a^{k_0}} + \frac{2}{k_0+24NV} + V\gamma_{T_a^{k_0}})$ .

**CLAIM 23.** *When  $T_b$  is sufficiently large,*

- $G_{T_a^k} \geq 4(F_{T_a^{k_0}} + \frac{2}{k_0+24NV} + V\gamma_{T_a^{k_0}})$  for every  $k \geq k_0$ .
- $\lim_{k \rightarrow \infty} G_{T_a^k} = 1$ .

**PROOF.** Since  $1 - V\gamma_{T_a^s} - \delta_{T_a^s} \rightarrow 1$  as  $T_b \rightarrow \infty$ , for sufficiently large  $T_b$  we have  $1 - V\gamma_{T_a^s} - \delta_{T_a^s} \geq 4(F_{T_a^{k_0}} + \frac{2}{k_0+24NV} + V\gamma_{T_a^{k_0}})$  and hence  $G_{T_a^k} \geq 4(F_{T_a^{k_0}} + \frac{2}{k_0+24NV} + V\gamma_{T_a^{k_0}})$ .

Now we prove  $\lim_{k \rightarrow \infty} G_{T_a^k} = 1$ . Consider the second term in  $G_{T_a^k}$ ,  $\sum_{s=k_0}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} (1 - V\gamma_{T_a^s} - \delta_{T_a^s})$ . Since

$$\sum_{s=\sqrt{k}}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} = \sum_{s=\sqrt{k}}^{k-1} \frac{2(s + 24NV) + 1}{(k + 24NV)^2} = \frac{(k + \sqrt{k} + 48NV)(k - \sqrt{k})}{(k + 24NV)^2} \rightarrow 1$$

and  $1 - V\gamma_{T_a^k} - \delta_{T_a^k} \rightarrow 1$  as  $k \rightarrow \infty$ , for any  $\varepsilon > 0$  we can always find  $K \geq k_0$  such that  $\sum_{s=\sqrt{k}}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} \geq 1 - \varepsilon/2$  for every  $k \geq K$  and  $1 - V\gamma_{T_a^s} - \delta_{T_a^s} \geq 1 - \varepsilon/2$  for every  $s \geq \sqrt{k}$ . Hence,

$$G_{T_a^k} \geq \sum_{s=\sqrt{k}}^{k-1} \frac{T_a^{s+1} - T_a^s}{T_a^k} (1 - V\gamma_{T_a^s} - \delta_{T_a^s}) \geq (1 - \varepsilon/2)(1 - \varepsilon/2) \geq 1 - \varepsilon,$$

In addition,  $G_{T_a^k} \leq 1$  when  $T_b$  is sufficiently large. Therefore  $\lim_{k \rightarrow \infty} G_{T_a^k} = 1$ .  $\square$

LEMMA 24. Fix any  $k$ . Suppose  $A_a^k$  holds and  $f_{T_a^k}(v^1 - 1) > G_{T_a^k}$  holds for both  $i \in M^1$ . Then, the following four events happen with probability at least  $1 - 3 \exp(-\frac{1}{2}|\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ :

- $A_a^{k+1}$ ;
- $f_{T_a^{k+1}}^i(v^1 - 1) > G_{T_a^{k+1}}$  holds for both  $i \in M^1$ ;
- $f_t^i(v^1 - 1) > (1 - \frac{2}{k+24NV})G_{T_a^k}$  holds for both  $i \in M^1$ , for any  $t \in \Gamma_a^{t+1}$ .
- $x_t^i(v^1 - 1) = \Pr[b_t^i = v^1 - 1 \mid H_{t-1}] \geq 1 - V\gamma_t$  for both  $i \in M^1$ , for any  $t \in \Gamma_a^{k+1}$ .

PROOF. By Lemma 20,  $A_a^{k+1}$  holds with probability at least  $1 - \exp(-\frac{1}{2}|\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ . Now we consider the second event. For every  $t \in \Gamma_a^{k+1}$ , we have

$$\begin{aligned} f_{t-1}^i(v^i - 1) &\geq \frac{T_a^k}{T_a^{k+1}} f_{T_a^k}(v^i - 1) \\ (\text{by condition}) &> \frac{T_a^k}{T_a^{k+1}} G_{T_a^k} \\ (\text{by Fact 18}) &\geq \left(1 - \frac{2}{k+24NV}\right) G_{T_a^k} \\ &\geq \frac{1}{2} G_{T_a^k} \\ (\text{by Claim 23}) &\geq 2 \left(F_{T_a^k} + \frac{2}{k+24NV} + V\gamma_{T_a^k}\right). \end{aligned} \tag{10}$$

In addition, according to Claim 19  $A_a^k$  implies

$$\frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq F_{T_a^k} + \frac{2}{k+24NV} \leq \frac{1}{2NV} - 2\gamma_t.$$

Using Claim 21 with  $X = F_{T_a^k} + \frac{2}{k+24NV}$ , we get  $\Pr[b_t^i = v^1 - 2 \mid H_{t-1}] \leq \gamma_t$ . Additionally, by Claim 8 we have  $\Pr[b_t^i \leq v^1 - 3 \mid H_{t-1}] \leq (V-1)\gamma_t$ . Therefore,

$$\Pr[b_t^i = v^1 - 1 \mid H_{t-1}] \geq 1 - V\gamma_t. \tag{11}$$

Using Azuma's inequality with  $\Delta = |\Gamma_a^{k+1}| \delta_{T_a^k}$ , we have with probability at least  $1 - \exp(-\frac{1}{2}|\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ ,

$$\sum_{t \in \Gamma_a^{k+1}} \mathbb{I}[b_t^i = v^1 - 1] > \sum_{t \in \Gamma_a^{k+1}} (1 - V\gamma_t - \delta_{T_a^k}) \geq |\Gamma_a^{k+1}|(1 - V\gamma_{T_a^k} - \delta_{T_a^k}).$$

It follows that

$$f_{T_a^{k+1}}^i(v^1 - 1) > \frac{1}{T_a^{k+1}} \left(T_a^k G_{T_a^k} + |\Gamma_a^{k+1}|(1 - V\gamma_{T_a^k} - \delta_{T_a^k})\right) = G_{T_a^{k+1}}$$

by definition.

