Incentive-aware Contextual Pricing with Non-parametric Market Noise

Negin Golrezaei
MIT Sloan

Patrick Jaillet
MIT EECS

Jason Cheuk Nam Liang
MIT Operations Research Center

Abstract

We consider a dynamic pricing problem for repeated contextual second-price auctions with multiple strategic buyers who aim to maximize their long-term time discounted utility. The seller has limited information on buyers' overall demand curves which depends on a non-parametric market-noise distribution, and buyers may potentially submit corrupted bids (relative to true valuations) to manipulate the seller's pricing policy for more favorable reserve prices in the future. We focus on designing the seller's learning policy to set contextual reserve prices where the seller's goal is to minimize regret compared to the revenue of a benchmark clairvoyant policy that has full information of buyers' demand. We propose a policy with a phased-structure that incorporates randomized "isolation" periods, during which a buyer is randomly chosen to solely participate in the auction. We show that this design allows the seller to control the number of periods in which buyers significantly corrupt their bids. We then prove that our policy enjoys a T-period regret of $\widetilde{\mathcal{O}}(\sqrt{T})$ facing strategic buyers. Finally, we conduct numerical simulations to compare our proposed algorithm to standard pricing policies. Our numerical results show that our algorithm outperforms these policies under various buyer bidding behavior.

1 INTRODUCTION

We study the problem of designing pricing policies for highly heterogeneous items against strategic agents. The motivation comes from the availability of massive amount of real-time data in online platforms and in particular, online advertising markets, where the seller has access to detailed information about item features/contexts. In such environments, designing optimal policies involves learning buyers' demand (which is a mapping from item features and offered prices to the likelihood of the item being sold) under limited understanding of buyers' behavior. Our key goal is to develop effective and robust dynamic pricing polices that facilitate such a complex learning process for very general non-parametric contextual demand curves facing strategic buyers.

Formally, we study the setting wherein any period t over a finite time horizon T, the seller sells one item to buyers via running a second price auction with a reserve price. The item is characterized by a d-dimensional context vector x_t, public to the seller and buyers. We consider an interdependent contextual valuation model in which a buyer's valuation for the item is the sum of common and private components. The common component determines the expected willingness-to-pay of buyers and is the inner product of the feature vector and a fixed "mean vector" β that is homogeneous across buyers; the private component, which captures buyers' idiosyncratic preferences, is independently sampled from an unknown *non-parametric* noise distribution F. We note that such a linear valuation model is very common in the literature of dynamic pricing; e.g. see Golrezaei et al. (2018); Javanmard and Nazerzadeh (2016); Kanoria and Nazerzadeh (2017) and Javanmard (2017).

Under this interdependent contextual valuation model, we study a *strategic setting* where buyers intend to maximize long-term discounted utility and may consequently submit *corrupted*, i.e., untruthful, bids. The motivation of this strategic setting comes from the repeated buyer-seller interactions when the seller does not possess full information on buyers' demand and aims to learn it using buyers' submitted bids. In a single-shot second price auction, where there is no repeated interactions between the seller and buyers, bidding truthfully is a buyer's weakly dominant action. However, this is no longer the case in our repeated second price auction setting: repeated auctions may incentivize the buyers to submit corrupted bids, rather than their true valuations, in order to manipulate seller's future reserve prices; e.g. by underbidding, buyers may trick the seller to lower future reserve prices.

In this work, we would like to design a reserve price pol-

Proceedings of the 26[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2023, Valencia, Spain. PMLR: Volume 206.

icy for the seller who does not know the mean vector β and the noise distribution F. The policy dynamically learns/optimizes contextual reserve prices while being robust to corrupted data (bids), submitted by strategic buyers. In particular, our objective is to minimize our policy's regret computed against a clairvoyant benchmark policy that knows both β and F. Designing low-regret policies in our setting involves overcoming the following challenges: (i) The demand curve is constantly shifting due to the change in contexts over time. (ii) The shape of the demand curve is unknown due to the lack of information on the market noise distribution F which may not enjoy a parametric functional form. Furthermore, we do not impose the *Monotone Hazard Rate* (MHR)[1] assumption on F. While the MHR assumption is common in the related literature and can significantly simplify reserve price optimization (see e.g. Remark 1), it has been shown to fail in practice (see Celis et al. (2014); Golrezaei et al. (2017)). (iii) As stated earlier, in our strategic setting, buyers may take advantage of the seller's lack of knowledge about buyers' demand and submit corrupted bids to manipulate future reserve prices.

Main contribution. We develop a policy called *Non-Parametric Contextual Policy against Strategic Buyers* (NPAC-S) that enables the seller to efficiently learn the optimal contextual reserve prices while being robust against buyers' corrupted bids. Our policy design incorporates two simple yet effective features, namely a *phased structure* and *random isolation*. First, NPAC-S partitions the entire horizon into consecutive phases, and then estimates the mean vector and the distributions of the second-highest and highest valuations only using data from the previous phase. This reduces the buyers' manipulating power on future reserve prices as past corrupted bids prior to the previous phase will not affect future pricing decisions. Second, the NPAC-S policy incorporates randomized isolation periods, that is, in each period with some probability the seller chooses a particular buyer at random and let her be the single participant of the auction during this period. In these isolation periods, the isolated buyer faces no competition from other buyers, and hence may incur large utility loss if a significantly corrupted bid is submitted.[2]

For our main theoretical results, we show that in virtue of our isolation periods in our design of NPAC-S, the number of past periods with large corruptions is $\mathcal{O}(\log(t))$ for any period t via leveraging the fact that buyers aim to maximize their long-term discounted utility. Furthermore, we present novel high probability bounds for our estimation errors in β and F which are estimated by ordinary least squares and empirical distributions, respectively, with the presence of corrupted bids. Finally, in Theorem 1, we show that the NPAC-S policy achieves a regret of $\widetilde{\mathcal{O}}(d\sqrt{T})$ for general non-parametric distributions F against a clairvoyant benchmark policy.

Related literature. Here we discuss related works that study dynamic pricing against strategic buyers with stochastic valuations, [3] and refer readers to Appendix A for broader related works.

Both Amin et al. (2013, 2014) study a dynamic pricing problem in a posted price auction against a single strategic buyer. Amin et al. (2013) addresses the non-contextual stochastic valuation setting, where as Amin et al. (2014) studies a linear contextual valuation model, but with no market noise disturbance. Amin et al. (2014) proposes an algorithm that achieves $\widetilde{\mathcal{O}}(T^{2/3})$ regret in contrast with our regret of $\widetilde{\mathcal{O}}(\sqrt{T})$ using the NPAC-S policy. We point out that this is because the seller in their setting only observes the outcome of the auction (i.e. bandit feedback), while in our setting we assume that seller can examine all submitted bids. Our setting is more complex compared to Amin et al. (2013, 2014) as we handle the contextual pricing problem against multiple strategic buyers, and also deals with the issue of learning a non-parametric distribution function in the presence of strategic buyer behavior. Kanoria and Nazerzadeh (2017) consider a contextual buyer valuation model similar to ours (but with the MHR assumption on the market noise distribution) and proposes a pricing algorithm that sets personalized reserve prices for individual buyers. They argue that the design of their algorithm induces an equilibrium where buyers always bid truthfully, and then further assume buyers act according to this equilibrium. Our work distinguishes itself from two aspects. First, setting personalized reserve prices in Kanoria and Nazerzadeh (2017) rely crucially on the MHR assumption, and in this paper we relax this assumption such that our methodology works for a larger class of market noise distributions. Second, we consider more general buyers who do not necessarily play any equilibrium and are forward looking. Golrezaei et al. (2018) study a similar interdependent contextual valuation model to ours, but with heterogeneous mean vector β across agents. Our work distinguishes itself from Golrezaei et al. (2018) in two major ways. First, they focus on optimizing contextual

[1]Distribution F is MHR if $\frac{f(z)}{1-F(z)}$ is non-decreasing in z, where f is the corresponding pdf.

[2]In the isolation periods, when the valuation of the isolated buyer is greater than the reserve price, significantly underbidding may cause the item to not be allocated; when the valuation of the isolated buyer is lower than the reserve price, overbidding results in the buyer paying much higher prices (relative to valuation) to achieve the item. In either case, the isolated buyer will incur a significant utility loss compared to truthful bidding.

[3]The general theme of learning in the presence of strategic agents or corrupted information has also been studied in other applications; see, for example, Chen and Keskin (2018); Birge et al. (2018); Feng et al. (2019). There are also related works that study adversarial buyer valuations. For example, Drutsa (2019) studies the seller's pricing problem for repeated second-price auctions facing multiple strategic buyers with private valuations fixed overtime. In addition, buyers in this work also seek to maximize cumulative discounted utility. The paper proposes an algorithm that achieves $\mathcal{O}(\log\log(T))$ regret for worst-case (adversarial) valuations.

reserve w.r.t. the worst-case distribution among a known class of MHR market noise distributions. In contrast, our work relaxes this constraint and does not require the seller to have any prior knowledge on the possibly non-parametric distribution. Second, in their setting, the seller only utilizes the outcome of the auctions to learn buyer demand and results in a regret of $\widetilde{\mathcal{O}}(T^{2/3})$.[4] In our work, we exploit the information of all submitted bids by taking advantage of the fact that buyers' utility-maximizing behaviour constrains their degree of corruption on bids. This eventually allows us to achieve an improved regret of $\widetilde{\mathcal{O}}(\sqrt{T})$. Nevertheless, our proposed algorithm cannot not handle heterogeneous β's, and hence this will be an interesting future research direction. Drutsa (2020) studies the posted price selling problem against a strategic agent with a non-linear (stochastic) contextual valuation model that satisfies some Lipschitz condition with no additive noise.

We summarize some key differences in the settings/results of the aforementioned works in Table 1.

2 PRELIMINARIES

Notation. For $a \in \mathbb{N}^+$, denote $[a] = \{1, 2, \ldots, a\}$. For two vectors $x, y \in \mathbb{R}^d$, denote $\langle x, y \rangle$ as their inner product. Finally, $\mathbb{I}\{\cdot\}$ is the indicator function: $\mathbb{I}\{\mathcal{A}\} = 1$ if event $\mathcal{A}$ occurs and 0 otherwise.

We consider a seller who runs repeated second price auctions over a horizon with length T that is unknown to the seller. In each auction $t \in [T]$, an item is sold to N buyers, where the item is characterized by a d-dimensional feature vector $x_t \in \mathcal{X} \subset \{x \in \mathbb{R}^d : \|x\|_\infty \leq x_{\max}\}$ where $0 < x_{\max} < \infty$. We assume that x_t is independently drawn from some distribution $\mathcal{D}$ unknown to the seller. We define Σ as the covariance matrix of distribution $\mathcal{D}$.[6] We assume that Σ is positive definite and unknown to the seller, and define the smallest eigenvalue of Σ to be $\lambda_0 > 0$.

Buyer valuation model. We focus on an interdependent valuation model where the valuation of buyer $i \in [N]$ at time $t \in [T]$is given by $v_{i,t} = \langle \beta, x_t \rangle + \epsilon_{i,t}$. Here, β is called the *mean vector* and is fixed over time and unknown to the seller, while $\epsilon_{i,t}$ is idiosyncratic market noise sampled independently over time and across buyers from some time-invariant distribution F with probability density function f, both unknown to the seller. Furthermore, F has bounded support $(-\epsilon_{\max}, \epsilon_{\max})$, in which its probability density function is bounded by $c_f := \sup_{z \in [-\epsilon_{\max}, \epsilon_{\max}]} f(z) \geq \inf_{z \in [-\epsilon_{\max}, \epsilon_{\max}]} f(z) > 0$. The support boundary $\epsilon_{\max}$ is not necessarily known to the seller. We assume there exist $v_{\max} > 0$ so that $v_{i,t} \in [0, v_{\max}]$ for all $i \in [N], t \in [T]$.

We highlight that our setting does not enforce distribution F to be parametric nor to satisfy the MHR assumption. This is because via analyzing real auction data sets, it has been shown that the MHR assumption does not necessarily hold in online advertising markets Celis et al. (2014); Golrezaei et al. (2017).

Repeated contextual second price auctions with reserve. The contextual second price auction with reserve is described as followed for $N \geq 2$ buyers: In any period $t \geq 1$, a context vector $x_t \sim \mathcal{D}$ is revealed to the seller and buyers. The seller then computes reserve price r_t, while simultaneously each buyer $i \in [N]$ forms individual valuations $v_{i,t}$ and submits a bid $b_{i,t}$ to the seller. Let $i^\star = \arg\max_{i \in [N]} b_{i,t}$ be the buyer who submitted the highest bid.[7] If $b_{i^\star,t} \geq r_t$, the item is allocated to buyer $i^\star$ and he is charged the maximum between the reserve price and second highest bid, i.e. buyer $i^\star$ pays $p_{i^\star,t} = \max\{r_t, \max_{i \neq i^\star} b_{i,t}\}$. For any other buyer $i \neq i^\star$, the payment $p_{i,t} = 0$. In the case where $b_{i^\star,t} < r_t$, the item is not allocated and all payments are zero.

Here, the seller's reserve price r_t can only depend on x_t and the history set $\mathcal{H}_{t-1} := \{(r_1, \{b_{i,1}\}_{i \in [N]}, x_1), \ldots, (r_{t-1}, \{b_{i,t-1}\}_{i \in [N]}, x_{t-1})\}$ which includes all information available to the seller up to period $t-1$.

Buyers' bidding behavior. In the setting where buyers are strategic, we assume that in any period t, each buyer $i \in [N]$ aims at maximizing his long-term discounted utility $U_{i,t}$:

$$U_{i,t} := \sum_{\tau=t}^{T} \eta^\tau \mathbb{E}\left[v_{i,\tau} w_{i,\tau} - p_{i,\tau}\right], \tag{1}$$

where $\eta \in (0, 1)$ is the discount factor, $w_{i,t} \in \{0, 1\}$ indicates whether buyer i wins the item; the expectation is taken with respect to the randomness due to the noise distribution F, the context distribution $\mathcal{D}$, buyers' bidding behavior, and the seller's pricing policy. We point out that this discounted utility model illustrates the fact that buyers are less patient than the seller, and is a common framework in many dynamic pricing literature; see Amin et al. (2013, 2014); Golrezaei et al. (2018), and Liu et al. (2018). The motivation lies in many applications in online advertisement markets wherein the user traffic is usually very uncertain and as a result, advertisers (buyers) would not like to miss out an opportunity of showing their ads to targeted users. An alternative interpretation for the above discounted utility model is that each buyer has probability η of leaving the

[4]A recent work Deng et al. (2019) builds on the result of Golrezaei et al. (2018) by considering a stronger benchmark that knows future buyer valuation distributions (noise distribution and all the future contextual information). They design robust pricing schemes whose regret is $\mathcal{O}(T^{5/6})$ against the aforementioned benchmark, confirming the generalizability of pricing schemes in Golrezaei et al. (2018).

[6]The covariance matrix of a distribution $\mathcal{P}$ on $\mathbb{R}^d$ is defined as $\mathbb{E}_{x \sim \mathcal{P}}[xx^\top] - \mu\mu^\top$, where $\mu = \mathbb{E}_{x \sim \mathcal{P}}[x]$.

[7]No ties will occur since we assume that no valuations and bids are the same.