Using a union bound, the first event  $A_a^{k+1}$  and the second event that  $f_{T_a^{k+1}}^i(v^1 - 1) > G_{T_a^{k+1}}$  holds for both  $i \in M^1$  happen with probability at least  $1 - 3 \exp(-\frac{1}{2}|\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ . The third event is given by (10) and the forth event is given by (11).  $\square$

We use Lemma 24 from  $k$  to  $\infty$ ; from its third and fourth events, combined with Claim 23, we get

$$\lim_{t \rightarrow \infty} f_t^i(v^1 - 1) \geq \lim_{k \rightarrow \infty} \left(1 - \frac{2}{k + 24NV}\right) G_{T_a^k} = 1 \text{ and } \lim_{t \rightarrow \infty} \mathbf{x}_t^i = \mathbf{1}_{v^1 - 1},$$

which happens with probability at least  $1 - 3 \sum_{k=0}^{\infty} \exp(-\frac{1}{2} |\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ . This concludes the analysis for Case 2.

Combining Case 1 and Case 2, we have that either  $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 2] = 1$  happens or  $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\forall i \in M^1, b_s^i = v^1 - 1] = 1$  happens (in which case we also have  $\lim_{t \rightarrow \infty} \mathbf{x}_t^i = \mathbf{1}_{v^1 - 1}$ ) with overall probability at least  $1 - \exp(-\frac{T_b}{24NV}) - 2 \exp(-\frac{T_b}{1152N^2V^2}) - 3 \sum_{k=0}^{\infty} \exp(-\frac{1}{2} |\Gamma_a^{k+1}| \delta_{T_a^k}^2)$ . Using Claim 17 concludes the proof.

#### A.1.3 The special case of $v^3 = v^1 - 1$ .

CLAIM 25. Given  $f_t^i(v^1 - 2) \geq 1 - \frac{1}{4+2NV}$  for all  $i \in M^1$ , we have  $\Pr[b_t^3 = v^1 - 2 | H_{t-1}] \geq 1 - V\gamma_t$ .

PROOF. If  $f_t^i(v^1 - 2) \geq 1 - \varepsilon$ ,  $\varepsilon = \frac{1}{4+2NV}$ , for all  $i \in M^1$  then the frequency of the maximum bid to be  $v^1 - 2$  is at least  $1 - 2\varepsilon$ , which implies

$$\alpha_{t-1}^3(v^1 - 2) \geq 2 \frac{1}{N}(1 - 2\varepsilon).$$

For any  $b \leq v^1 - 3$ ,

$$\alpha_{t-1}^3(b) \leq V2\varepsilon.$$

Since  $\gamma_t < \frac{1}{12N^2V^2} < \frac{1}{NV}$ , we have  $\alpha_{t-1}^3(v^1 - 2) - \alpha_{t-1}^3(b) \geq 2 \frac{1}{N}(1 - 2\varepsilon) - 2V\varepsilon > V\gamma_t$ , which implies, according to mean-based property,

$$\Pr[b_t^3 = v^1 - 2] \geq 1 - V\gamma_t. \quad \square$$

CLAIM 26. If history  $H_{t-1}$  satisfies  $f_{t-1}^i(v^1 - 2) \geq \frac{9}{10}$  for  $i \in M^1$  and  $f_{t-1}^3(v^1 - 2) \geq \frac{9}{10}$ , then  $\Pr[b_{t-1}^{i'} = v^1 - 2 | H_{t-1}] \leq \gamma_t$ .

PROOF. If  $f_{t-1}^i(v^1 - 2) \geq \frac{9}{10}$  for  $i \in M^1$  and  $f_{t-1}^3(v^1 - 2) \geq \frac{9}{10}$ , then we have

$$\frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[|\{i \notin M^1 : b_s^i = v^1 - 2\}| \geq 2] \geq 1 - 2 \times \frac{1}{10} = \frac{4}{5},$$

$$P_{t-1}^{i'}(0 : v^1 - 3) \leq 1 - f_{t-1}^3(v^1 - 2) \leq \frac{1}{10}.$$

Recall that  $P_t^i(k) = \frac{1}{t} \sum_{s=1}^t \mathbb{I}[\max_{j \neq i} b_s^j = k]$ . By  $P_t^i(0 : k)$  we mean  $\sum_{\ell=0}^k P_t^i(\ell)$ . And we can calculate

$$\begin{aligned} & \alpha_{t-1}^{i'}(v^1 - 1) - \alpha_{t-1}^{i'}(v^1 - 2) \\ & \geq P_{t-1}^{i'}(v^1 - 1) \times (\frac{1}{2} - 0) + \frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[|\{i \notin M^1 : b_s^i = v^1 - 2\}| \geq 2] \times (1 - \frac{2}{3}) \\ & \quad + P_{t-1}^{i'}(0 : v^1 - 3) \times (1 - 2) \\ & \geq 0 + \frac{1}{3} \times \frac{4}{5} - \frac{1}{10} = \frac{1}{6} \\ & > V\gamma_t, \end{aligned}$$

which implies  $\Pr[b_{t-1}^{i'} = v^1 - 2 | H_{t-1}] \leq \gamma_t$  according to mean-based property.  $\square$

We only provide a proof sketch here; the formal proof is complicated but similar to the above proof for Case 2 and hence omitted. We prove by contradiction. Suppose Case 1 happens, that is, at each time step  $T_a^k$  the frequency of  $v^1 - 1$  for both bidders  $i \in M^1$ ,  $f_{T_a^k}^i(v^1 - 1)$ , is upper bounded by the threshold  $16(F_{T_a^k} + \frac{2}{k+24NV} + V\gamma_{T_a^k})$ , which approaches 0 as  $k \rightarrow \infty$ . Assuming  $A_a^0, \dots, A_a^k$  happen (which happens with high probability), the frequency of  $0 : v^1 - 3$  is also low. Thus,  $f_t^i(v^1 - 2)$  must be close to 1. Then, according to Claim 25, bidder 3 will bid  $v^1 - 2$  with high probability. Using Azuma's inequality, with high probability, the frequency of bidder 3 bidding  $v^1 - 2$  in all future periods will be approximately 1, which increases  $f_t^3(v^1 - 2)$  to be close to 1 after several periods. Then, according to Claim 26, bidder  $i \in M^1$  will switch to bid  $v^1 - 1$ . After several periods, the frequency  $f_{T_a^k}^i(v^1 - 1)$  will exceed  $16(F_{T_a^k} + \frac{2}{k+24NV} + V\gamma_{T_a^k})$  and thus satisfy Case 2. This leads to a contradiction.