Table 1: Summary of settings and results for seller algorithms that sell against strategic agents with stochastic valuations. Note that the Discount util. column indicates whether the algorithm deals with buyers who discount their long-term utilities. Note that HO-SERPKanoria and Nazerzadeh (2021) and SCORP Golrezaei et al. (2018) set *personalized reserve prices* for each buyer, whereas NPAC-S sets a single reserve for all buyers. PELS in Drutsa (2020) learns a non-linear contextual valuation model and hence yields larger regret. Among all algorithms, only SCORP Golrezaei et al. (2018) handles heterogeneous β across buyers.

Algorithm	# buyers	Context	Noise/value dist.	Discount util.	Regret
Phased Amin et al. (2013)	1	False	Lipschitz	True	Sublinear[5]
LEAP Amin et al. (2014)	1	True	No additive noise	True	$\mathcal{O}(T^{2/3})$
PELS Drutsa (2020)	1	True	No additive noise	True	$\mathcal{O}(T^{d/(d+1)})$
HO-SERP Kanoria and Nazerzadeh (2021)	≥ 2	True	MHR	False	$\mathcal{O}(\sqrt{T})$
SCORP Golrezaei et al. (2018)	≥ 2	True	MHR	True	$\mathcal{O}(T^{2/3})$
NPAC-S (this work)	≥ 2	True	Non-parametric	True	$\mathcal{O}(\sqrt{T})$

repeated auctions, and thereby the expected cumulative utility of each bidder is exactly Eq. (1). It is worth noting that Amin et al. (2013) showed, in the case of a single buyer, it is not possible to obtain a no-regret policy when $\eta = 1$, that is, when the buyer is as patient as the seller.

Furthermore, we assume buyers corrupt their true valuations in an additive manner:

$$\forall i \in [N], t \in [T] \quad b_{i,t} = v_{i,t} - a_{i,t} \ \text{ where } \ |a_{i,t}| \leq a_{\max} .$$

Here, $a_{i,t}$ is called the degree of corruption, and we refer to the buyer behavior of submitting a bid $b_{i,t} \neq v_{i,t}$ (i.e., $a_{i,t} \neq 0$) as "corrupted bidding". Note that when $a_{i,t} > 0$, the buyer shades her bid, and when $a_{i,t} < 0$, the buyer overbids. Essentially, a buyer i's strategic behavior is equivalent to deciding on a non-zero value of $a_{i,t}$. In this work, we impose no restrictions on the degree of corruption $a_{i,t}$ for a buyer i in period t other than it is bounded. [8]

3 BENCHMARK AND SELLER'S REGRET

The seller's revenue in period $t \in [T]$ is the sum of total payments from all buyers, and the expected revenue given context $x_t \in \mathcal{X}$ and reserve price r_t is

$$\mathsf{rev}_t(r_t) := \mathbb{E}\Big[\sum_{i\in[N]} p_{i,t} \;\Big|\; x_t, r_t\Big], \qquad (2)$$

$$\text{where } \ p_{i,t} \ = \ \max\{b_t^-, r_t\}\mathbb{I}\{b_{i,t} \geq \max\{b_t^+, r_t\}\} .$$

Here, b_t^- and b_t^+ are the second-highest and highest bids in period t, respectively; the expectation is taken with respect to the noise distribution in period t and any randomness in the reserve price r_t as well as bid values submitted by buyers in period t (as buyers' bidding strategies may be randomized).

[8] A bound for the degree of corruption is natural as buyers always submit non negative bids and all bids are bounded by $v_{\max}$.

The seller's objective is to maximize his expected revenue over a fixed time horizon T through optimizing contextual reserve prices r_t for any $t \in [T]$. To evaluate any seller pricing policy, we compare its total revenue against that of a benchmark policy run by a clairvoyant seller who knows the mean vector β and the non-parametric noise distribution F. This clairvoyant seller's benchmark policy sets the "optimal" contextual reserve price in each period to obtain the maximum achievable revenue $\max_r \mathsf{rev}_t(r)$ in each period, and hence facing such a seller there will be no incentive for buyers' to corrupt their bids. To provide a more formal definition for the revenue of the clairvoyant seller as well as "optimality" in contextual reserve prices, we rely on the following proposition that characterizes the seller's conditional expected revenue when buyers bid truthfully.

Proposition 1 (Seller's Revenue with Truthful Buyers)**.** *Consider the case of $N \geq 2$ buyers who bid their true valuations, i.e., $v_{i,t} = b_{i,t}$ for any $i \in [N]$ and $t \in [T]$. Conditioned on the reserve price r_t and the current context $x_t \in \mathcal{X}$, the seller's single period expected revenue in Equation (2) is*

$$\int_{-\infty}^{\infty} z dF^-(z) + \langle \beta, x_t \rangle + \int_0^{r_t} F^-(z - \langle \beta, x_t \rangle) dz - r_t \left(F^+(r_t - \langle \beta, x_t \rangle)\right) , \qquad (3)$$

where for any $z \in \mathbb{R}$, $F^-(z) := NF^{N-1}(z) - (N-1)F^N(z)$ and $F^+(z) := F^N(z)$.

The proof for this proposition is detailed in Appendix B. In Proposition 1, $F^+(\cdot)$ and $F^-(\cdot)$ are the cumulative distribution functions of $\epsilon_t^+ := v_t^+ - \langle \beta, x_t \rangle$ and $\epsilon_t^- := v_t^- - \langle \beta, x_t \rangle$ respectively, where v_t^+ and v_t^- are the highest and second highest valuations in period $t \in [T]$.

In light of Proposition 1, we define the benchmark policy of the clairvoyant seller as followed,

Definition 1 (Benchmark Policy)**.** *The benchmark policy knows the mean vector β and noise distribution F, and sets* *the reserve price for a context vector $x \in \mathcal{X}$ as*

$$r^\star(x) = \arg\max_{y \geq 0} \int_0^y F^-(z - \langle \beta, x\rangle)dz - y\left(F^+(y - \langle \beta, x\rangle)\right) . \quad (4)$$

Therefore, the benchmark reserve price in period t, denoted by $r_t^\star$, is $r^\star(x_t)$, and the corresponding optimal revenue, denoted by $\text{REV}_t^\star$, is equal to

$$\int_{-\infty}^{\infty} z dF^-(z) + \langle \beta, x_t\rangle + \int_0^{r^\star(x_t)} F^-(z - \langle \beta, x_t\rangle)dz - r^\star(x_t)\left(F^+(r^\star(x_t) - \langle \beta, x_t\rangle)\right) .$$

Remark 1. *When distribution F satisfies the MHR assumption, the objective function of the optimization problem in Equation (4) is unimodal in the decision variable y, and according to Golrezaei et al. (2018), $r^\star(x)$ can be simplified as follows: $r^\star(x) = \arg\max_{y\geq 0} y(1 - F(y - \langle \beta, x\rangle))$. In words, the MHR assumption decouples the reserve price optimization problem for multiple agents to the much simpler monopolistic pricing for each individual agent.*

We observe this benchmark provides an optimal mapping from the feature vector x_t to reserve price $r^\star(x_t)$, which remains unchanged over time as the mean vector β and noise distribution F are time-invariant. This echoes our earlier point that pricing is challenging in our contextual setting since we would need to approximate or learn the optimal mapping $r^\star(\cdot)$, whereas in non-contextual environments it is sufficient to learn a single optimal reserve price value.

We now proceed to define the regret of a policy π (possibly random) when the regret is measured against the benchmark policy. Suppose that in any period t, policy π selects reserve price r_t^π. Then, the regret of policy π in period t and its cumulative T-period regret are defined as:

$$\mathsf{Regret}^\pi(T) = \sum_{t\in[T]} \mathbb{E}\left[\text{REV}_t^\star - \mathsf{rev}_t(r_t^\pi)\right] , \quad (5)$$

where the optimal revenue $\text{REV}_t^\star$ is given in Definition 1, and the expectation is taken with respect to the context distribution $\mathcal{D}$ as well as the possible randomness in the actual reserve price r_t^π. Our goal is to design a policy that obtains a low regret for any β, F, and context distribution $\mathcal{D}$.

4 THE NPAC-S POLICY

In this section, we first propose a policy called *Non-Parametric Contextual Policy against Strategic Buyers* (NPAC-S) to maximize seller's expected revenue in our strategic setting. Then, we provide insights into how our design in NPAC-S makes the policy robust to buyer strategic behavior, and in turn allows the policy to learn the mean vector β and noise distribution F efficiently. Finally, we present theoretical regret guarantees for NPAC-S against the clairvoyant benchmark described in Definition 1 that sets the optimal contextual reserve price defined in Equation (4).

The NPAC-S policy. The detailed NPAC-S policy is shown in Algorithm 1, and consists of three main components. *(i) Phased Structure:* NPAC-S partitions T into consecutive phases, where each phase $\ell \geq 1$, denoted as E_ℓ, has length $T^{1-2^{-\ell}}$. This implies $|E_1| = \sqrt{T}$ and $|E_\ell|/\sqrt{|E_{\ell-1}|} = \sqrt{T}$. Here, we can establish that the total number of phases can be upper bounded by $\lceil \log_2(\log_2(T))\rceil + 1$. *(ii) Estimation for β, F^- and F^+:* At the end of each phase, NPAC-S uses the submitted bids from the pervious phase and employs Ordinary Least Squares (OLS) and empirical distributions to estimate the mean vector β as well as F, respectively. *(iii) Random isolation:* NPAC-S incorporates random isolation periods in which a single buyer is chosen at random, and the item is auctioned to this isolated buyer (i.e. the seller only considers the bid of the isolated buyer and ignores bid from other buyers).[9] Note that when a buyer i is isolated, the buyer wins the item if and only if his bid is greater than the reserve price, and pays the reserve price if he wins. Here, the seller's pricing policy is announced to all buyers (at $t = 0$) so that buyers examine the policy and have the freedom to adopt any bidding strategy to maximize their long term discounted utility.

Remark 2. *Here, we comment on how one can solve the reserve price optimization problem in Equation (7). The key observation is that for any period t, $\widehat{F}_\ell(\cdot)$ is a step function with jumps at points in the finite set $\mathcal{C}_\ell := \{b_{i,\tau} - \langle \widehat{\beta}_\ell, x_\tau\rangle\}_{i\in[N],\tau\in E_{\ell-1}}$. This implies that in order to solve for r_t in Equation (7), it suffices to conduct a grid search for $\forall y \in \mathcal{C}_\ell$. More specifically, we let $\{z^{(0)}, z^{(1)}, \ldots z^{(M)}\}$ be the ordered list (in increasing order) of all elements in $\mathcal{C}_\ell \cup \{0\}$, where $z^{(0)} := 0$ and $M := |\mathcal{C}_\ell|$ (here, we assumed that $0 \notin \mathcal{C}_\ell$ without loss of generality). Hence, r_t is equal to*

$$\arg\max_{m\in[M]} \sum_{j=1}^{m} \widehat{F}_\ell^-(z^{(j)} - \langle \widehat{\beta}_\ell, x_t\rangle) \cdot (z^{(j)} - z^{(j-1)}) - z^{(m)} \widehat{F}_\ell^+(z^{(m)} - \langle \widehat{\beta}_\ell, x_t\rangle) .$$

This shows that the complexity to solve Equation (7) is $\mathcal{O}(M^2)$. More detailed discussions and efficient algorithms regarding related problems can be found in Mohri and Medina (2016).

[9] The seller discloses her commitment to the random isolation protocol to all buyers at $t = 0$, and it is not necessary for the seller to reveal, during an isolation period, which buyer is being isolated.

[a] For a matrix A, $A^\dagger$ represents its pseudo inverse, so if A is invertible, we have $A^\dagger = A^{-1}$. In Lemma 4 of Appendix C, we show that with high probability $\sum_\tau x_\tau x_\tau^\top$ is positive definite, and hence invertible.

Algorithm 1: Non-Parametric Contextual Policy against Strategic Buyers (NPAC-S)

Initialize $\widehat{\beta}_1 = 0$, and $\widehat{F}_1^-(z) = \widehat{F}_1^+(z) = 0$ for $\forall z \in \mathbb{R}$.
for phase $\ell \geq 1$ **do**
  for $t \in E_\ell$ **do**
    Isolation: With probability $1/|E_\ell|$, choose one buyer uniformly at random and offer price

$$r_t^u \sim \text{Uniform}(0, v_{\max})\,. \tag{6}$$

    No Isolation: With probability $1 - 1/|E_\ell|$, set reserve price for all buyers as

$$\begin{aligned}\widehat{r}_t = \arg\max_{y\in[0,v_{\max}]} \int_0^y \widehat{F}_\ell^-(z - \langle\widehat{\beta}_\ell, x_t\rangle)dz \\ - y\cdot \widehat{F}_\ell^+(y - \langle\widehat{\beta}_\ell, x_t\rangle)\,.\end{aligned} \tag{7}$$

    Observe all bids $\{b_{i,t}\}_{i\in[N]}$
  end for
  Update estimate of the mean vector β**:** [a]

$$\widehat{\beta}_{\ell+1} = (\sum_{\tau\in E_\ell} x_\tau x_\tau^\top)^\dagger \cdot (\sum_{\tau\in E_\ell} x_\tau \bar{b}_\tau)\,, \tag{8}$$

  where $\bar{b}_\tau = \frac{1}{N}\sum_{i\in[N]} b_{i,\tau}$.
  Update the estimate of F^+ **and** F^-**:**

$$\begin{aligned}\widehat{F}_{\ell+1}^-(z) &= N\widehat{F}_{\ell+1}^{N-1}(z) - (N-1)\widehat{F}_{\ell+1}^N(z)\\ \widehat{F}_{\ell+1}^+(z) &= \widehat{F}_{\ell+1}^N(z)\,.\end{aligned} \tag{9}$$

  where $\widehat{F}_{\ell+1}(z)$ is defined as

$$\widehat{F}_{\ell+1}(z) = \frac{1}{N|E_\ell|}\sum_{\tau\in E_\ell}\sum_{i\in[N]} \mathbb{I}(b_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z), \tag{10}$$

end for

Motivation for design of NPAC-S. Here we provide some insights into the design of the NPAC-S policy, particularly the phased structure and the incorporation of random isolation periods.

Due to the phased structure of the algorithm, our estimates for β, F^-, and F^+ only depend on the bids and contextual features in the previous phase. Thus, corrupted bids submitted by buyers in past periods will have no impact on future estimates as well as pricing decisions. One can think of this as erasing all memory prior to the previous phase and restarting the algorithm, which can potentially reduce buyers' manipulating power on our estimates and reserve prices.

We now discuss the impact of having isolation periods. As all buyers are aware of the randomized isolation protocol, the presence of isolation periods restricts buyers from significantly corrupting their bids too often as by doing so they may suffer a substantial utility loss when they are isolated. To illustrate this point with an example, compare the following scenarios: (i) if there are no isolation periods, a buyer having the lowest valuation among all buyers may submit a bid by adding large corruption, but still ending up not being the second highest or highest bidder. Assuming that other buyers bid truthfully, such a scenario will not lead to any changes in utility of any buyer, but introduces a large outlier to the set of data points used in our estimations. In words, when no isolation occurs, buyers may be able to distort the seller's learning process without facing unfavorable consequences; (ii) during an isolation period when a buyer is isolated, corrupting her bid may perhaps result in significant utility loss, e.g., losing the item by underbidding when her true valuation is greater than the reserve price, or winning the item by overbidding when her true valuation is less than the reserve price. Therefore, randomized isolation incentivizes utility-maximizing buyers to reduce the frequency of corrupting their bids. Mathematically, we characterize this statement in the following Lemma 1.