## A.2 Proof of Proposition 6

We consider a simple case where there are only two bidders with the same type  $v^1 = v^2 = 3$ . Let  $V = 3$ . The set of possible bids is  $\mathcal{B}^1 = \mathcal{B}^2 = \{0, 1, 2\}$ . Denote  $f_t^i(b) = \frac{1}{t} \sum_{s=1}^t \mathbb{I}[b_s^i = b]$  the frequency of bidder  $i$ 's bid in the first  $t$  rounds.

CLAIM 27. For  $i \in \{1, 2\}$ ,  $\alpha_t^i(1) - \alpha_t^i(2) = f_t^{3-i}(0) - \frac{f_t^{3-i}(2)}{2}$  and  $\alpha_t^i(1) - \alpha_t^i(0) = f_t^{3-i}(1) + \frac{f_t^{3-i}(0)}{2}$ .

PROOF. We can express  $\alpha_t^i(b)$  using the frequencies as the following.

$$\begin{aligned}\alpha_t^i(0) &= \frac{3f_t^{3-i}(0)}{2}; \\ \alpha_t^i(1) &= f_t^{3-i}(1) + 2f_t^{3-i}(0) = 1 + f_t^{3-i}(0) - f_t^{3-i}(2); \\ \alpha_t^i(2) &= \frac{f_t^{3-i}(2)}{2} + 1 - f_t^{3-i}(2).\end{aligned}$$

Then the claim follows from direct calculation.  $\square$

We construct a  $\gamma_t$ -mean-based algorithm Alg (Algorithm 1) with  $\gamma_t = O(\frac{1}{t^{1/4}})$  such that, with constant probability,  $\lim_{t \rightarrow \infty} f_t^i(1) = 1$  but in infinitely many rounds the mixed strategy  $x_t^i = 1_2$ . The key idea is that, when  $\alpha_t^i(1) - \alpha_t^i(2)$  is positive but lower than  $V\gamma_t$  in some round  $t$  (which happens infinitely often), we let the algorithm bid 2 with certainty in round  $t + 1$ . This does not violate the mean-based property.

---

### Algorithm 1 A mean-based bidding algorithm

---

**Require:**  $T_0 > 640$  such that  $\exp\left(-\frac{T_0^{1/3}}{900}\right) \leq \frac{1}{16}$ .

- 1: **for**  $t = 1, 2, \dots$  **do**
- 2:   **if**  $t \leq T_0 - T_0^{2/3}$  **then**
- 3:     Bid  $b_t = 1$ .
- 4:   **else if**  $T_0 - T_0^{2/3} + 1 \leq t \leq T_0$  **then**
- 5:     Bid  $b_t = 0$ .
- 6:   **else**
- 7:     Find  $k$  such that  $32^k T_0 + 1 \leq t \leq 32^{k+1} T_0$ .
- 8:     **if**  $t = 32^k T_0 + 1$ ,  $\text{argmax}_b \alpha_{t-1}(b) = 1$ , and  $\alpha_{t-1}^i(1) - \alpha_{t-1}^i(2) < V\gamma_t$  **then**
- 9:       Bid  $b_t = 2$ .
- 10:      **else**
- 11:       Bid  $b_t = \text{argmax}_{b \in \{0, 1, 2\}} \alpha_{t-1}(b)$  (break ties arbitrarily) with probability  $1 - T_{k+1}^{-1/3}$  and 0 with probability  $T_{k+1}^{-1/3}$ .
- 12:      **end if**
- 13:     **end if**
- 14:   **end for**

---

We note that this algorithm has no randomness in the first  $T_0$  rounds. It bids 1 in the first  $T_0 - T_0^{2/3}$  rounds and bid 0 in the remaining  $T_0^{2/3}$  rounds. Define round  $T_k = 32^k T_0$  for  $k \geq 0$ . Let  $\gamma_t = 1$  for  $1 \leq t \leq T_0$  and  $\gamma_t = T_k^{-1/4} = O(t^{-1/4})$  for  $t \in [T_k + 1, T_{k+1}]$  and all  $k \geq 0$ .

CLAIM 28. Algorithm 1 is a  $\gamma_t$ -mean-based algorithm with  $\gamma_t = O(t^{-1/4})$ .

PROOF. We only need to verify the mean-based property in round  $t \geq T_0 + 1$  since  $\gamma_t = 1$  for  $t \leq T_0$ . The proof follows by the definition and is straightforward: If the condition in Line 8 holds, where  $\text{argmax}_b \alpha_{t-1}(b) = 1$  and  $\alpha_{t-1}^i(1) - \alpha_{t-1}^i(2) \leq V\gamma_t$ , then the mean-based property does not apply to bids 1 and 2 and the algorithm bids 0 with probability  $0 \leq \gamma_t$ . Otherwise, according to Line 11, the algorithm bids  $b' \notin \text{argmax}_b \alpha_{t-1}(b)$  with probability at most  $T_{k+1}^{-1/3} \leq \gamma_t$ .  $\square$

For  $k \geq 0$ , denote  $A_k$  the event that for both  $i \in \{1, 2\}$ , it holds that  $T_k^{-\frac{1}{3}} \leq f_{T_k}^i(0) \leq 2T_k^{-\frac{1}{3}}$  and  $f_{T_k}^i(2) = \frac{k}{T_k}$ . Since both bidders submit deterministic bids in the first  $T_0$  rounds, it is easy to check that  $A_0$  holds probability 1.

The following two claims show that if  $A_0, A_1, \dots$  all happen, then the dynamics time-average converges to 1 while in the meantime, both of the bidders bid 2 at round  $T_k + 1$  for all  $k \geq 0$ .

CLAIM 29. If  $A_k$  happens, then both of the bidders bid 2 in round  $T_k + 1$ .

PROOF. According to Claim 27, we know that for any  $i \in \{1, 2\}$  and any  $t > T_0$ ,

$$\alpha_{t-1}^i(1) - \alpha_{t-1}^i(0) = f_{t-1}^{3-i}(1) + \frac{f_{t-1}^{3-i}(0)}{2} > 0.$$

Thus  $\operatorname{argmax}_b \{\alpha_{t-1}^i(b)\} \neq 0$  for any history  $H_{t-1}$ . Again by Claim 27, we have for any  $i \in \{1, 2\}$ .

$$0 < T_k^{-\frac{1}{3}} - \frac{k}{T_k} \leq \alpha_{T_k}^i(1) - \alpha_{T_k}^i(2) = f_{T_k}^{3-i}(0) - \frac{f_{T_k}^{3-i}(2)}{2} \leq f_{T_k}^{3-i}(0) \leq 2T_{k+1}^{-\frac{1}{3}} < 3T_{k+1}^{-\frac{1}{4}} = 3\gamma_{T_{k+1}}.$$

It follows by the definition of Algorithm 1 that both bidders bid 2 in round  $T_k + 1$ .  $\square$

**CLAIM 30.** *For any  $k \geq 0$  and  $i \in \{1, 2\}$ , if  $A_{k+1}$  holds, then  $f_t^i(1) \geq 1 - 64T_{k+1}^{-\frac{1}{3}} - \frac{32k}{T_{k+1}}$  holds for any  $t \in [T_k, T_{k+1}]$ .*

**PROOF.** Let  $A_{k+1}$  holds. Then

$$2T_{k+1}^{-\frac{1}{3}} \geq f_{T_{k+1}}^i(0) \geq \frac{tf_t^i(0)}{T_{k+1}} \geq \frac{f_t^i(0)}{32},$$

which implies that  $f_{T_k}^i(0) \leq 64T_{k+1}^{-\frac{1}{3}}$ . Similarly, we have  $f_t^i(2) \leq \frac{32k}{T_{k+1}}$ . The claim follows by  $f_t^i(1) = 1 - f_t^i(0) - f_t^i(2)$ .  $\square$

We now bound the probability of  $A_{k+1}$  given the fact that  $A_k$  happens, which is used later to derive a constant lower bound on the probability that  $A_k$  happens for all  $k \geq 0$ .

**CLAIM 31.** *For any  $k \geq 0$ ,*

$$\Pr[A_{k+1} \mid A_k] \geq 1 - 4 \exp\left(\frac{T_{k+1}^{\frac{1}{3}}}{900}\right).$$

**PROOF.** Suppose that  $A_k$  happens. We know from Claim 29 that both bidders bid 2 in round  $T_k + 1$ . The following claim shows the behaviour of the algorithm in rounds  $[T_k + 2, T_{k+1}]$ .

**CLAIM 32.** *For any  $i \in \{1, 2\}$  and any  $t \in [T_k + 2, T_{k+1}]$ ,*

$$\begin{aligned} \Pr[b_t^i = 1 \mid A_k] &= 1 - T_{k+1}^{-\frac{1}{3}} \\ \Pr[b_t^i = 0 \mid A_k] &= T_{k+1}^{-\frac{1}{3}}. \end{aligned}$$

**PROOF.** According to the definition of Algorithm 1, it suffices to prove that for any  $t \in [T_k + 2, T_{k+1}]$  and  $i \in \{1, 2\}$ ,  $\operatorname{argmax}_b \{\alpha_{t-1}^i(b)\} = 1$  holds.

We prove it by induction. For the base case, it is easy to verify that  $\alpha_{T_{k+1}}^i(1) - \alpha_{T_{k+1}}^i(2) = f_{T_{k+1}}^{3-i}(0) - \frac{f_{T_{k+1}}^{3-i}(2)}{2} > 0, \forall i \in \{1, 2\}$ . Suppose the claim holds for all of the rounds  $[T_k + 2, t]$ . Then none of the bidders bids 2 in rounds  $[T_k + 2, t]$ . It follows that for any  $i \in \{1, 2\}$ ,

$$\begin{aligned} \alpha_t^i(1) - \alpha_t^i(2) &= f_t^{3-i}(0) - \frac{f_t^{3-i}(2)}{2} \\ &\geq \frac{f_{T_k}^{3-i}(0)}{32} - \frac{k+1}{2T_k} \\ &\geq \frac{1}{32T_k^{\frac{1}{3}}} - \frac{k+1}{T_k} \\ &> 0 \text{ (since } T_0 > 64^{\frac{3}{2}}\text{).} \end{aligned}$$

Therefore  $\operatorname{argmax}_b \{\alpha_{t-1}^i(b)\} = 1$ . This completes the induction step.  $\square$

From the above proof we can also conclude that for  $i \in \{1, 2\}$ ,  $f_{T_{k+1}}^i(2) = \frac{k+1}{T_{k+1}}$ .

Note that the bidding strategies of a bidder at different rounds in  $[T_k + 2, T_{k+1}]$  are independent. According to Chernoff bound, we have for  $i \in \{1, 2\}$ ,

$$\begin{aligned} \Pr\left[\frac{29}{30} \frac{T_{k+1} - T_k - 1}{T_{k+1}^{\frac{1}{3}}} \leq \sum_{s=T_k+2}^{T_{k+1}} \mathbf{1}[b_s^i = 0] \leq \frac{31}{30} \frac{T_{k+1} - T_k - 1}{T_{k+1}^{\frac{1}{3}}} \mid A_k\right] &\geq 1 - 2 \exp\left(-\frac{T_{k+1} - T_k - 1}{450T_{k+1}^{\frac{2}{3}}}\right) \\ &\geq 1 - 2 \exp\left(-\frac{T_{k+1}^{\frac{1}{3}}}{900}\right). \end{aligned}$$