Lemma 1 (Bounding number of significantly corrupted bids)**.** *For $i \in [N]$ and phase $\ell \geq 1$ define*

$$\begin{aligned}\mathcal{S}_{i,\ell} &:= \left\{t \in E_\ell : |a_{i,t}| \geq \frac{1}{|E_\ell|}\right\}\\ L_\ell &:= \log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)/\log(1/\eta)\,,\end{aligned} \tag{11}$$

where $\mathcal{S}_{i,\ell}$ is the set of all periods in phase E_ℓ during which buyer i significantly corrupts his bids. Then, we have $\mathbb{P}(|\mathcal{S}_{i,\ell}| > L_\ell) \leq 1/|E_\ell|$.

The proof of this lemma is shown in Appendix C.1.

Bounding the regret of NPAC-S. Here, we first present the regret of NPAC-S. Then we introduce several key results that are crucial to proving the regret bound of NPAC-S and also comment on how they resolve challenges that arise due to buyers' strategic behavior.

Theorem 1 (Regret of NPAC-S Policy)**.** *Suppose that the length of the horizon $T \geq \max\{\left(\frac{8x_{\max}^2}{\lambda_0^2}\right)^4, 9\}$ where λ_0^2 is the minimum eigenvalue of covariance matrix Σ. Then, in the strategic setting, the T-period regret of the NPAC-S policy is in the order of $\mathcal{O}\Big(c_f\sqrt{dN^3\log(T)}\cdot\log(\log(T))\left(\sqrt{T} + \frac{\sqrt{N^3\log(T)}T^{\frac{1}{4}}}{\log(1/\eta)}\right)\Big)$, where regret is computed against the benchmark policy in Definition 1 that knows the mean vector β and noise distribution F. Here, recall $c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z) > 0$ where f is the the pdf of F.*

Remark 3. *The proof of this theorem is presented in Appendix C. In the regret of NPAC-S, the factor $1/\log(1/\eta)$ serves as a worse case guarantee for the amount of corruption that buyers' can apply to their bids throughout the entire horizon T. As buyers get less patient, i.e., as η decreases, buyers are less willing to forgo current utility in the current period. Thus, in the presence of randomized* *isolation periods, impatient buyers are less likely to significantly corrupt bids, which translates into lower regret. The* $\log(\log(T))$ *factor corresponds to the information loss due to the policy's phased structure, which "restarts" the algorithm at the beginning of each of* $\mathcal{O}(\log(\log(T)))$ *phases and relies only on the information of the previous phase.*

The regret of NPAC-S can be decomposed into two parts: (i) the estimation errors in β, F^- and F^+, which result in the posted reserve price r_t deviating from the optimal reserve price $r_t^\star$, and hence incur a revenue loss compared to the clairvoyant benchmark; and (ii) the revenue loss due to *allocation mismatch* in the auction outcome because of buyers' strategic bidding behaviour. Here, allocation mismatch refers to the phenomenon where a bidder would have won (lost) the auctioned item had she bid truthfully, but instead lost (won) the item as she submitted a corrupted bid in reality.

We first comment on several challenges with respect to bounding the estimation errors in β, F^- and F^+. First, the OLS estimator and empirical distributions to estimate the mean vector and distributions F^- and F^+, respectively are extremely vulnerable to corrupted data (outliers), and hence standard high probability bounds are invalid for our setting. Additionally, there exists a complication in terms of bounding the estimation errors in F^- and F^+ because estimation errors for β will further propagate into the estimation errors in F and consequently impacting the estimates for F^- and F^+. To illustrate this point, consider the ideal scenario where all bids are truthful (i.e. $v_{i,t} = b_{i,t}$ for all $i \in [N]$ and $t \in [T]$). Even in this scenario, the terms $v_{i,\tau} - \langle \widehat{\beta}_\ell, x_\tau \rangle$ in the expressions for $\widehat{F}_\ell(\cdot)$ are not realizations of $\epsilon_{i,\tau}$ due to estimation errors in the mean vector $\widehat{\beta}_\ell$. Hence, the estimate $\widehat{F}_\ell(\cdot)$ evaluated at any point $z \in \mathbb{R}$ is biased, i.e. $\mathbb{E}[\widehat{F}_\ell(z - \langle \widehat{\beta}_\ell, x_t \rangle)] \neq F(z - \langle \widehat{\beta}_{\ell+1}, x_t \rangle)$. Furthermore, the estimates $\widehat{F}_\ell^+(\cdot)$ and $\widehat{F}_\ell^-(\cdot)$ are evaluated at points which may be random variables since $\widehat{\beta}_\ell$ is a random variable that depends on the history of the previous phase.

In light of such challenges in bounding estimation errors, as one of our main contributions, the following Lemma 2 provides good estimation error guarantees for β, F^- and F^+ in the presence of corrupted bids and the aforementioned error propagation phenomena.

Lemma 2 (Bounding estimation errors in β, F^- and F^+)**.** *For any phase* E_ℓ*, with probability at least* $1 - \Theta(1/|E_\ell|)$*, the following events hold: (i)* $\|\widehat{\beta}_{\ell+1} - \beta\|_1 = \mathcal{O}\big(\frac{1}{\sqrt{|E_\ell|}} + \frac{\log(|E_\ell|)}{\log(1/\eta)|E_\ell|}\big)$*; (ii) for any* $z \in \mathbb{R}$, $|\widehat{F}_{\ell+1}^-(z) - F^-(z)| = \mathcal{O}\big(\frac{N^2}{\sqrt{|E_\ell|}} + \frac{N^2\log(|E_\ell|)}{\log(1/\eta)|E_\ell|}\big)$ *and* $|\widehat{F}_{\ell+1}^+(z) - F^+(z)| = \mathcal{O}\big(\frac{N}{\sqrt{|E_\ell|}} + \frac{N\log(|E_\ell|)}{\log(1/\eta)|E_\ell|}\big)$*. Here, recall the discount factor* $\eta \in (0,1)$.

We refer readers to Lemma 4 and Lemma 5 in Appendix C.3 for more detailed statements on our high probability bounds regarding estimation errors in β, F^- and F^+.

In addition to inaccurate estimates for β, F^- and F^+, the allocation mismatch phenomenon due to strategic bidding also contributes to the regret of NPAC-S. For example, suppose that the highest valuation is greater than the reserve price. In that case, if buyers were truthful, the item would be allocated and the seller would gain positive revenue. Now, if buyers shade their bids, the auctioned item may not get allocated, resulting in zero revenue for the seller. In the following Lemma 3, we show that the number of allocation mismatch periods for each buyer is bounded with high probability.

Lemma 3 (Bounding allocation mismatch periods)**.** *Define the following two sets of time periods:*

$$\begin{aligned}
&\mathcal{B}_{i,\ell}^s = \{t \in E_\ell : v_{i,t} \geq D_t\ ,\ b_{i,t} \leq D_t\} \quad \textit{and} \\
&\mathcal{B}_{i,\ell}^o = \{t \in E_\ell : v_{i,t} \leq D_t\ ,\ b_{i,t} \geq D_t\} \\
&\textit{where } D_t = \max\{b_{-i,t}^+, \widehat{r}_t\}\,.
\end{aligned} \tag{12}$$

Here, $b_{-i,t}^+$ *is the highest among all bids excluding that submitted by buyer* i*, and* $\widehat{r}_t$ *is the reserve price offered to all buyers if no isolation occurs (defined in Equation (7)). Then, for* $\mathcal{B}_{i,\ell} := \mathcal{B}_{i,\ell}^s \cup \mathcal{B}_{i,\ell}^o$*, we have* $\mathbb{P}(|\mathcal{B}_{i,\ell}| \leq 2L_\ell + 4c_f + 8\log(|E_\ell|)) \geq 1 - 4/|E_\ell|$*, and* L_ℓ *is defined in Equation (11). Here, the probability is taken with respect to the randomness in* $\{(x_\tau, \epsilon_{i,\tau}, a_{i,\tau})\}_{\tau \in E_\ell, i \in [N]}$.

Note that $\mathcal{B}_{i,\ell}^s$ represents the set of all periods in phase ℓ during which buyer i should have won the item if she bid truthfully, but in reality lost due to shading her bid (i.e. allocation mismatch due to shading), while similarly $\mathcal{B}_{i,\ell}^o$ is the periods of allocation mismatch due to overbidding. Therefore, $\mathcal{B}_{i,\ell} := \mathcal{B}_{i,\ell}^s \cup \mathcal{B}_{i,\ell}^o$ can be interpreted as the set of all periods in phase ℓ when an allocation mismatch occurs for buyer i. The detailed proof is provided in Appendix C.2.

NPAC-S against Truthful Buyers. Here, we make a remark that in a hypothetical world where buyers are truthful (i.e. $v_{i,t} = b_{i,t}$ or equivalently the degree of corruption $a_{i,t} = 0$ for all $i \in [N]$, $t \in [T]$), our proposed NPAC-S policy achieves a regret of $\mathcal{O}(c_f\sqrt{dN^3T\log(T)} \cdot \log\log(T))$ compared to the clairvoyant benchmark policy in Definition 1. Intuitively, this is easy to see because the set of all periods in phase E_ℓ during which a buyer i significantly corrupts his bids, namely $\mathcal{S}_{i,\ell}$ defined in Lemma 1, will be empty. As a result, there will be no allocation mismatch periods, and the $1/\log(1/\eta)$ terms in the estimation errors in β, F^-, F^+ will vanish (see Lemma 2). The proof for the regret bounds of NPAC-S against truthful buyers would thus be a simplification to the proof of Theorem 1, and hence will be omitted.

5 NUMERICAL STUDY

Here, we present numerical simulations to compare the performance of NPAC-S with several baseline seller policies. In particular, consider the following baseline policies: (i) NAIVE which always sets a 0 reserve price; (ii) CONTHEDGE which runs an independent version of the HEDGE algorithm for every distinct context vector (see an introduction of HEDGE for the adversarial multi-arm bandit problem in Auer et al. (1995)). The "arms" of HEDGE correspond to potential reserve price options. Note that HEDGE is a special case of the well-known EXP3 algorithm which is a simple off-the-shelf algorithm that not only has good theoretical guarantees, but has also been applied (or its variations/generalizations have been adopted) in many areas in online advertising (see e.g. Zimmert and Seldin (2019); Balseiro and Gur (2019); Han et al. (2020)). (iii) HO-SERP, which sets personalized reserve prices for each buyer using "rolling window" estimates of β and F w.r.t other buyers' submitted bids (see Kanoria and Nazerzadeh (2021)). Here we consider HO-SERP as a baseline because among all seller algorithms in related works that study pricing in a contextual, stochastic, and strategic buyer setting similar to ours (see Table 1), HO-SERP achieves nearly the best theoretical performance. Note HO-SERP requires the noise distribution to be MHR.

To model buyers' strategic behavior, instead of restricting buyers to bid according to a specific strategy to maximizes long-term discounted utility, we instead mimic the outcome of some general class of such strategies (parameterized by η) via randomly selecting periods over the entire horizon and have buyers significantly corrupt bids in these periods. We will refer to these randomly selected periods as *corruption periods*. When this randomization procedure is repeated over many trials, we believe the average bidding outcome would serve as a relatively accurate approximation to the outcomes of a general class of strategies for utility-discounting buyers. Furthermore, inspired by Lemma 1 which suggests that the number of periods when a buyer significantly corrupts her bid is bounded, we let the selected number of corruption periods be L_ℓ defined in Equation (11) . Note that L_ℓ is increasing in η and represents the fact that more patient buyers (i.e. larger η) value long term utility more and hence would be willing to corrupt bids more frequently with the aim of achieving higher future utility.

Out detailed experimental setup is as followed. We consider a horizon of length $T = 5,000$, $N = 2$ buyers, context vectors of dimension $d = 4$, $v_{\max} = 10$ and $v_{\min} = 0$. For each $\eta \in \{0.2, 0.4, 0.6, 0.8\}$, repeat the following procedure for $n = 50$ trials, each including T periods:

For each phase E_ℓ ($\ell \geq 1$), [10] sample L_ℓ corruption periods uniformly at random. Then, regarding buyer's valuations, we generate $\beta \in [0,1]^d$, where each entry is sampled independently according to a uniform distribution on [0,1], i.e., $U(0,1)$. We further normalize β with the sum of all entries so that $\|\beta\|_2 = 1$. We then generate 10 distinct contexts vectors $\mathcal{X} = \{X^j\}_{j\in[10]}$, where each entry for any distinct context vector is sampled independently from $U\left(\frac{v_{\max}}{3}, \frac{2v_{\max}}{3}\right)$. Then, for every period $t \in [T]$, sample x_t uniformly at random from $\mathcal{X}$, and sample $\epsilon_{i,t}$ for all $i \in [N]$ independently from $U\left(-\frac{v_{\max}}{3}, \frac{v_{\max}}{3}\right)$. Note that our construction guarantees $v_{i,t} = \langle \beta, x_t \rangle + \epsilon_{i,t} \in [v_{\min}, v_{\max}]$, and the noise distribution is uniform which satisfies the MHR assumption (so the application of the HO-SERP is valid). If t is a corruption period, we let buyers submit a bid of value 0 to model the behavior of significant bid-shading; otherwise, we let buyers bid their true valuations.[11]

For comprehensiveness, we also consider the truthful setting by repeating the above valuation generation procedure for another $n = 50$ trials and have buyers always submit their true valuations. Finally, for each of the aforementioned trials, we run the NPAC-S as well as other baseline algorithms independently and simply record the realized revenue of each algorithm across all repeated auctions.

We report the average per-period revenue loss compared to the benchmark policy (Definition 1) for each algorithm in Figure 1.

We observe that our proposed NPAC-S algorithm not only outperforms CONTHEDGE in all settings consistently by a $3\% \sim 4\%$ and NAIVE in the truthful setting by $6\% \sim 7\%$, NPAC-S also generally yields more stable outcomes as measured by the standard deviation of per-period revenue loss across n trials. Compared to HO-SERP, NPAC-S slightly outperforms HO-SERP in the truthful setting and for $\eta =$ 0.2, 0.4, 0.6. Nevertheless, we point out that our experimental setting inherently favors HO-SERP since performance guarantees of this algorithm relies on the noise distribution being MHR, which is the case for our uniform noise. Moreover, the comparison with HO-SERP also demonstrates the advantages of NPAC-S from a practical viewpoint, since NPAC-S, unlike HO-SERP, sets a single reserve price for all buyers and still matches or improves upon the performance of HO-SERP.

[10] For fixed T, since length of phase $\ell \geq 1$ is $T^{1-2^{-\ell}}$, in our case when $T = 5,000$ we have 4 phases whose phase lengths are $70, 594, 1724, 2612$, respectively, where the last phase is truncated.

[11] We remark that our numerical experiments focus on buyers' bid-shading behavior. This is mainly because empirical studies found that shading is prevalent in repeated auctions on modern online advertising platforms and theoretical works have demonstrated various versions of bid-shading strategies can help buyers achieve near-optimal performances in a variety of practical settings, such as buyers being constrained by a limited budget or target return on investment (see e.g. Zeithammer (2007); Golrezaei et al. (2021b); Balseiro and Gur (2019)).

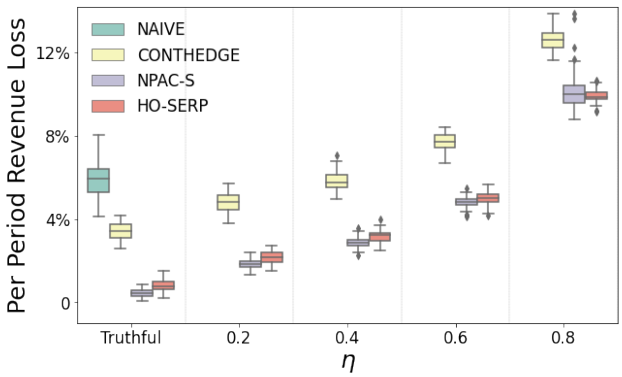

Figure 1: **Performance comparison with baselines.** This figure displays the average per-period revenue loss compared to the benchmark policy (Definition 1). Each box plot corresponds to $n = 50$ trials. NAIVE is only run for the truthful setting because buyers will have no incentive to bid untruthfully when there is no reserve price. CONTHEDGE is run with"arms" $\{0, 0.5, 1, \dots 10\}$, where each arm corresponds to a reserve price option.

References

Amin, K., Rostamizadeh, A., and Syed, U. (2013). Learning prices for repeated auctions with strategic buyers. In *Advances in Neural Information Processing Systems*, pages 1169–1177.

Amin, K., Rostamizadeh, A., and Syed, U. (2014). Repeated contextual auctions with strategic buyers. In *Advances in Neural Information Processing Systems*, pages 622–630.

Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. (1995). Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of IEEE 36th Annual Foundations of Computer Science*, pages 322–331. IEEE.

Balseiro, S. R. and Gur, Y. (2019). Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9):3952–3968.

Bastani, H. and Bayati, M. (2015). Online decision-making with high-dimensional covariates. *Available at SSRN 2661896*.

Besbes, O. and Zeevi, A. (2009). Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research*, 57(6):1407–1420.

Birge, J. R., Feng, Y., Keskin, N. B., and Schultz, A. (2018). Dynamic learning and market making in spread betting markets with informed bettors. *Available at SSRN 3283392*.

Broder, J. and Rusmevichientong, P. (2012). Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4):965–980.

Celis, L. E., Lewis, G., Mobius, M., and Nazerzadeh, H. (2014). Buy-it-now or take-a-chance: Price discrimination through randomized auctions. *Management Science*, 60(12):2927–2948.

Cesa-Bianchi, N., Gentile, C., and Mansour, Y. (2015). Regret minimization for reserve prices in second-price auctions. *IEEE Transactions on Information Theory*, 61(1):549–564.

Chen, H. and Keskin, N. B. (2018). Markdown policies for demand learning and strategic customer behavior. *Available at SSRN 3299819*.

Chen, N. and Gallego, G. (2018). Nonparametric learning and optimization with covariates.

Cohen, M., Lobel, I., and Paes Leme, R. (2016). Feature-based dynamic pricing. *Available at SSRN 2737045*.

den Boer, A. V. and Zwart, B. (2013). Simultaneously learning and optimizing using controlled variance pricing. *Management science*, 60(3):770–783.

Deng, Y., Lahaie, S., and Mirrokni, V. (2019). Robust pricing in non-clairvoyant dynamic mechanism design. *Available at SSRN*.

Drutsa, A. (2019). Reserve pricing in repeated second-price auctions with strategic bidders. *arXiv preprint arXiv:1906.09331*.

Drutsa, A. (2020). Optimal non-parametric learning in repeated contextual auctions with strategic buyer. In *International Conference on Machine Learning*, pages 2668–2677. PMLR.

Dvoretzky, A., Kiefer, J., Wolfowitz, J., et al. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, 27(3):642–669.

Fan, J., Guo, Y., and Yu, M. (2022). Policy optimization using semiparametric models for dynamic pricing. *Journal of the American Statistical Association*, pages 1–29.

Feng, Z., Parkes, D. C., and Xu, H. (2019). The intrinsic robustness of stochastic bandits to strategic manipulation. *arXiv preprint arXiv:1906.01528*.

Golrezaei, N., Jaillet, P., Liang, J. C. N., and Mirrokni, V. (2021a). Bidding and pricing in budget and roi constrained markets. *arXiv preprint arXiv:2107.07725*.

Golrezaei, N., Javanmard, A., and Mirrokni, V. (2018). Dynamic incentive-aware learning: Robust pricing in contextual auctions.

Golrezaei, N., Lin, M., Mirrokni, V., and Nazerzadeh, H. (2017). Boosted second price auctions: Revenue optimization for heterogeneous bidders.

Golrezaei, N., Lobel, I., and Paes Leme, R. (2021b). Auction design for roi-constrained buyers. In *Proceedings of the international conference on World Wide Web*.

Han, Y., Zhou, Z., Flores, A., Ordentlich, E., and Weissman, T. (2020). Learning to bid optimally and efficiently in adversarial first-price auctions. *arXiv preprint arXiv:2007.04568*.

Javanmard, A. (2017). Perishability of data: dynamic pricing under varying-coefficient models. *The Journal of Machine Learning Research*, 18(1):1714–1744.

Javanmard, A. and Nazerzadeh, H. (2016). Dynamic pricing in high-dimensions. *arXiv preprint arXiv:1609.07574*.

Kanoria, Y. and Nazerzadeh, H. (2017). Dynamic reserve prices for repeated auctions: Learning from bids. *Available at SSRN 2444495*.

Kanoria, Y. and Nazerzadeh, H. (2021). Incentive-compatible learning of reserve prices for repeated auctions. *Operations Research*, 69(2):509–524.

Kleinberg, R. and Leighton, T. (2003). The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 594–605. IEEE.

Koufogiannakis, C. and Young, N. E. (2014). A nearly linear-time ptas for explicit fractional packing and covering linear programs. *Algorithmica*, 70(4):648–674.

Leme, R. P. and Schneider, J. (2018). Contextual search via intrinsic volumes. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 268–282. IEEE.

Liu, J., Huang, Z., and Wang, X. (2018). Learning optimal reserve price against non-myopic bidders. In *Advances in Neural Information Processing Systems*, pages 2038–2048.

Lobel, I., Leme, R. P., and Vladu, A. (2018). Multidimensional binary search for contextual decision-making. *Operations Research*, 66(5):1346–1361.

Luo, Y., Sun, W. W., et al. (2021). Distribution-free contextual dynamic pricing. *arXiv preprint arXiv:2109.07340*.

Mahdian, M., Mirrokni, V., and Zuo, S. (2017). Incentive-aware learning for large markets. In *Proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee*.

McSherry, F. and Talwar, K. (2007). Mechanism design via differential privacy. In *FOCS*, volume 7, pages 94–103.

Mohri, M. and Medina, A. M. (2016). Learning algorithms for second-price auctions with reserve. *The Journal of Machine Learning Research*, 17(1):2632–2656.

Shah, V., Blanchet, J., and Johari, R. (2019). Semi-parametric dynamic contextual pricing. *arXiv preprint arXiv:1901.02045*.

Tropp, J. A. et al. (2015). An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230.

Xu, J. and Wang, Y.-X. (2022). Towards agnostic feature-based dynamic pricing: Linear policies vs linear valuation with unknown noise. In *International Conference on Artificial Intelligence and Statistics*, pages 9643–9662. PMLR.

Zeithammer, R. (2007). Research note—strategic bid-shading and sequential auctioning with learning from past prices. *Management Science*, 53(9):1510–1519.

Zimmert, J. and Seldin, Y. (2019). An optimal algorithm for stochastic and adversarial bandits. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 467–475. PMLR.

Appendices for

Incentive-aware Contextual Pricing with Non-parametric Market Noise

Our Appendix is organized as followed. In Appendix A, we include an extended literature review that discusses broader related works. Appendix B includes the all proofs of the results in Section 3. Appendix C is dedicated to Section 4. In particular, Appendix C proves Theorem 1 which shows a regret bound for the NPAC-S policy against strategic buyers.

A EXTENDED LITERATURE REVIEW

There has been a large body of literature that considers the problem of non-contextual dynamic pricing with non-strategic buyers. Kleinberg and Leighton (2003) studies repeated non-contextual posted price auctions with a single buyer whose valuations are fixed, drawn from a fixed but unknown distribution, and chosen by an adversary who is oblivious to the seller's algorithm. den Boer and Zwart (2013); Besbes and Zeevi (2009); Broder and Rusmevichientong (2012) study non-contextual dynamic pricing with demand uncertainty, where they estimate unknown model parameters using estimation techniques such as maximum likelihood. Golrezaei et al. (2021a) considers a seller repeatedly pricing against a buyer who is subject to budget and return-on-investment (ROI) constraints. Cesa-Bianchi et al. (2015) considers the dynamic pricing problem in non-contextual repeated second-price auctions with multiple buyers whose bids are drawn from some unknown and possibly non-parametric distribution. In addition, they also consider bandit feedback where the seller only observes realized revenues instead of all submitted bids. In their non-contextual setup, the seller's revenue-maximizing price is fixed throughout the entire time horizon, and the key is to approximate this optimal price by estimating the valuation distribution. In our setting, however, the optimal reserve prices are context-dependent, which means the seller is required to estimate (i) the distributional form of valuations and (ii) buyers' willingness-to-pay that varies in each period according to different contexts.

Another line of research studies the problem of contextual dynamic pricing with non-strategic buyer behavior. Cohen et al. (2016); Lobel et al. (2018); Leme and Schneider (2018) propose learning algorithms based on binary search methods when the context vector is chosen adversarially in each round. Chen and Gallego (2018) consider the problem where a learner observes contextual features and optimizes an objective by experimenting with a fixed set of decisions. Their tree-based non-parametric learning policy is designed to handle very general objectives and not specifically tailored to pricing problems. Thus, in pricing problems, its performance deteriorates as the dimension of the feature vector increases. Javanmard and Nazerzadeh (2016) also considered a contextual pricing problem with an unknown but parametric noise distribution, and uses a maximum likelihood estimator to jointly estimate the mean vector and distributions parameters. Finally, there is a recent line of work that considers contextual dynamic pricing with non-strategic buyers and non-parametric market noise. Shah et al. (2019) studied a dynamic posted price problem where the relationship between the expectation of the logarithm of buyer valuation and the contextual features is linear, while the market noise distribution is non-parametric. This logarithmic form of the valuation model allows them to separate the noise term from the context, which makes it possible to independently estimate the noise distribution and expected buyer valuation. In our setting, however, the context is embedded within the noise distribution, and our estimation errors in the mean vector β will propagate into the estimation error in the noise distribution, making the learning task more difficult, compared to that in Shah et al. (2019). On the other hand, Luo et al. (2021); Fan et al. (2022); Xu and Wang (2022) all study a repeated posted price problem with context and additive non-parametric noise similar to the buyer valuation model studied in this paper. In addition, these works consider censored feedback where the seller only observes the binary purchasing decision in each period. Our work distinguishes itself by considering a multi-buyer setting with strategic behavior.

Finally, our work is also related to the recent literature within the domain of mechanism design and online learning that adopt methodologies from differential privacy to deal with strategic agents; see, for example, McSherry and Talwar (2007); Mahdian et al. (2017); Liu et al. (2018).

B APPENDIX FOR SECTION 3: PROOF OF PROPOSITION 1

Let $Q_t(\cdot)$ be the distributions of a buyer's valuation when we condition on the feature vector x_t. Further, let $Q_t^-(\cdot)$ be the distribution of v_t^-, which is the second highest valuation at time t. Then, we have $Q_t(z) = F(z - \langle\beta, x_t\rangle)$ and $Q_t^-(z) = F^-(z - \langle\beta, x_t\rangle)$. When $N \geq 2$ and all buyers bid truthfully, according to Equations (2) , the seller's expected revenue conditioned on x_t by setting reserve price r_t is:

$$\begin{aligned}\mathsf{rev}_t(r_t) &= \mathbb{E}\left[\max\{r_t, v_t^-\}\mathbb{I}\{v_t^+ \geq r_t\} \mid x_t, r_t\right] \\ &= \mathbb{E}\left[r_t \mathbb{I}\{v_t^+ \geq r_t \geq v_t^-\} + v_t^- \mathbb{I}\{v_t^+ \geq v_t^- \geq r_t\} \mid x_t, r_t\right] ,\end{aligned} \tag{13}$$

where v_t^+ is the highest valuation at time t. The first term within the expectation, conditioned on x_t and r_t, is

$$\mathbb{E}\left[r_t \mathbb{I}\{v_t^+ \geq r_t \geq v_t^-\} \mid x_t, r_t\right] = r_t N \left[Q_t(r_t)\right]^{N-1}\left[1 - Q_t(r_t)\right], \tag{14}$$

where we used the fact that r_t is independent of v_t^+ and v_t^- since the seller sets reserve price r_t based on only the past history $\mathcal{H}_{t-1} = \{(r_1, v_1, x_1), (r_2, v_2, x_2), \ldots, (r_{t-1}, v_{t-1}, x_{t-1})\}$, and both v_t^+ and v_t^-, conditioned on x_t, are independent of the past. The second term within the expectation of Equation (13) is

$$\begin{aligned}\mathbb{E}\left[v_t^- \mathbb{I}\{v_t^+ \geq v_t^- \geq r_t\} \mid x_t, r_t\right] &= \mathbb{E}\left[v_t^- \mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] \\ &= \mathbb{E}\left[(v_t^- - r_t)\mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] + r_t \mathbb{E}\left[\mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] \\ &= \int_0^\infty \mathbb{P}\left(v_t^- - r_t \geq z\right) dz + r_t\left[1 - Q_t^-(r_t)\right] \\ &= \int_{r_t}^\infty \left[1 - Q_t^-(z)\right] dz + r_t\left[1 - Q_t^-(r_t)\right] \\ &= \mathbb{E}\left[v_t^- \mid x_t, r_t\right] - \int_0^{r_t}\left[1 - Q_t^-(z)\right] dz + r_t\left[1 - Q_t^-(r_t)\right] \\ &= \mathbb{E}\left[v_t^- \mid x_t\right] + \int_0^{r_t} Q_t^-(z) dz - r_t Q_t^-(r_t) .\end{aligned} \tag{15}$$

Note that the integration starts from 0 because all valuations are considered to be positive. Since $F^-(\tilde{z}) := NF^{N-1}(\tilde{z}) - (N-1)F^N(\tilde{z})$ for any $\tilde{z} \in \mathbb{R}$, we have

$$Q_t^-(r_t) = N\left[Q_t(r_t)\right]^{N-1}\left[1 - Q_t(r_t)\right] + \left[Q_t(r_t)\right]^N . \tag{16}$$

Hence, combining Equations (13), (14), (15), and (16), we have

$$\begin{aligned}\mathsf{rev}_t(r_t) &= \mathbb{E}\left[v_t^- \mid x_t\right] + \int_0^{r_t} Q_t^-(z) dz - r_t\left[Q_t(r_t)\right]^N \\ &= \mathbb{E}\left[v_t^- \mid x_t\right] + \int_0^{r_t} F^-(z - \langle\beta, x_t\rangle) dz - r_t\left[F^+(r_t - \langle\beta, x_t\rangle)\right] \\ &= \int_{-\infty}^\infty z dF^-(z) + \langle\beta, x_t\rangle + \int_0^{r_t} F^-(z - \langle\beta, x_t\rangle) dz - r_t\left[F^+(r_t - \langle\beta, x_t\rangle)\right] .\end{aligned}$$

□

C APPENDIX FOR SECTION 4: PROOF OF THEOREM 1

We first introduce some definitions that we will extensively rely on throughout our proof of Theorem 1. We start off with the "good" events $\xi_{\ell+1}$, $\xi_{\ell+1}^-$ and $\xi_{\ell+1}^+$ for $\ell \geq 1$ in which the estimates of β, F^- and F^+ are accurate:

$$\xi_{\ell+1} = \left\{\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}}\right\} \tag{17}$$

$$\text{where } \delta_\ell := \frac{\sqrt{2d\log(|E_\ell|)\epsilon_{\max} x_{\max}^2}}{\lambda_0^2\sqrt{N|E_\ell|}} + \frac{\sqrt{d}\left(NL_\ell a_{\max} + 1\right) x_{\max}^2}{|E_\ell|\lambda_0^2} , \tag{18}$$

$$\xi_{\ell+1}^- = \left\{\left|\widehat{F}_{\ell+1}^-(z) - F^-(z)\right| \leq 2N^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + NL_\ell}{|E_\ell|}\right)\right\} , \tag{19}$$

$$\xi_{\ell+1}^+ = \left\{\left|\widehat{F}_{\ell+1}^+(z) - F^+(z)\right| \leq N\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + NL_\ell}{|E_\ell|}\right)\right\} , \tag{20}$$

where $a_{\max}$ is the maximum possible corruption, $\gamma_\ell = \sqrt{\log(|E_\ell|)}/\sqrt{2N|E_\ell|}$, λ_0^2 is the minimum eigenvalue of covariance matrix Σ, and $c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z) \geq \inf_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z) > 0$. Furthermore,

$$L_\ell = \frac{\log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)}{\log(1/\eta)} = \mathcal{O}\left(\frac{\log(|E_\ell|)}{\log(1/\eta)}\right),$$

where $|E_\ell| = T^{1-2^{-\ell}}$ is the length of the ℓ^{th} phase.