Therefore, with probability at least  $1 - 4 \exp\left(-\frac{T_{k+1}^{\frac{1}{3}}}{900}\right)$ , both of the above event happens. It implies that for  $i \in \{1, 2\}$

$$\begin{aligned} f_{T_{k+1}}^i(0) &\geq \frac{1}{T_{k+1}} \left( T_k f_{T_k}^i(0) + \frac{29}{30} \frac{T_{k+1} - T_k - 1}{T_{k+1}^{\frac{1}{3}}} \right) \\ &\geq \frac{1}{T_{k+1}} \left( \frac{T_k}{T_k^{\frac{1}{3}}} + \frac{29}{30} \frac{T_{k+1} - T_k - 1}{T_{k+1}^{\frac{1}{3}}} \right) \\ &\geq \frac{32^{\frac{1}{3}}}{32T_{k+1}^{\frac{1}{3}}} + \frac{29}{32T_{k+1}^{\frac{1}{3}}} \\ &\geq \frac{1}{T_{k+1}^{\frac{1}{3}}}, \end{aligned}$$

and

$$\begin{aligned} f_{T_{k+1}}^i(0) &\leq \frac{1}{T_{k+1}} \left( T_k f_{T_k}^i(0) + \frac{31}{30} \frac{T_{k+1} - T_k - 1}{T_{k+1}^{\frac{1}{3}}} \right) \\ &\leq \frac{1}{T_{k+1}} \left( \frac{2T_k}{T_k^{\frac{1}{3}}} + \frac{31}{30} \frac{T_{k+1} - T_k - 1}{T_{k+1}^{\frac{1}{3}}} \right) \\ &\leq \frac{2 \times 32^{\frac{1}{3}}}{32T_{k+1}^{\frac{1}{3}}} + \frac{31}{30T_{k+1}^{\frac{1}{3}}} \\ &\leq \frac{2}{T_{k+1}^{\frac{1}{3}}}. \end{aligned}$$

Therefore,  $A_{k+1}$  holds. This completes the proof.  $\square$

Using a union bound, we have

$$\begin{aligned} \Pr[\forall k \geq 0, A_k \text{ holds}] &\geq \Pr[A_0] \prod_{k=0}^{\infty} \Pr[A_{k+1} | A_k] \\ &\geq 1 - 4 \sum_{j=1}^{\infty} \exp\left(-\frac{T_j^{\frac{1}{3}}}{900}\right) \\ &\geq 1 - 4 \sum_{j=1}^{\infty} \exp\left(-\frac{T_0^{\frac{1}{3}} 3^j}{900}\right) \\ &= 1 - 4 \exp\left(-\frac{T_0^{\frac{1}{3}}}{300}\right) \left(1 + \sum_{j=2}^{\infty} \exp\left(-\frac{T_0^{\frac{1}{3}} (3^j - 3)}{900}\right)\right) \\ &\geq 1 - 8 \exp\left(-\frac{T_0^{\frac{1}{3}}}{300}\right) \\ &\geq \frac{1}{2}. \end{aligned}$$

Therefore, with probability at least  $\frac{1}{2}$ , the dynamics time-average converges to the equilibrium of 1, while both bidders' mixed strategies do not converge in the last-iterate sense. This completes the proof.

### A.3 Proof of Example 7

We only need to verify that the 0-mean-based property is satisfied for player 1 because players 2 and 3 always get zero utility no matter what they bid. Let  $q_t$  denote the fraction of the first  $t$  rounds where one of players 2 and 3 bids 6 (in the other  $1 - q_t$  fraction of rounds both players 2 and 3 bid 1); clearly,  $q_t \geq \frac{2}{3}$  for any  $t \geq 1$ . For player 1, at each round  $t$  her average utility by bidding 7 is  $\alpha_{t-1}^1(7) = 10 - 7 = 3$ ; by bidding 6,  $\alpha_{t-1}^1(6) = (10 - 6)(\frac{1}{2}q_{t-1} + (1 - q_{t-1})) = 4(1 - \frac{q_{t-1}}{2}) \leq \frac{8}{3} < 3$ ; by bidding 2,  $\alpha_{t-1}^1(2) = (10 - 2)(1 - q_{t-1}) \leq \frac{8}{3} < 3$ ; and clearly  $\alpha_{t-1}^1(b) < 3$  for any other bid. Hence,  $7 = \operatorname{argmax}_{b \in \mathcal{B}^1} \{\alpha_{t-1}^1(b)\}$ .

## B MISSING PROOFS FROM SECTION 4

### B.1 Proof of Claim 9

Let  $\Gamma = \{s \leq t-1 \mid \exists i \in M^1, b_s^i \leq v^1 - 3\}$ . It follows that the premise of the claim becomes  $\frac{|\Gamma|}{t-1} \leq \frac{1}{3NV}$ . First, note that

$$\begin{aligned} P_{t-1}^i(0 : v^1 - 3) &= \frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\max_{i' \neq i} b_s^{i'} \leq v^1 - 3] \\ &\leq \frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] = \frac{|\Gamma|}{t-1} \leq \frac{1}{3NV}. \end{aligned} \quad (12)$$

Then, according to (4),

$$\begin{aligned} \alpha_{t-1}^i(v^1 - 1) - \alpha_{t-1}^i(v^1 - 2) \\ = Q_{t-1}^i(v^1 - 1) + P_{t-1}^i(v^1 - 2) - 2Q_{t-1}^i(v^1 - 2) - P_{t-1}^i(0 : v^1 - 3). \end{aligned} \quad (13)$$

Using  $Q_{t-1}^i(v^1 - 1) \geq \frac{1}{N}P_{t-1}^i(v^1 - 1)$  and  $Q_{t-1}^i(v^1 - 2) \leq \frac{1}{2}P_{t-1}^i(v^1 - 2)$  from (3), we can lower bound (13) by

$$\frac{1}{N}P_{t-1}^i(v^1 - 1) - P_{t-1}^i(0 : v^1 - 3).$$

With (12), we get

$$\alpha_{t-1}^i(v^1 - 1) - \alpha_{t-1}^i(v^1 - 2) \geq \frac{1}{N}P_{t-1}^i(v^1 - 1) - \frac{1}{3NV}.$$

If  $\frac{1}{N}P_{t-1}^i(v^1 - 1) - \frac{1}{3NV} > V\gamma_t$ , then  $\alpha_{t-1}^i(v^1 - 1) - \alpha_{t-1}^i(v^1 - 2) > V\gamma_t$ . Therefore,  $\Pr[b_t^i = v^1 - 2 \mid H_{t-1}] \leq \gamma_t$ .