We also define the event that the number of periods in phase E_ℓ during which buyer i submits significantly corrupted bids is bounded by L_ℓ:

$$\mathcal{G}_{i,\ell} := \{|\mathcal{S}_{i,\ell}| \leq L_\ell\} \ . \tag{21}$$

Here, $\mathcal{S}_{i,\ell} = \left\{t \in E_\ell : |a_{i,t}| \geq \frac{1}{|E_\ell|}\right\}$ is the set of all periods in phase E_ℓ during which buyer i extensively corrupts her bids.

We are now equipped to show Theorem 1 according to the following steps:

(i) Decompose the single period regret into $\mathcal{R}_t^{(1)}$ and $\mathcal{R}_t^{(2)}$, where $\mathcal{R}_t^{(1)}$ bounds the expected revenue loss due to the discrepancy between the actual reserve price r_t and the optimal reserve price $r_t^\star$ and $\mathcal{R}_t^{(2)}$, which bounds the expected revenue loss due to allocation mismatches. Note that $\mathcal{R}_t^{(1)}$ is a result of the estimation inaccuracies in β, F^- and F^+.

(ii) Bound $\mathcal{R}_t^{(1)}$ using Lemmas 1, 4, 5, and 6.

(iii) Bound $\mathcal{R}_t^{(2)}$ using Lemmas 1 and 3.

(iv) Sum up $\mathcal{R}_t^{(1)}$ and $\mathcal{R}_t^{(2)}$ to bound the cumulative expected regret over a phase E_ℓ and the entire horizon T.

(i) Decomposing single period regret into $\mathcal{R}_t^{(1)}$ and $\mathcal{R}_t^{(2)}$: According to the NPAC-S policy detailed in Algorithm 1, the expected revenue in period t is given by

$$\mathsf{rev}_t(r_t) \ = \ \mathbb{E}\left[\max\{b_t^-, \widehat{r}_t\}\mathbb{I}\{b_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} + \sum_{i\in[N]} r_t^u \mathbb{I}\{b_{i,t} > r_t^u\}\mathbb{I}\{i \text{ is isolated}\} \ \mid x_t, r_t\right], \tag{22}$$

where the expectation is taken with respect to $\{(x_\tau, \epsilon_{i,\tau}, a_{i,\tau})\}_{\tau\in[t],i\in[N]}$ and $\widehat{r}_t, r_t^u$ are defined in Equations (6) and (7) respectively. Hence, the regret is given by

$$\begin{aligned}
\mathsf{Regret}_t \ &= \ \mathbb{E}\left[\mathrm{REV}_t^\star - \mathsf{rev}_t(r_t)\right] \\
&= \ \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \mathsf{rev}_t(r_t)\right] \\
&= \ \left(\mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right]\right) \\
&\quad + \left(\mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} - \mathsf{rev}_t(r_t)\right]\right) \\
&:= \ \mathcal{R}_t^{(1)} + \mathcal{R}_t^{(2)},
\end{aligned} \tag{23}$$

where the expectation is taken with respect the context $x_t \sim \mathcal{D}$ and the randomness in r_t; $r_t^\star$ is the optimal reserve price (defined in Equation (4)) if the seller has full knowledge of F and β; and we defined:

$$\begin{aligned}
\mathcal{R}_t^{(1)} &:= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
\mathcal{R}_t^{(2)} &:= \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} - \mathsf{rev}_t(r_t)\right]
\end{aligned} \tag{24}$$

(ii) Bounding $\mathcal{R}_t^{(1)}$: We start by upper bounding $\mathcal{R}_t^{(1)}$ for a period $t \in E_{\ell+1}$ where $\ell \geq 1$.

$$\begin{aligned}
\mathcal{R}_t^{(1)} &= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
&= \mathbb{E}\left[\left(\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right)\mathbb{I}\{\text{no isolation in } t\}\right] \\
&\quad + \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\left(1 - \mathbb{I}\{\text{no isolation in } t\}\right)\right] \\
&= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right]\left(1 - \frac{1}{|E_\ell|}\right) \\
&\quad + \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right]\cdot\frac{1}{|E_\ell|} \\
&\leq \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right] + \frac{v_{\max}}{|E_\ell|}\,, \qquad (25)
\end{aligned}$$

where the third equality is because an isolation event is independent of any other event, and the final inequality follows from a simple observation that $\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} \leq v_{\max}$.

For simplicity, we define

$$\widetilde{\mathcal{R}}_t^{(1)} := \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\;\middle|\; x_t, \widehat{r}_t\right],$$

so Equation (25) yields

$$\mathcal{R}_t^{(1)} \leq \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)}\right] + \frac{v_{\max}}{|E_\ell|}\,, \qquad (26)$$

where the expectation is taken with respect to the context x_t and reserve price $\widehat{r}_t$. Notice that $\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}$ is exactly the revenue difference $\mathsf{rev}_t(r_t^\star) - \mathsf{rev}_t(r_t)$ had the seller set reserve prices $r_t^\star$ or r_t when all buyers bid truthfully. Hence, by applying Proposition 1 we obtain

$$\widetilde{\mathcal{R}}_t^{(1)} = \int_0^{r_t^\star} F^-(z - \langle\beta, x_t\rangle)dz - r_t^\star\left[F^+(r_t^\star - \langle\beta, x_t\rangle)\right] - \int_0^{\widehat{r}_t} F^-(z - \langle\beta, x_t\rangle)dz + \widehat{r}_t\left[F^+(\widehat{r}_t - \langle\beta, x_t\rangle)\right].$$

Note that we can apply Proposition 1 because $\widehat{r}_t$ is the reserve price set according to the NPAC-S policy when no isolation occurs, and only depends on the current context x_t and the past $\mathcal{H}_{t-1} = \{(r_1, b_1, x_1), (r_2, b_2, x_2), \ldots, (r_{t-1}, b_{t-1}, x_{t-1})\}$.

By defining $y_t := \langle\beta, x_t\rangle$, $\widehat{y}_t := \langle\widehat{\beta}_\ell, x_t\rangle$ and

$$\rho_t(r, y, F^{(1)}, F^{(2)}) := \int_0^r F^{(2)}(z - y)dz - r\left[F^{(1)}(r - y)\right], \qquad (27)$$

we can rewrite $\widetilde{\mathcal{R}}_t^{(1)}$ as the following:

$$\begin{aligned}
\widetilde{\mathcal{R}}_t^{(1)} &= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\;\middle|\; x_t, \widehat{r}_t\right] \\
&= \rho_t(r_t^\star, y_t, F^-, F^+) - \rho_t(\widehat{r}_t, y_t, F^-, F^+) \\
&= \rho_t(r_t^\star, y_t, F^-, F^+) - \rho_t(r_t^\star, \widehat{y}_t, F^-, F^+) \\
&\quad + \rho_t(r_t^\star, \widehat{y}_t, F^-, F^+) - \rho_t(r_t^\star, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) \\
&\quad + \rho_t(r_t^\star, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) - \rho_t(\widehat{r}_t, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) \\
&\quad + \rho_t(\widehat{r}_t, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) - \rho_t(\widehat{r}_t, \widehat{y}_t, F^-, F^+) \\
&\quad + \rho_t(\widehat{r}_t, \widehat{y}_t, F^-, F^+) - \rho_t(\widehat{r}_t, y_t, F^-, F^+)\,.
\end{aligned} \qquad (28)$$

We now invoke Lemma 6, where we show that when events $\xi_{\ell+1}$, $\xi_{\ell+1}^-$ and $\xi_{\ell+1}^+$ (see definition in Equation (17),(18), (19) and (20)) happen for some phase $\ell \geq 1$, we have for $r \in \{r_t^\star, \widehat{r}_t\}$,

(i) $|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, F^-, F^+)| \leq 3rc_f N^2\delta_\ell$ a.s.

(ii) $\left|\rho_t(r,\widehat{y}_t,F^-,F^+)-\rho_t(r,\widehat{y}_t,\widehat{F}^-_{\ell+1},\widehat{F}^+_{\ell+1})\right| \;\leq\; 3rN^2\left(\gamma_\ell+c_f\delta_\ell+\frac{c_f+L_\ell}{|E_\ell|}\right)$ a.s.

Note that the first inequality bounds the impact of errors β and the second bounds the impact of errors in the distributions. Applying these bounds in (28), we get

$$\begin{aligned}\widetilde{\mathcal{R}}^{(1)}_t\cdot\mathbb{I}\left\{\xi_{\ell+1}\cap\xi^-_{\ell+1}\cap\xi^+_{\ell+1}\right\} \;\leq\; & 3(r^\star_t+\widehat{r}_t)c_fN^2\delta_\ell\\ &+3(r^\star_t+\widehat{r}_t)N^2\left(\gamma_\ell+c_f\delta_\ell+\frac{c_f+L_\ell}{|E_\ell|}\right)\\ &+\rho_t(r^\star_t,\widehat{y}_t,\widehat{F}^-_{\ell+1},\widehat{F}^+_{\ell+1})-\rho_t(\widehat{r}_t,\widehat{y}_t,\widehat{F}^-_{\ell+1},\widehat{F}^+_{\ell+1})\,. \end{aligned} \tag{29}$$

We recall that the seller's pricing decision $\widehat{r}_t$ when no isolation occurs is defined in Equation (7), and realize that in fact $\widehat{r}_t=\arg\max_{r\in(0,v_{\max}]}\rho_t(r,\widehat{y}_t,\widehat{F}^-_{\ell+1},\widehat{F}^+_{\ell+1})$. So, by the optimality of $\widehat{r}_t$ and $r^\star_t\leq v_{\max}$, we obtain the fact that $\rho_t(r^\star_t,\widehat{y}_t,\widehat{F}^-_{\ell+1},\widehat{F}^+_{\ell+1})-\rho_t(\widehat{r}_t,\widehat{y}_t,\widehat{F}^-_{\ell+1},\widehat{F}^+_{\ell+1})\leq 0$. Using this inequality in (29), we get

$$\begin{aligned}&\widetilde{\mathcal{R}}^{(1)}_t\cdot\mathbb{I}\left\{\xi_{\ell+1}\cap\xi^-_{\ell+1}\cap\xi^+_{\ell+1}\right\}\\ \leq\;& 6v_{\max}c_fN^2\delta_\ell+6v_{\max}N^2\left(\gamma_\ell+c_f\delta_\ell+\frac{c_f+L_\ell}{|E_\ell|}\right)\\ =\;& 12v_{\max}c_fN^2\delta_\ell+6v_{\max}N^2\left(\frac{\sqrt{\log(|E_\ell|)}}{\sqrt{2N|E_\ell|}}+\frac{c_f+L_\ell}{|E_\ell|}\right)\\ =\;& 12v_{\max}c_fN^2\delta_\ell+\frac{6v_{\max}\sqrt{N^3\log(|E_\ell|)}}{\sqrt{2E_\ell}}+\frac{6v_{\max}N^2(c_f+L_\ell)}{|E_\ell|}\,, \end{aligned} \tag{30}$$

where we used the fact that $r^\star_t,\widehat{r}_t\;\leq\;v_{\max}$ in the inequality. Note that $L_\ell\;=\;\log\left(v^2_{\max}N|E_\ell|^4-1\right)/\log(\frac{1}{\eta})\;=\;\mathcal{O}\left(\log(T)/\log(1/\eta)\right)$, since we recall that $|E_\ell|=T^{1-2^{-\ell}}$.

To complete the bound for $\mathcal{R}^{(1)}_t$ in period $t\in E_{\ell+1}$, we continue to bound Equation (26):

$$\begin{aligned}\mathcal{R}^{(1)}_t\;\leq\;&\mathbb{E}\left[\widetilde{\mathcal{R}}^{(1)}_t\right]+\frac{v_{\max}}{|E_\ell|}\\ =\;&\mathbb{E}\left[\widetilde{\mathcal{R}}^{(1)}_t\cdot\mathbb{I}\left\{\xi_{\ell+1}\cap\xi^-_{\ell+1}\cap\xi^+_{\ell+1}\right\}\right]+\mathbb{E}\left[\widetilde{\mathcal{R}}^{(1)}_t\cdot\mathbb{I}\left\{\xi^c_{\ell+1}\cup\left(\xi^-_{\ell+1}\right)^c\cup\left(\xi^+_{\ell+1}\right)^c\right\}\right]+\frac{v_{\max}}{|E_\ell|}\\ \leq\;&\mathbb{E}\left[\widetilde{\mathcal{R}}^{(1)}_t\cdot\mathbb{I}\left\{\xi_{\ell+1}\cap\xi^-_{\ell+1}\cap\xi^+_{\ell+1}\right\}\right]+v_{\max}\mathbb{P}\left(\xi^c_{\ell+1}\cup\left(\xi^-_{\ell+1}\right)^c\cup\left(\xi^+_{\ell+1}\right)^c\right)+\frac{v_{\max}}{|E_\ell|}\\ \leq\;&12v_{\max}c_fN^2\delta_\ell+\frac{6v_{\max}\sqrt{N^3\log(|E_\ell|)}}{\sqrt{2E_\ell}}+\frac{v_{\max}\left(6N^2(c_f+L_\ell)+9N+15d+9\right)}{|E_\ell|}\,, \end{aligned} \tag{31}$$

where the second inequality follows from a simple observation that $\widetilde{\mathcal{R}}^{(1)}_t\leq v_{\max}$ almost surely, and the third inequality uses Equation (30) and Lemma 7, which shows $\mathbb{P}\left(\xi^c_{\ell+1}\cup\left(\xi^-_{\ell+1}\right)^c\cup\left(\xi^+_{\ell+1}\right)^c\right)\;\leq\;(9N+15d+8)/|E_\ell|$,

(iii) Bounding $\mathcal{R}^{(2)}_t$: So far, we have bounded $\mathcal{R}^{(1)}_t$ for $t\in E_{\ell+1}$ ($\ell\geq 1$), and will move on to bound $\mathcal{R}^{(2)}_t$ defined in Equation (23) for $t\in E_\ell$ for any $\ell\geq 1$. We define

$$b^+_{-i,t}=\max_{j\neq i}b_{j,t}\quad\text{and}\quad v^+_{-i,t}=\max_{j\neq i}v_{j,t}\,, \tag{32}$$

which represent the highest bid excluding that of buyer i, and the highest valuation excluding that of buyer i, respectively.