Suppose  $\frac{1}{N}P_{t-1}^i(v^1 - 1) - \frac{1}{3NV} \leq V\gamma_t$ , which is equivalent to

$$P_{t-1}^i(v^1 - 1) \leq \frac{1}{3V} + NV\gamma_t.$$

Consider  $Q_{t-1}^i(v^1 - 2)$ . By the definition of  $\Gamma$ , in all rounds  $s \notin \Gamma$  and  $s \leq t-1$ , we have that all bidders in  $M^1$  bid  $v^1 - 2$  or  $v^1 - 1$ . If bidder  $i$  wins with bid  $v^1 - 2$  in round  $s \notin \Gamma$ , she must be tied with at least two other bidders in  $M^1$  since  $|M^1| \geq 3$ ; if bidder  $i$  wins with bid  $v^1 - 2$  (tied with at least one other bidder) in round  $s \in \Gamma$ , that round contributes at most  $\frac{1}{2}$  to the summation in  $Q_{t-1}^i(v^1 - 2)$ . Therefore,

$$Q_{t-1}^i(v^1 - 2) \leq \frac{1}{t-1} \left( \frac{(t-1) - |\Gamma|}{3} + \frac{|\Gamma|}{2} \right) = \frac{1}{3} + \frac{1}{6} \frac{|\Gamma|}{t-1} \leq \frac{1}{3} + \frac{1}{18NV}. \quad (14)$$

We then consider  $P_{t-1}^i(v^1 - 2)$ . Since  $P_{t-1}^i(0 : v^1 - 3) + P_{t-1}^i(v^1 - 2) + P_{t-1}^i(v^1 - 1) = 1$ , and recalling that  $P_{t-1}^i(0 : v^1 - 3) \leq \frac{1}{3NV}$  and  $P_{t-1}^i(v^1 - 1) \leq \frac{1}{3V} + NV\gamma_t$ , we get

$$P_{t-1}^i(v^1 - 2) = 1 - P_{t-1}^i(0 : v^1 - 3) - P_{t-1}^i(v^1 - 1) \geq 1 - \frac{1}{3NV} - \frac{1}{3V} - NV\gamma_t. \quad (15)$$

Combining (13) with (12), (14), and (15), we get

$$\begin{aligned} \alpha_{t-1}^i(v^1 - 1) - \alpha_{t-1}^i(v^1 - 2) \\ \geq 0 + \left( 1 - \frac{1}{3NV} - \frac{1}{3V} - NV\gamma_t \right) - 2 \left( \frac{1}{3} + \frac{1}{18NV} \right) - \frac{1}{3NV} \\ = \frac{1}{3} - \frac{3N+7}{9NV} - NV\gamma_t \\ \geq \frac{1}{3} - \frac{13}{18V} - \frac{1}{12NV} \quad (\text{since } N \geq 2 \text{ and } \gamma_t \leq \frac{1}{12N^2V^2}) \\ \geq \frac{5}{54} - \frac{1}{12NV} \quad (\text{since } V \geq 3) \geq \frac{1}{12NV} > V\gamma_t. \end{aligned}$$

Therefore, by the mean-based property,  $\Pr[b_t^i = v^1 - 2 \mid H_{t-1}] \leq \gamma_t$ .

## B.2 Proof of Corollary 13

Using Lemma 10 and Lemma 11 from  $k = 0$  to  $v_1 - 4$ , we get

$$\Pr[A_{v^1-3}] \geq \Pr[A_0, A_1, \dots, A_{v^1-3}] \geq 1 - \exp\left(-\frac{T_b}{24NV}\right) - \sum_{k=0}^{v^1-4} \sum_{j=1}^d \exp\left(-\frac{|\Gamma_k^j|}{1152N^2V^2}\right).$$

Note that  $|\Gamma_k^j| = T_k^j - T_k^{j-1} = cT_k^{j-1} - cT_k^{j-2} = c|\Gamma_k^{j-1}|$ , for any  $k \in \{0, 1, 2, \dots, v^1 - 4\}$  and  $j \in \{2, \dots, d\}$ , and that  $|\Gamma_k^1| = c|\Gamma_{k-1}^d|$  for any  $k \in \{1, 2, \dots, v^1 - 4\}$ . We also note that  $|\Gamma_0^1| = (c-1)T_0 = T_b$ . Thus,

$$\sum_{k=0}^{v^1-4} \sum_{j=1}^d \exp\left(-\frac{|\Gamma_k^j|}{1152N^2V^2}\right) = \sum_{s=0}^{(v^1-3)d-1} \exp\left(-\frac{c^s T_b}{1152N^2V^2}\right).$$

We then upper bound the above equation by

$$\begin{aligned} &\leq \sum_{s=0}^{\infty} \exp\left(-\frac{c^s T_b}{1152N^2V^2}\right) \\ &= \exp\left(-\frac{T_b}{1152N^2V^2}\right) \left(1 + \sum_{s=1}^{\infty} \exp\left(-\frac{(c^s - 1)T_b}{1152N^2V^2}\right)\right). \end{aligned}$$

It suffices to prove that  $\sum_{s=1}^{\infty} \exp\left(-\frac{(c^s - 1)T_b}{1152N^2V^2}\right) \leq 1$ . Since  $c^s - 1 \geq c - 1 + (s-1)(c^2 - c)$ ,  $\forall s \geq 1$ , we have

$$\begin{aligned} &\sum_{s=1}^{\infty} \exp\left(-\frac{(c^s - 1)T_b}{1152N^2V^2}\right) \\ &\leq \sum_{s=1}^{\infty} \exp\left(-\frac{(c-1)T_b}{1152N^2V^2}\right) \left(\exp\left(-\frac{(c^2 - c)T_b}{1152N^2V^2}\right)\right)^{s-1} \\ &\leq \sum_{s=1}^{\infty} \left(\frac{1}{2}\right)^s = 1, \end{aligned}$$

where the second inequality holds because  $\exp\left(-\frac{(c^2 - c)T_b}{1152N^2V^2}\right) \leq \exp\left(-\frac{(c-1)T_b}{1152N^2V^2}\right) \leq \frac{1}{2}$  by the assumption on  $T_b$ .  $\square$