We then have

$$
\begin{aligned}
\mathcal{R}_t^{(2)} &= \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+>\widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}-\mathsf{rev}_t(r_t)\right] \\
&\le \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+>\widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right]-\mathbb{E}\left[\max\{b_t^-,r_t\}\mathbb{I}\{b_t^+>\widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
&= \left(\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+>\widehat{r}_t\}\right]-\mathbb{E}\left[\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_t^+>\widehat{r}_t\}\right]\right)\cdot\left(1-\frac{1}{|E_\ell|}\right) \\
&< \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+>\widehat{r}_t\}\right]-\mathbb{E}\left[\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_t^+>\widehat{r}_t\}\right] \\
&= \sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t}>\max\{v_{-i,t}^+,\widehat{r}_t\}\}-\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t}>\max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&= \sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{v_{-i,t}^+,\widehat{r}_t\}<v_{i,t}<\max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\quad -\sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{b_{-i,t}^+\widehat{r}_t\}<v_{i,t}<\max\{v_{-i,t}^+,\widehat{r}_t\}\}\right] \\
&\quad +\sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t}>\max\{b_{-i,t}^+\widehat{r}_t\}\}-\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t}>\max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\le \sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{v_{-i,t}^+,\widehat{r}_t\}<v_{i,t}<\max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\quad +\sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t}>\max\{b_{-i,t}^+\widehat{r}_t\}\}-\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t}>\max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\le \sum_{i\in[N]}v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v_{-i,t}^+,\widehat{r}_t\}<v_{i,t}<\max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\quad +\sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t}>\max\{b_{-i,t}^+\widehat{r}_t\}\}-\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t}>\max\{b_{-i,t}^+\widehat{r}_t\}\}\right]\,, \qquad (33)
\end{aligned}
$$

where the first inequality follows from Equation (22); the third inequality is due to the fact that $\sum_{i\in[N]}\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{b_{-i,t}^+\widehat{r}_t\}<v_{i,t}<\max\{v_{-i,t}^+,\widehat{r}_t\}\}\right] \ge 0$; and the last inequality holds because $\max\{v_t^-,\widehat{r}_t\}\le v_{\max}$. To continue the bound for Equation (33), we use the definition of $\mathcal{B}_{i,\ell}:=\mathcal{B}_{i,\ell}^s\cup\mathcal{B}_{i,\ell}^o$ in Lemma 3, where

$$
\begin{aligned}
\mathcal{B}_{i,\ell}^s &= \left\{t\in E_\ell:\mathbb{I}\left\{v_{i,t}>\{b_{-i,t}^+,\widehat{r}_t\}\right\}=1\,,\ \mathbb{I}\left\{b_{i,t}>\{b_{-i,t}^+,\widehat{r}_t\}\right\}=0\right\} \\
\mathcal{B}_{i,\ell}^o &= \left\{t\in E_\ell:\mathbb{I}\left\{v_{i,t}>\{b_{-i,t}^+,\widehat{r}_t\}\right\}=0\,,\ \mathbb{I}\left\{b_{i,t}>\{b_{-i,t}^+,\widehat{r}_t\}\right\}=1\right\}\,.
\end{aligned}
$$

Here, $\mathcal{B}_{i,\ell}^s$ represents the periods during which buyer i could have won the auction had she bid truthfully but in reality lost since she shaded her bid (allocation mismatch due to shading), while $\mathcal{B}_{i,\ell}^o$ represents the periods when buyer i would have lost the auction had she bid truthfully, but instead won the item due to overbidding (allocation mismatch due to overbidding). Hence, for any period $t\in E_\ell/\mathcal{B}_{i,\ell}=\left\{t\in E_\ell:\mathbb{I}\left\{v_{i,t}>\{b_{-i,t}^+,\widehat{r}_t\}\right\}=\mathbb{I}\left\{b_{i,t}>\{b_{-i,t}^+,\widehat{r}_t\}\right\}\right\}$ (which means in period $t\in E_\ell/\mathcal{B}_{i,\ell}$ the outcome for buyer i would not have changed even if she bid truthfully), we have

$\mathbb{I}\{v_{i,t} > \max\{b^+_{-i,t}, \widehat{r}_t\}\} = \mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}, \widehat{r}_t\}\}$. Therefore, defining $\mathcal{B}_\ell := \cup_{i\in[N]}\mathcal{B}_{i,\ell}$, we have

$$\begin{aligned}
&\mathcal{R}^{(2)}_t \mathbb{I}\{t \in E_\ell/\mathcal{B}_\ell\} \\
\leq\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}\widehat{r}_t\}\}\right] \\
&\quad + \sum_{i\in[N]} \mathbb{E}\left[\max\{v^-_t, \widehat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{b^+_{-i,t}\widehat{r}_t\}\} - \max\{b^-_t, \widehat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}\widehat{r}_t\}\}\right]\mathbb{I}\{t \in E_\ell/\mathcal{B}_\ell\} \\
=\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}\widehat{r}_t\}\}\right] \\
&\quad + \sum_{i\in[N]} \mathbb{E}\left[\left(\max\{v^-_t, \widehat{r}_t\} - \max\{b^-_t, \widehat{r}_t\}\right)\mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}, \widehat{r}_t\}\}\right] \\
\leq\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}, \widehat{r}_t\}\}\right] + \mathbb{E}\left[\max\{v^-_t, \widehat{r}_t\} - \max\{b^-_t, \widehat{r}_t\}\right] \\
\leq\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}, \widehat{r}_t\}\}\right] + \mathbb{E}\left[\left(v^-_t - b^-_t\right)^+\right].
\end{aligned}$$

The first inequality follows from Equation (33); the first equality follows from the fact that $t \in E_\ell/\mathcal{B}_\ell$; the second inequality holds because $\sum_{i\in[N]}\mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}\widehat{r}_t\}\} \leq \sum_{i\in[N]}\mathbb{I}\{b_{i,t} > b^+_{-i,t}\}\} = 1$; the third inequality applies the fact that $\max\{a,c\} - \max\{b,c\} \leq (a-b)^+$ for any $a,b,c \in \mathbb{R}$. Denoting $i^\star := \arg\max_{i\in[N]} v_{i,t}$, we have

$$\begin{aligned}
&\sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}, \widehat{r}_t\}\}\right] \\
=\ & v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t}, \widehat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t}, \widehat{r}_t\}\}\right]
\end{aligned}$$

since $\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t}\} = 0$ if $i \neq i^\star$. Therefore

$$\mathcal{R}^{(2)}_t \mathbb{I}\{t \in E_\ell/\mathcal{B}_\ell\} \ \leq\ v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t}, \widehat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t}, \widehat{r}_t\}\}\right] + \mathbb{E}\left[\left(v^-_t - b^-_t\right)^+\right], \tag{34}$$

To bound the first term in Equation (34), we again evoke the inequality $\max\{a,c\} - \max\{b,c\} = (a-b)^+$ for any $a,b,c \in \mathbb{R}$ and get $\max\{b^+_{-i^\star,t}, \widehat{r}_t\} - \max\{v^+_{-i^\star,t}, \widehat{r}_t\} \leq \left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+$. Hence,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t}, \widehat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t}, \widehat{r}_t\}\}\right] \\
\leq\ & \mathbb{E}\left[\mathbb{I}\{\max\{b^+_{-i^\star,t}, \widehat{r}_t\} - \left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+ < v_{i^\star,t} < \max\{b^+_{-i^\star,t}, \widehat{r}_t\}\}\right] \\
=\ & \mathbb{E}\left[\mathbb{E}\left[\mathbb{I}\{\max\{b^+_{-i^\star,t}, \widehat{r}_t\} - \left(b^+_{-i^\star,t} - v^+_{-i,t}\right)^+ < v_{i^\star,t} < \max\{b^+_{-i^\star,t}, \widehat{r}_t\}\} \,\middle|\, b^+_{-i^\star,t}, v^+_{-i^\star,t}\right]\right] \\
=\ & \mathbb{E}\left[\int_{\max\{b^+_{-i^\star,t},\widehat{r}_t\} - \left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+ - \langle\beta, x_t\rangle}^{\max\{b^+_{-i^\star,t},\widehat{r}_t\} - \langle\beta, x_t\rangle} f(z)dz\right] \\
\leq\ & c_f\mathbb{E}\left[\left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+\right].
\end{aligned} \tag{35}$$

Now, set $j \in [N]$ such that $b^+_{-i^\star,t} = b_{j,t}$ ($j \neq i^\star$), i.e. j is the highest bidder among all buyers excluding $i^\star$. Then $b^+_{-i^\star,t} - v^+_{-i^\star,t} = b_{j,t} - v^+_{-i^\star,t} \leq b_{j,t} - v_{j,t} = -a_{j,t}$, where the inequality follows from the fact that $v^+_{-i^\star,t}$ is the highest valuation among all buyers excluding $i^\star$ (which includes j as $j \neq i^\star$). Therefore, continuing the bound in Equation (35), we have

$$\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t}, \widehat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t}, \widehat{r}_t\}\}\right] \ \leq\ c_f(-a_{j,t})^+ \ \leq\ c_f\sum_{i\in[N]}(-a_{i,t})^+. \tag{36}$$

To bound the second term in Equation (34), namely $\mathbb{E}\left[\left(v^-_t - b^-_t\right)^+\right]$, without loss of generality assume $v_{1,t} \geq v_{2,t} \geq \cdots \geq v_{N,t}$. Hence $v^-_t = v_{2,t}$. If $b_{2,t} \leq b^-_t$, we have $v^-_t - b^-_t \leq v_{2,t} - b_{2,t} = a_{2,t}$. Otherwise if $b_{2,t} > b^-_t$, then buyer 2 submitted the highest bid, so $b_{i,t} \le b_t^-$ for any $i \neq 2$ and thus, $v_t^- - b_t^- \le v_{1,t} - b_t^- \le v_{1,t} - b_{1,t} = a_{1,t}$. Hence,

$$\mathbb{E}\left[\left(v_t^- - b_t^-\right)^+\right] \le \max_{j\in[N]}(a_{j,t})^+ \le \sum_{j\in[N]}(a_{j,t})^+ . \tag{37}$$

Finally, combining Equations (34), (36), and (37), we have for any $t \in E_\ell$ and $\ell \ge 1$

$$\mathcal{R}_t^{(2)}\mathbb{I}\{t \in E_\ell/\mathcal{B}_\ell\} \le v_{\max}c_f \sum_{i\in[N]}(-a_{i,t})^+ + \sum_{i\in[N]}(a_{i,t})^+ \le (v_{\max}c_f + 1)\sum_{i\in[N]}|a_{i,t}| \tag{38}$$

iv. Bounding Cumulative Regret: We now bound the cumulative expected regret in a phase $E_{\ell+1}$ ($\ell \ge 1$) via first bounding $\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(1)}$ and $\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(2)}$ respectively.

$$\begin{aligned}
&\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(1)}\\
\le\ &\sum_{t\in E_{\ell+1}}\left(12v_{\max}c_fN^2\delta_\ell + \frac{6v_{\max}\sqrt{N^3\log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f+L_\ell)+9N+15d+9\right)}{|E_\ell|}\right)\\
=\ &|E_{\ell+1}|\left(12v_{\max}c_fN^2\delta_\ell + \frac{6v_{\max}\sqrt{N^3\log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f+L_\ell)+9N+15d+9\right)}{|E_\ell|}\right)\\
=\ &|E_{\ell+1}|\cdot\frac{3v_{\max}\sqrt{2N^3\log(|E_\ell|)}}{\sqrt{|E_\ell|}}\left(\frac{4c_f\epsilon_{\max}x_{\max}^2\sqrt{d}}{\lambda_0^2}+1\right)\\
&+\frac{|E_{\ell+1}|}{|E_\ell|}\left(\frac{12v_{\max}c_fN^2\sqrt{d}\left(NL_\ell a_{\max}+1\right)x_{\max}^2}{\lambda_0^2} + v_{\max}\left(6N^2(c_f+L_\ell)+9N+15d+9\right)\right)\\
\le\ &c_1^1c_f\sqrt{dTN^3\log(|E_\ell|)} + c_2^2c_f\sqrt{d}N^3L_\ell T^{\frac{1}{4}}\\
\le\ &c_1c_f\sqrt{dN^3\log(|E_\ell|)}\left(\sqrt{T}+\frac{\sqrt{N^3\log(|E_\ell|)}T^{\frac{1}{4}}}{\log(1/\eta)}\right),
\end{aligned} \tag{39}$$

for some absolute constants $c_1^1, c_1^2, c_1 > 0$. The first inequality follows from Equation (31). In the second equality, we then used the definition of $\delta_\ell = \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max}x_{\max}^2}{\lambda_0^2\sqrt{N|E_\ell|}} + \frac{\sqrt{d}(NL_\ell a_{\max}+1)x_{\max}^2}{|E_\ell|\lambda_0^2}$, defined in Equation (18). In the second inequality, we relied on the construction of the length of phases in Algorithm 1, i.e. $|E_\ell| = T^{1-2^{-\ell}}$ so that $|E_{\ell+1}|/\sqrt{|E_\ell|} = \sqrt{T}$ and $|E_{\ell+1}|/|E_\ell| = T^{2^{-(\ell+1)}} \le T^{\frac{1}{4}}$. The last inequality follows from the fact that $L_\ell = \log\left(v_{\max}^2N|E_\ell|^4 - 1\right)/\log(\frac{1}{\eta})$.