## B.3 Proof of Claim 14

Since  $\delta_{T_a^0} \rightarrow 0$  and  $\gamma_{T_a^0} \rightarrow 0$  as  $T_b \rightarrow \infty$ , when  $T_b$  is sufficiently large we have

$$F_{T_a^1} = \frac{1}{c} \frac{1}{4NV} + \frac{c-1}{c} \left(\delta_{T_a^0} + |M^1|V\gamma_{T_a^0}\right) \leq \frac{1}{c} \frac{1}{4NV} + \frac{c-1}{c} \frac{1}{4NV} \leq \frac{1}{4NV} = F_{T_a^0}.$$

By definition, for every  $k \geq 1$

$$F_{T_a^{k+1}} = \frac{1}{c} F_{T_a^k} + \frac{c-1}{c} \left(\delta_{T_a^k} + |M^1|V\gamma_{T_a^k}\right), \quad F_{T_a^k} = \frac{1}{c} F_{T_a^{k-1}} + \frac{c-1}{c} \left(\delta_{T_a^{k-1}} + |M^1|V\gamma_{T_a^{k-1}}\right).$$

Using the fact that  $F_{T_a^k} \leq F_{T_a^{k-1}}$  and that  $\delta_{T_a^k} + |M^1|V\gamma_{T_a^k}$  is decreasing in  $k$ , we have  $F_{T_a^{k+1}} \leq F_{T_a^k} \leq \frac{1}{4NV}$ . Similarly, we have  $\tilde{F}_{T_a^{k+1}} \leq \tilde{F}_{T_a^k}$  for any  $k \geq 0$ .

Note that  $\delta_{T_a^k} \rightarrow 0$  and  $\gamma_{T_a^0} \rightarrow 0$  as  $k \rightarrow +\infty$ . Therefore, for any  $0 < \varepsilon \leq \frac{1}{4NV}$ , we can find  $k$  sufficiently large such that  $\frac{1}{c^{k/2}} \leq \frac{\varepsilon}{6}$ ,  $\delta_{T_a^k} \leq \frac{\varepsilon}{6}$ , and  $\gamma_{T_a^k} \leq \frac{\varepsilon}{6|M^1|V}$ . Then we have

$$\begin{aligned} F_{T_a^k} &\leq \tilde{F}_{T_a^k} = \frac{1}{c^k} + \sum_{s=0}^{k-1} \frac{c-1}{c^{k-s}} \delta_{T_a^s} + \sum_{s=0}^{k-1} |M^1|V \frac{c-1}{c^{k-s}} \gamma_{T_a^s} \\ &\leq \frac{\varepsilon}{3} + 2 \sum_{s=0}^{k/2-1} \frac{c-1}{c^{k-s}} + \sum_{s=k/2}^{k-1} \frac{c-1}{c^{k-s}} (\delta_{T_a^{k/2}} + |M^1|V \frac{c-1}{c^{k-s}} \gamma_{T_a^{k/2}}) \\ &\leq \frac{\varepsilon}{3} + 2 \frac{1}{c^{k/2}} + \frac{\varepsilon}{3} \sum_{s=k/2}^{k-1} \frac{c-1}{c^{k-s}} \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Thus for any  $l \geq k$ , we have  $F_{T_a^l} \leq \tilde{F}_{T_a^l} \leq \varepsilon$ . Since  $F_{T_a^k}$  and  $\tilde{F}_{T_a^k}$  are both positive, we have  $\lim_{k \rightarrow \infty} F_{T_a^k} = \lim_{k \rightarrow \infty} \tilde{F}_{T_a^k} = 0$ .  $\square$

## B.4 Proof of Lemma 15

We use an induction to prove the following:

$$\Pr[A_a^{k+1}] \geq 1 - \exp\left(-\frac{T_b}{24NV}\right) - 2 \exp\left(-\frac{T_b}{1152N^2V^2}\right) - \sum_{s=0}^k \exp\left(-\frac{1}{2}|\Gamma_a^{s+1}|\delta_{T_a^s}^2\right).$$

We do not assume  $|M^1| \geq 3$  for the moment. The base case follows from Corollary 11 because  $A_a^0$  is the same as  $A_{v^1-3}$ . Suppose  $A_a^k$  happens. Consider  $A_a^{k+1}$ . For any round  $t \in \Gamma_a^{k+1}$ ,

$$\begin{aligned} P_{t-1}^i(0 : v^1 - 3) &\leq \frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \\ &= \frac{1}{t-1} \left( \sum_{s=1}^{T_a^k} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] + \sum_{s=T_a^k+1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \right) \\ (F_{T_a^k} \leq \frac{1}{4NV}) &\leq \frac{1}{t-1} \left( \frac{T_a^k}{4NV} + (t-1 - T_a^k) \right) \\ (T_a^k \leq t-1 \leq T_a^{k+1}) &\leq \frac{1}{T_a^k} \left( \frac{T_a^k}{4NV} + T_a^{k+1} - T_a^k \right) \\ (T_a^{k+1} = cT_a^k) &= \frac{1}{3NV} \stackrel{(\gamma_t < \frac{1}{12NV})}{<} \frac{1}{2NV} - 2\gamma_t. \end{aligned}$$

By Claim 8 and a similar analysis to Claim 12, for any history  $H_{t-1}$  that satisfies  $A_a^k$ ,

$$\Pr[\exists i \in M^1, b_t^i \leq v^1 - 3 \mid H_{t-1}, A_a^k] \leq |M^1|V\gamma_t. \quad (16)$$

Let  $Z_t = \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] - |M^1|V\gamma_t$  and let  $X_t = \sum_{s=T_a^k+1}^t Z_s$ . We have  $\mathbb{E}[Z_t \mid A_a^k, H_{t-1}] \leq 0$ . Therefore, the sequence  $X_{T_a^k+1}, X_{T_a^k+2}, \dots, X_{T_a^{k+1}}$  is a supermartingale (with respect to the sequence of history  $H_{T_a^k}, H_{T_a^k+1}, \dots, H_{T_a^{k+1}-1}$ ). By Azuma's inequality, for any  $\Delta > 0$ , we have

$$\Pr\left[\sum_{t \in \Gamma_a^{k+1}} Z_t \geq \Delta \mid A_a^k\right] \leq \exp\left(-\frac{\Delta^2}{2|\Gamma_a^{k+1}|}\right).$$

Let  $\Delta = |\Gamma_a^{k+1}|\delta_{T_a^k}$ . Then with probability at least  $1 - \exp(-\frac{1}{2}|\Gamma_a^{k+1}|\delta_{T_a^k}^2)$ , we have

$$\sum_{t \in \Gamma_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] < \Delta + |M^1|V \sum_{t \in \Gamma_a^{k+1}} \gamma_t \leq |\Gamma_a^{k+1}|\delta_{T_a^k} + |M^1|V|\Gamma_a^{k+1}|\gamma_{T_a^k}, \quad (17)$$

which implies

$$\begin{aligned} \frac{1}{T_a^{k+1}} \sum_{t=1}^{T_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] &= \frac{1}{T_a^{k+1}} \left( \sum_{t=1}^{T_a^k} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] + \sum_{t \in \Gamma_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 3] \right) \\ &\leq \frac{1}{T_a^{k+1}} \left( T_a^k F_{T_a^k} + |\Gamma_a^{k+1}|\delta_{T_a^k} + |M^1|V|\Gamma_a^{k+1}|\gamma_{T_a^k} \right) \\ (\text{since } T_a^{k+1} = cT_a^k) &= \frac{1}{c} F_{T_a^k} + \frac{c-1}{c} \delta_{T_a^k} + |M^1|V \frac{c-1}{c} \gamma_{T_a^k} \\ (\text{by definition}) &= F_{T_a^{k+1}} \end{aligned}$$

and thus  $A_a^{k+1}$  holds.

Now we suppose  $|M^1| \geq 3$ , then we can change (16) to

$$\Pr[\exists i \in M^1, b_t^i \leq v^1 - 2 \mid H_{t-1}, A_a^k] \leq |M^1|V\gamma_t$$

because of Claim 9 and the fact that  $\frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{I}[\exists i \in M^1, b_s^i \leq v^1 - 3] \leq \frac{1}{3NV}$ . The definition of  $Z_t$  is changed accordingly, and (17) becomes

$$\sum_{t \in \Gamma_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 2] < |\Gamma_a^{k+1}| \delta_{T_a^k} + |M^1|V|\Gamma_a^{k+1}| \gamma_{T_a^k},$$

which implies

$$\frac{1}{T_a^{k+1}} \sum_{t=1}^{T_a^{k+1}} \mathbb{I}[\exists i \in M^1, b_t^i \leq v^1 - 2] \leq \frac{1}{T_a^{k+1}} \left( T_a^k \tilde{F}_k + |\Gamma_a^{k+1}| \delta_{T_a^k} + |M^1|V|\Gamma_a^{k+1}| \gamma_{T_a^k} \right) = \tilde{F}_{T_a^{k+1}}.$$

To conclude, by induction,

$$\begin{aligned} \Pr[A_a^{k+1}] &= \Pr[A_a^k] \Pr[A_a^{k+1}|A_a^k] \geq \Pr[A_a^k] - \exp\left(-\frac{1}{2} |\Gamma_a^{k+1}| \delta_{T_a^k}^2\right) \\ &\geq 1 - \exp\left(-\frac{T_b}{24NV}\right) - 2 \exp\left(-\frac{T_b}{1152N^2V^2}\right) - \sum_{s=0}^k \exp\left(-\frac{1}{2} |\Gamma_a^{s+1}| \delta_{T_a^s}^2\right). \end{aligned}$$

As  $\delta_t = (\frac{1}{t})^{\frac{1}{3}}$  and  $|\Gamma_a^s| = c^{s+d(v^1-3)-1}(c-1)T_0$ ,  $T_a^s = c^{s+d(v^1-3)}T_0$  (we abuse the notation and let  $v^1 - 3 = 0$  if  $v^1 < 3$ ), we have

$$\begin{aligned} &\sum_{s=0}^k \exp\left(-\frac{1}{2} |\Gamma_a^{s+1}| \delta_{T_a^s}^2\right) \\ &= \sum_{s=0}^k \exp\left(-\frac{1}{2} c^{\frac{1}{3}(s+d(v^1-3))} (c-1)(T_0)^{\frac{1}{3}}\right) \\ &= \exp\left(-\frac{1}{2} c^{\frac{1}{3}d(v^1-3)} (c-1)(T_0)^{\frac{1}{3}}\right) \left(1 + \sum_{s=1}^k \exp\left(-\frac{1}{2} c^{\frac{1}{3}d(v^1-3)} (c-1)(T_0)^{\frac{1}{3}} (c^{\frac{s}{3}} - 1)\right)\right) \\ &\leq \exp\left(-\frac{1}{2} c^{\frac{1}{3}d(v^1-3)} (c-1)(T_0)^{\frac{1}{3}}\right) \left(1 + \sum_{s=1}^k \exp\left(-\frac{1}{2} c^{\frac{1}{3}d(v^1-3)} (c-1)(T_0)^{\frac{1}{3}} s(c^{\frac{1}{3}} - 1)\right)\right) \\ &\leq \exp\left(-\frac{1}{2} c^{\frac{1}{3}d(v^1-3)} (c-1)(T_0)^{\frac{1}{3}}\right) \left(1 + \sum_{s=1}^k \left(\frac{1}{2}\right)^s\right) \\ &\leq 2 \exp\left(-\frac{1}{2} c^{\frac{1}{3}d(v^1-3)} (c-1)(T_0)^{\frac{1}{3}}\right), \end{aligned}$$

where in the last but one inequality we suppose that  $T_0$  is large enough so that  $\exp\left(-\frac{1}{2} c^{\frac{1}{3}d(v^1-3)} (c-1)(T_0)^{\frac{1}{3}} s(c^{\frac{1}{3}} - 1)\right) \leq \frac{1}{2}$ . Substituting  $T_0 = 12NVT_b = \frac{1}{c-1}T_b$ ,  $c = 1 + \frac{1}{12NV}$ , and  $c^d = 8NV$  gives

$$\begin{aligned} \sum_{s=0}^k \exp\left(-\frac{1}{2} |\Gamma_a^{s+1}| \delta_{T_a^s}^2\right) &\leq 2 \exp\left(-\left(\frac{(8NV)^{(v^1-3)} T_b}{1152N^2V^2}\right)^{\frac{1}{3}}\right) \\ &\leq 2 \exp\left(-\left(\frac{T_b}{1152N^2V^2}\right)^{\frac{1}{3}}\right), \end{aligned}$$

concluding the proof.  $\square$