On the other hand, to bound $\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(2)}$, we again utilize the definition of $\mathcal{B}_{i,\ell} := \mathcal{B}_{i,\ell}^s \cup \mathcal{B}_{i,\ell}^o$ and $\mathcal{B}_\ell := \cup_{i\in[N]}\mathcal{B}_{i,\ell}$ where $\mathcal{B}_{i,\ell}^s$ and $\mathcal{B}_{i,\ell}^o$ are defined in Equation (12) of Lemma 3. Denote $K_{\ell+1} := 2L_{\ell+1} + 4c_f + 8\log(|E_{\ell+1}|)$. Then, we have

$$\begin{aligned}\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(2)} &= \mathbb{E}\left[\sum_{t\in\mathcal{B}_{\ell+1}} \mathcal{R}_t^{(2)}\right] + \mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}} \mathcal{R}_t^{(2)}\right]\\
&\le v_{\max}\mathbb{E}\left[|\mathcal{B}_{\ell+1}|\cdot\mathbb{I}\{|\mathcal{B}_{\ell+1}|\le NK_{\ell+1}\right] + v_{\max}\mathbb{E}\left[|\mathcal{B}_{\ell+1}|\cdot\mathbb{I}\{|\mathcal{B}_{\ell+1}|> NK_{\ell+1}\}\right]\\
&\quad + (v_{\max}c_f+1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right]\\
&\le v_{\max}NK_{\ell+1} + v_{\max}|E_{\ell+1}|\cdot\mathbb{P}\left(|\mathcal{B}_{\ell+1}|>NK_{\ell+1}\right) + (v_{\max}c_f+1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right]\\
&\le v_{\max}NK_{\ell+1} + 4v_{\max}N + (v_{\max}c_f+1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right]\\
&\le v_{\max}N(K_{\ell+1}+4) + (v_{\max}c_f+1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right], \end{aligned} \tag{40}$$

where the first inequality follows from Equation (38) and uses the fact that $\mathcal{R}_t^{(2)} \le v_{\max}$; the second inequality is because $|\mathcal{B}_{\ell+1}| \le |E_{\ell+1}|$; the third inequality applies Lemma 3 which shows $\mathbb{P}\left(|\mathcal{B}_{i,\ell+1}| > K_{\ell+1}\right) \le 4/|E_{\ell+1}|$, and hence $\mathbb{P}\left(|\mathcal{B}_{\ell+1}| \le NK_{\ell+1}\right) \ge \mathbb{P}\left(\cap_{i\in[N]}\{|\mathcal{B}_{i,\ell+1}|\le K_{\ell+1}\}\right) \ge 1-4N/|E_{\ell+1}|$. To bound $\mathbb{E}\left[\sum_{t\in E_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right]$, we recall $\mathcal{S}_{\ell+1} := \cup_{i\in[N]}\mathcal{S}_{i,\ell+1}$ where $\mathcal{S}_{i,\ell+1}$ is defined in Equation (11), and consider the following

$$\begin{aligned}\mathbb{E}\left[\sum_{t\in E_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right] &\le \mathbb{E}\left[\sum_{t\in\mathcal{S}_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right] + \mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{S}_{\ell+1}}\sum_{i\in[N]}\frac{1}{|E_{\ell+1}|}\right]\\
&\le Na_{\max}\mathbb{E}\left[|\mathcal{S}_{\ell+1}|\right] + N\\
&= Na_{\max}\mathbb{E}\left[|\mathcal{S}_{\ell+1}|\cdot\left(\mathbb{I}\{|\mathcal{S}_{\ell+1}|\le NL_{\ell+1}\} + \mathbb{I}\{|\mathcal{S}_{\ell+1}|> NL_{\ell+1}\}\right)\right] + N\\
&\le Na_{\max}\left(NL_{\ell+1} + |E_{\ell+1}|\cdot\mathbb{P}\left(|\mathcal{S}_{\ell+1}|>NL_{\ell+1}\right)\right) + N\\
&\le N^2a_{\max}\left(L_{\ell+1}+1\right) + N\,, \end{aligned}\tag{41}$$

where the first inequality holds because $|a_{i,t}| \le 1/|E_{\ell+1}|$ for all $t \in E_{\ell+1}/\mathcal{S}_{\ell+1}$ and the fourth inequality follows from Lemma 1 that shows $\mathbb{P}\left(|\mathcal{S}_{i,\ell+1}| > L_{\ell+1}\right) \le 1/|E_{\ell+1}|$, which implies $\mathbb{P}\left(|\mathcal{S}_{\ell+1}| \le NL_{\ell+1}\right) \ge \mathbb{P}\left(\cap_{i\in[N]}\{|\mathcal{S}_{i,\ell+1}|\le L_{\ell+1}\}\right) \ge 1-N/|E_{\ell+1}|$.

Hence, Equations (40) and (41) show that $\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(2)}$ is upper bounded as

$$\begin{aligned}\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(2)} &\le v_{\max}N(K_\ell+4) + (v_{\max}c_f+1)\left(N^2a_{\max}\left(L_{\ell+1}+1\right)+N\right)\\
&\le c_2c_fN^2\cdot\frac{\log(|E_{\ell+1}|)}{\log(1/\eta)}\,,\end{aligned}\tag{42}$$

for some absolute constant $c_2 > 0$. Combining this with the upper bound

$$c_1c_f\sqrt{dN^3\log(|E_\ell|)}\left(\sqrt{T}+\frac{\sqrt{N^3\log(|E_\ell|)}T^{\frac{1}{4}}}{\log(1/\eta)}\right)$$

shown in Equation (39), the expected cumulative regret in phase $E_{\ell+1}$ is

$$\sum_{t\in E_{\ell+1}}\mathsf{Regret}_t \le c_3c_f\sqrt{dN^3\log(T)}\left(\sqrt{T}+\frac{\sqrt{N^3\log(T)}T^{\frac{1}{4}}}{\log(1/\eta)}\right),$$

for some absolute constant $c_3 > 0$. Finally, since the total number of phases is upper bounded by $\lceil \log\log(T) \rceil + 1$, the cumulative expected regret over the entire horizon T is

$$\begin{aligned}\mathsf{Regret}(T) \;&\leq\; v_{\max}|E_1| + \sum_{\ell=2}^{\lceil \log\log(T) \rceil} c_3 c_f \sqrt{dN^3 \log(T)} \left(\sqrt{T} + \frac{\sqrt{N^3 \log(T)} T^{\frac{1}{4}}}{\log(1/\eta)} \right) \\ &=\; \mathcal{O}\left(c_f \sqrt{dN^3 \log(T)} \cdot \log(\log(T)) \left(\sqrt{T} + \frac{\sqrt{N^3 \log(T)} T^{\frac{1}{4}}}{\log(1/\eta)} \right) \right).\end{aligned}$$

□

C.1 Proof of Lemma 1

According to the definitions of the cumulative discounted utility defined in Equation (1) and the NPAC-S policy in Algorithm 1, buyer i's utility for submitting a bid $b \in [0, v_{\max}]$ in period $t \in [T]$ conditioning on $v_{i,t}, b^+_{-i,t}, r_t$ is given by

$$u_{i,t}(b) = \begin{cases} \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right) \mathbb{I}\{b > \max\{r_t, b^+_{-i,t}\}\} & \text{no isolation} \\ (v_{i,t} - r_t)\, \mathbb{I}\{b > r_t\} & i \text{ is isolated} \\ 0 & j \neq i \text{ is isolated} \end{cases}, \tag{43}$$

where $b^+_{-i,t}$ is the highest bid excluding that of buyer i, and the reserve price $r_t = \widehat{r}_t \mathbb{I}\{\text{no isolation in } t\} + r^u_t (1 - \mathbb{I}\{\text{no isolation in } t\})$ ($\widehat{r}_t$ and r^u_t are defined in Equations (6) and (7) of the NPAC-S policy respectively). Note that $u_{i,t}(b)$ is a random variable that depends on the $x_t, \{\epsilon_{i,t}\}_{i\in[N]}, b^+_{-i,t}$ and r_t. The undiscounted utility loss $u^-_{i,t}$ for buyer i if he submits a bid $b_{i,t}$ compared to bidding truthfully is $u^-_{i,t} = u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t})$.

Now, when any buyer $j \neq i$ is isolated, the utility for buyer i is always 0 regardless of what he submits, so there is no utility loss due to bidding behaviour. We now consider the scenarios when no isolation occurs and when buyer i is isolated, respectively, using the definition of utility in Equation (1).

No isolation occurs: The undiscounted utility loss for bidding untruthfully is

$$\begin{aligned} u^-_{i,t} \mathbb{I}\{\text{no isolation in } t\} \;&=\; (u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t}))\, \mathbb{I}\{\text{no isolation in } t\} \\ &=\; \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right) \mathbb{I}\{v_{i,t} > \max\{r_t, b^+_{-i,t}\}\} \\ &\quad\; - \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right) \mathbb{I}\{b_{i,t} > \max\{r_t, b^+_{-i,t}\}\} \\ &=\; \left|v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right| \mathbb{I}\{v_{i,t} > \max\{r_t, b^+_{-i,t}\} > b_{i,t}\} \\ &\quad\; + \left|v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right| \mathbb{I}\{v_{i,t} < \max\{r_t, b^+_{-i,t}\} < b_{i,t}\} \\ &\geq\; 0\,. \end{aligned} \tag{44}$$

Isolating buyer i: The undiscounted utility for submitting any bid $b \in \mathbb{R}$ for any given r_t is $(v_{i,t} - r_t)\, \mathbb{I}\{b > r_t\}$. Hence,

$$\begin{aligned} u^-_{i,t} \mathbb{I}\{i \text{ is isolated}\} \;&=\; (u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t}))\, \mathbb{I}\{i \text{ is isolated}\} \\ &=\; (v_{i,t} - r_t)\, \mathbb{I}\{v_{i,t} > r_t\} - (v_{i,t} - r_t)\, \mathbb{I}\{b_{i,t} > r_t\} \\ &=\; (v_{i,t} - r_t)\, \mathbb{I}\{v_{i,t} > r_t > b_{i,t}\} + (-v_{i,t} + r_t)\, \mathbb{I}\{v_{i,t} < r_t < b_{i,t}\}\,. \end{aligned} \tag{45}$$

The NPAC-S policy offers a price r_t drawn from Uniform$(0, v_{\max})$ to the isolated buyer i with probability $1/|E_\ell|$, where i is chosen uniformly among all buyers. So, the expected utility loss $u^-_{i,t}$ for a buyer $i \in [N]$ conditioned on the fact that the buyer lies by an amount of $a_{i,t}$ is

$$
\begin{aligned}
&\mathbb{E}[u_{i,t}^- \mid a_{i,t}] \\
=\ &\mathbb{E}[u_{i,t}^- \mathbb{I}\{i \text{ is isolated}\} + u_{i,t}^- \mathbb{I}\{\text{no isolation in } t\} \mid a_{i,t}] \\
\geq\ &\mathbb{E}[u_{i,t}^- \mathbb{I}\{i \text{ is isolated}\} \mid a_{i,t}] \\
=\ &\frac{1}{N|E_\ell|}\mathbb{E}\left[(v_{i,t} - r_t)\,\mathbb{I}\{v_{i,t} > r_t > b_t\} + (-v_{i,t} + r_t)\,\mathbb{I}\{b_t < r_t < v_{i,t}\} \mid a_{i,t}\right] \\
=\ &\frac{1}{v_{\max}N|E_\ell|}\mathbb{E}\left[\mathbb{E}\left[\int_{v_{i,t}-a_{i,t}}^{v_{i,t}} (v_{i,t} - r)dr + \int_{v_{i,t}}^{v_{i,t}+a_{i,t}} (-v_{i,t} + r)dr \;\Big|\; a_{i,t}, v_{i,t}\right] \;\Big|\; a_{i,t}\right] \\
=\ &\frac{(a_{i,t})^2}{v_{\max}N|E_\ell|}\,.
\end{aligned} \tag{46}
$$

The first inequality follows from $u_{i,t}^- \mathbb{I}\{i \text{ is isolated}\} \geq 0$ as demonstrated in Equation (44). Now we lower bound the total expected utility loss in phase E_ℓ. First, by Equations (44) and (45), we know that $u_{i,t}^- \geq 0$ for $\forall i, t$. Therefore, denoting $s_{\ell+1}$ as the first period of phase $E_{\ell+1}$, for any $\tilde{z} > 0$ we have

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t\in E_\ell} \eta^t u_{i,t}^-\right] &\geq \mathbb{E}\left[\sum_{t\in \mathcal{S}_{i,\ell}} \eta^t u_{i,t}^-\right] \\
&\geq \mathbb{E}\left[\sum_{t\in \mathcal{S}_{i,\ell}} \eta^t u_{i,t}^- \mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{t\in \mathcal{S}_{i,\ell}} \eta^t u_{i,t}^- \;\Big|\; \{a_{i,t}\}_{t\in E_\ell}\right] \mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&\geq \mathbb{E}\left[\sum_{t\in \mathcal{S}_{i,\ell}} \frac{\eta^t}{v_{\max}N|E_\ell|^3} \cdot \mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&\geq \mathbb{E}\left[\sum_{t=s_{\ell+1}-|\mathcal{S}_{i,\ell}|}^{s_{\ell+1}-1} \frac{\eta^t}{v_{\max}N|E_\ell|^3} \cdot \mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&\geq \mathbb{E}\left[\sum_{t=s_{\ell+1}-\tilde{z}}^{s_{\ell+1}-1} \frac{\eta^t}{v_{\max}N|E_\ell|^3} \cdot \mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&= \frac{\eta^{s_{\ell+1}}\left(1 - \eta^{-\tilde{z}}\right)}{(1-\eta)v_{\max}N|E_\ell|^3}\mathbb{P}\left(|\mathcal{S}_{i,\ell}| \geq \tilde{z}\right)\,,
\end{aligned} \tag{47}
$$

where the first equality holds because $|\mathcal{S}_{i,\ell}| = \sum_{t\in E_\ell} \mathbb{I}\{a_{i,t} > 1/E_\ell\}$ is a function of $\{a_{i,t}\}_{t\in E_\ell}$; the third inequality follows from Equation (46) and $a_{i,t} \geq 1/|E_\ell|$ for any $t \in \mathcal{S}_{i,\ell}$; and the fourth inequality is because $\eta \in (0, 1)$.

Furthermore, corrupting a bid at time $t \in E_\ell$ will only impact the prices offered by the seller in future phases, i.e., phase $\ell + 1, \ell + 2, \dots$, so the utility gain due to lying in phase ℓ, denoted as $U_{i,\ell}^+$ is upper bounded by $v_{\max} \sum_{t\geq s_{\ell+1}} \eta^t = v_{\max}\eta^{s_{\ell+1}}/(1 - \eta)$. Since the buyer is utility maximizing, the net utility gain due to lying in phase ℓ should be greater than 0, otherwise the buyer can choose to always bid 0 in phase ℓ which is equivalent to not participating in the auctions. Hence,

$$
\mathbb{E}\left[U_{i,\ell}^+ - \sum_{t\in E_\ell} \eta^t u_{i,t}^-\right] \geq 0\,.
$$

Combining this with $U_{i,\ell}^+ \leq v_{\max}\eta^{s_{\ell+1}}/(1 - \eta)$ and the lower bound for $\mathbb{E}\left[\sum_{t\in E_\ell} u_{i,t}^-\right]$ shown in Equation (47), we have

$$
\frac{v_{\max}\eta^{s_{\ell+1}}}{1-\eta} \geq \frac{\eta^{s_{\ell+1}}\left(1 - \eta^{-\tilde{z}}\right)}{(1-\eta)v_{\max}N|E_\ell|^3}\mathbb{P}\left(|\mathcal{S}_{i,\ell}| \geq \tilde{z}\right)\,,
$$

which holds for any $\tilde{z} > 0$. Taking $\tilde{z} = \log\left(v_{\max}^2 N|E_\ell|^4 - 1\right) / \log(1/\eta)$ and by rearranging terms, the inequality above yields

$$\mathbb{P}\left(|\mathcal{S}_{i,\ell}| \geq \frac{\log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)}{\log(\frac{1}{\eta})}\right) \leq \frac{1}{|E_\ell|}\,.$$

□

C.2 Proof of Lemma 3

Defining $\mathcal{H}_{i,t} := \{(b_{-i,\tau}^+, \widehat{r}_\tau, x_\tau)\}_{\tau\in[t]}$, we have

$$\begin{aligned}
& \mathbb{E}\left[\mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s\} \mid \mathcal{H}_{i,t}\right] \\
= \; & \mathbb{P}\left(t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s \mid \mathcal{H}_{i,t}\right) \\
= \; & \mathbb{P}\left(v_{i,t} \geq \max\{b_{-i,t}^+, \widehat{r}_t\}\ ,\ b_{i,t} < \max\{b_{-i,t}^+, \widehat{r}_t\}\ ,\ a_{i,t} \in (0, 1/|E_\ell|) \mid \mathcal{H}_{i,t}\right) \\
= \; & \mathbb{P}\left(\max\{b_{-i,t}^+, \widehat{r}_t\} - \langle x_t, \beta\rangle \leq \epsilon_{i,t} \leq \max\{b_{-i,t}^+, \widehat{r}_t\} - \langle x_t, \beta\rangle + a_{i,t}\ ,\ a_{i,t} \in (0, 1/|E_\ell|) \mid \mathcal{H}_{i,t}\right) \\
\leq \; & \mathbb{P}\left(\max\{b_{-i,t}^+, \widehat{r}_t\} - \langle x_t, \beta\rangle \leq \epsilon_{i,t} \leq \max\{b_{-i,t}^+, \widehat{r}_t\} - \langle x_t, \beta\rangle + 1/|E_\ell| \mid \mathcal{H}_{i,t}\right) \\
= \; & \mathbb{E}\left[\int_{\max\{b_{-i,t}^+, \widehat{r}_t\} - \langle x_t, \beta\rangle}^{\max\{b_{-i,t}^+, \widehat{r}_t\} - \langle x_t, \beta\rangle + 1/|E_\ell|} f(z)dz \,\middle|\, \mathcal{H}_{i,t}\right] \\
\leq \; & \frac{c_f}{|E_\ell|}\,.
\end{aligned} \tag{48}$$

The last inequality uses the fact that $c_f = \sup_{\tilde{z}\in[-\epsilon_{\max}, \epsilon_{\max}]} f(\tilde{z})$.

Define $\zeta_t = \mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s\}$ and $\phi_t = \mathbb{E}\left[\mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s\} \mid \mathcal{H}_{i,t}\right]$. Then $\mathbb{E}[\zeta_t \mid \mathcal{H}_{i,t}] = \phi_t$, which implies $\mathbb{E}[\zeta_t - \phi_t \mid \sum_{\tau<t}\zeta_\tau, \sum_{\tau<t}\phi_\tau] = \mathbb{E}\left[\mathbb{E}\left[\zeta_t - \phi_t \mid \mathcal{H}_{i,t}\right] \mid \sum_{\tau<t}\zeta_\tau, \sum_{\tau<t}\phi_\tau\right] = 0$. Hence, in the context of the multiplicative Azuma inequality described in Lemma 11, by setting $z_{1,t} = \zeta_t$, $z_{2,t} = \phi_t$, $\tilde{\gamma} = 1/2$ and $A = 2\log(|E_\ell|)$ we have $|z_{1,t} - z_{2,t}| \leq 1$

$$\mathbb{P}\left(\frac{1}{2}\sum_{t\in E_\ell}\zeta_t \geq \sum_{t\in E_\ell}\phi_t + 2\log(|E_\ell|)\right) \leq \exp\left(-\log(|E_\ell|)\right)\,. \tag{49}$$

Now, according to Equation (48), we have $\phi_t \leq c_f/|E_\ell|$, so $\sum_{t\in E_\ell}\phi_t \leq c_f$. Moreover, $|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s| = \sum_{t\in E_\ell}\zeta_t$. Hence, following Equation (49), we have

$$\begin{aligned}
\mathbb{P}\left(|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s| \geq 2c_f + 4\log(|E_\ell|)\right) \; & \leq \; \mathbb{P}\left(\frac{1}{2}\sum_{t\in E_\ell}\zeta_t \geq \sum_{t\in E_\ell}\phi_t + 2\log(|E_\ell|)\right) \\
& \leq \; \exp\left(-\log(|E_\ell|)\right) = \frac{1}{|E_\ell|}\,.
\end{aligned} \tag{50}$$

When the event $\mathcal{G}_{i,t} = \{|\mathcal{S}_{i,\ell}| \leq L_\ell\}$ occurs, where $L_\ell = \log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)/\log(1/\eta)$, we have $|\mathcal{B}_{i,\ell}^s| \leq |\mathcal{S}_{i,\ell}| + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s| \leq L_\ell + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s|$. Therefore when event $\mathcal{G}_{i,t}$ occurs,

$$\begin{aligned}
& \mathbb{P}\left(|\mathcal{B}_{i,\ell}^s| \leq L_\ell + 2c_f + 4\log(|E_\ell|)\right) \\
\geq \; & \mathbb{P}\left(\left\{|\mathcal{B}_{i,\ell}^s| \leq L_\ell + 2c_f + 4\log(|E_\ell|)\right\} \bigcap \mathcal{G}_{i,t}\right) \\
\geq \; & \mathbb{P}\left(\left\{|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s| \leq 2c_f + 4\log(|E_\ell|)\right\} \bigcap \mathcal{G}_{i,t}\right) \\
\geq \; & 1 - \mathbb{P}\left(|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s| \geq 2c_f + 4\log(|E_\ell|)\right) - \mathbb{P}\left(\mathcal{G}_{i,t}^c\right) \\
\geq \; & 1 - \frac{2}{|E_\ell|}\,.
\end{aligned}$$

The second inequality follows from $|\mathcal{B}_{i,\ell}^s| \leq L_\ell + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}_{i,\ell}^s|$ when the event $\mathcal{G}_{i,t}$ occurs; the third inequality applies the union bound, and the final inequality follows from Equation (50) and Lemma 1.

Similarly, we can show the same probability upper bound for $|\mathcal{B}_{i,\ell}^o|$. Finally, using the fact that $\mathcal{B}_{i,\ell} = \mathcal{B}_{i,\ell}^s \cup \mathcal{B}_{i,\ell}^o$ and applying a union bound would yield the desired expression. □

C.3 Other Lemmas for proving Theorem 1

Lemma 4 (Bounding Estimation Errors in β)**.** *For any phase E_ℓ and $\gamma > 0$, we have*

$$\mathbb{P}\left(\|\widehat{\beta}_{\ell+1}-\beta\|_1 \le \gamma + \frac{d\left(NL_\ell a_{\max}+1\right)x_{\max}}{|E_\ell|\lambda_0^2}\right)$$
$$\ge\ 1-2d\exp\left(-\frac{N\gamma^2\lambda_0^4|E_\ell|}{2{\epsilon_{\max}}^2x_{\max}^2 d}\right)-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)-\frac{N}{|E_\ell|}\,,$$

where λ_0^2 is the minimum eigenvalue of the covariance matrix Σ, $\widehat{\beta}_{\ell+1}$ is defined in Equation (8), and $L_\ell = \log\left(v_{\max}^2 N|E_\ell|^4-1\right)/\log(1/\eta)$. Furthermore, setting $\gamma = \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max}x_{\max}}{\lambda_0^2\sqrt{N|E_\ell|}}$ and denoting $\delta_\ell = \gamma\cdot x_{\max} + \frac{d(NL_\ell a_{\max}+1)x_{\max}^2}{|E_\ell|\lambda_0^2}$, we have

$$\mathbb{P}\left(\|\widehat{\beta}_{\ell+1}-\beta\|_1 \le \frac{\delta_\ell}{x_{\max}}\right)\ge 1-\frac{2d+N}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\,.$$

Proof. Proof of Lemma 4.

The proof of Lemma 4 is inspired by Lemma EC.7.2 in Bastani and Bayati (2015), but here we made substantial modifications to resolve the issues that arise when estimating β in the presence of corrupted bids submitted by buyers.

First, recall that the smallest eigenvalue λ_0^2 of the covariance matrix Σ of $x\sim\mathcal{D}$ is greater than 0. Since the second moment matrix $\mathbb{E}[x_t x_t^\top]=\Sigma+\mathbb{E}[x]\mathbb{E}[x]^\top$, we know that the smallest eigenvalue of $\mathbb{E}[x_t x_t^\top]$ is at least $\lambda_0^2>0$. We denote the design matrix of all the features in phase E_ℓ as $X\in\mathbb{R}^{|E_\ell|\times d}$, and $\bar{\epsilon}_\tau=\frac{\sum_{i\in[N]}\epsilon_{i,\tau}}{N}$ for $\forall\tau\in E_\ell$.

We first consider the case where the smallest eigenvalue of the second moment matrix $\lambda_{\min}\left(X^\top X/|E_\ell|\right)\ge\lambda_0^2/2$, which implies that $(X^\top X)^{-1}$ exists and $(X^\top X)^{-1}=(X^\top X)^\dagger$. By the definition $b_{i,t}=v_{i,t}-a_{i,t}$, and the definition of $\bar{b}_\tau$ for any $\tau\in[T]$ in Equation (8) we have

$$\begin{aligned}\widehat{\beta}_{\ell+1}=\left(X^\top X\right)^{-1}X^\top\begin{pmatrix}\bar{b}_1\\ \vdots\\ \bar{b}_t\end{pmatrix}&=\left(X^\top X\right)^{-1}X^\top\begin{pmatrix}\frac{\sum_{i\in[N]}v_{i,1}-a_{i,1}}{N}\\ \vdots\\ \frac{\sum_{i\in[N]}v_{i,t}-a_{i,t}}{N}\end{pmatrix}\\&=\beta+\left(X^\top X\right)^{-1}X^\top\begin{pmatrix}\frac{\sum_{i\in[N]}\epsilon_{i,1}-a_{i,1}}{N}\\ \vdots\\ \frac{\sum_{i\in[N]}\epsilon_{i,t}-a_{i,t}}{N}\end{pmatrix}\\&=\beta+\left(X^\top X\right)^{-1}X^\top\left(\bar{\mathcal{E}}-A\right)\,,\end{aligned}\tag{51}$$

where $\bar{\mathcal{E}}$ is the column vector consisting of all $\bar{\epsilon}_\tau:=\frac{\sum_{i\in[N]}\epsilon_{i,\tau}}{N}$, and A is the column vector consisting of all $\bar{a}_\tau := \frac{\sum_{i\in[N]}a_{i,\tau}}{N}$ for $\forall\tau\in[t]$. Therefore,

$$\begin{aligned}\|\widehat{\beta}_{\ell+1}-\beta\|_2\ &=\ \|\left(X^\top X\right)^{-1}X^\top\left(\bar{\mathcal{E}}-A\right)\|_2\\&\le\ \frac{1}{|E_\ell|\lambda_0^2}\left(\|X^\top\bar{\mathcal{E}}\|_2+\|X^\top A\|_2\right)\,.\end{aligned}\tag{52}$$

Denote X^j as the jth column of X, i.e. the jth row of $X^\top$ for $j=1,2\ldots d$, we now bound $\|X^\top\bar{\mathcal{E}}\|_2$ and $\|X^\top A\|_2$ separately. First, since $\|X^\top\bar{\mathcal{E}}\|_2^2=\sum_{j\in[d]}\left|\bar{\mathcal{E}}^\top X^j\right|^2$, for any $\gamma>0$,

$$\bigcap_{j\in[d]}\left\{\left|\bar{\mathcal{E}}^\top X^j\right|\le\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right\}\subseteq\left\{\frac{1}{|E_\ell|\lambda_0^2}\cdot\|X^\top\bar{\mathcal{E}}\|_2\le\gamma\right\}\,.\tag{53}$$

We observe that $\bar{\mathcal{E}}^\top X^j=\frac{\sum_{\tau\in E_\ell}\sum_{i\in[N]}\epsilon_{i,\tau}X_{\tau j}}{N}$, where all $\epsilon_{i,\tau}X_{\tau j}$ are 0-mean and $\epsilon_{\max}x_{\max}$-subgaussion random variables. Therefore by Hoeffding's inequality, for any $\tilde{\gamma}>0$

$$\mathbb{P}\left(\left|N\bar{\mathcal{E}}^\top X^j\right|\le\tilde{\gamma}\right)\ge 1-2\exp\left(-\frac{\tilde{\gamma}^2}{2{\epsilon_{\max}}^2x_{\max}^2|E_\ell|N}\right)\,.\tag{54}$$

Replacing $\tilde{\gamma}$ with $N|E_\ell|\lambda_0^2\gamma/\sqrt{d}$ and using Equation (53) yields:

$$\begin{aligned}\mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\cdot\|X^\top\bar{\mathcal{E}}\|_2\le\gamma\right\}\right) \;&\ge\; \mathbb{P}\left(\bigcap_{j\in[d]}\left\{\left|\bar{\mathcal{E}}^\top X^j\right|\le\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right\}\right)\\ &\ge\; 1-\sum_{j\in[d]}\mathbb{P}\left(\left|\bar{\mathcal{E}}^\top X^j\right|>\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right)\\ &\ge\; 1-2d\exp\left(-\frac{N\gamma^2\lambda_0^4|E_\ell|}{2{\epsilon_{\max}}^2x_{\max}^2d}\right)\,,\end{aligned}\tag{55}$$

where the first inequality follows from Equation (53), the second inequality applies the union bound, and the last inequality follows from Equation (54).

In the following, we show a high probability bound for $\|X^\top A\|_2^2$ by using the fact that $|a_{i,t}|\le 1/|E_\ell|$ for any $t\in E_\ell/\mathcal{S}_{i,\ell}$, where $\mathcal{S}_{i,\ell}=\{t\in E_\ell:|a_{i,t}|>1/|E_\ell|\}$, and $\mathcal{S}_{i,\ell}\le L_\ell$ with high probability.

Recall the event $\mathcal{G}_{i,\ell}=\{|\mathcal{S}_{i,\ell}|\le L_\ell\}$, and in Lemma 1 we showed that $\mathbb{P}\left(\mathcal{G}_{i,\ell}^c\right)=\mathbb{P}\left(|\mathcal{S}_{i,\ell}|>L_\ell\right)\le\frac{1}{|E_\ell|}$. We now bound $\|X^\top A\|_2$ under the occurrence of $\mathcal{G}_{i,\ell}$ for all i.

$$\begin{aligned}\|X^\top A\|_2^2=\sum_{j\in[d]}\left|A^\top X^j\right|^2 \;&=\; \sum_{j\in[d]}\left(\frac{\sum_{\tau\in E_\ell}\sum_{i\in[N]}a_{i,\tau}X_{\tau j}}{N}\right)^2\\ &\le\; \sum_{j\in[d]}\left(\frac{\sum_{\tau\in E_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\right)^2\,.\end{aligned}\tag{56}$$

For periods in $S_\ell:=\cup_{i\in[N]}\mathcal{S}_{i,\ell}$, we have,

$$\frac{\sum_{\tau\in S_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\;\le\;\sum_{\tau\in S_\ell}a_{\max}x_{\max}\;\le\;NL_\ell a_{\max}x_{\max}\,,\tag{57}$$

where the last inequality holds because events $\mathcal{G}_{i,\ell}$ occurs for all i. On the other hand, recall that $|a_{i,t}|\ge 1/|E_\ell|$ for any i and $t\in\mathcal{S}_{i,\ell}$. Hence, $|a_{i,t}|\le 1/|E_\ell|$ for periods in $E_\ell/\mathcal{S}_\ell$,

$$\frac{\sum_{\tau\in E_\ell/\mathcal{S}_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\;\le\;\sum_{\tau\in E_\ell/\mathcal{S}_\ell}\frac{x_{\max}}{|E_\ell|}\;\le\;\sum_{\tau\in E_\ell}\frac{x_{\max}}{|E_\ell|}=x_{\max}\,.\tag{58}$$

Combining Equations (56), (57), and (58), we have

$$\|X^\top A\|_2\;\le\;\sqrt{d\left(\frac{\sum_{\tau\in[t]}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\right)^2}\;\le\;\sqrt{d}\,(NL_\ell a_{\max}+1)\,x_{\max}\,.\tag{59}$$

Now it only remains to show $\lambda_{\min}\left(X^\top X/|E_\ell|\right)\;\ge\;\lambda_0^2/2$ with high probability, which can be achieved by applying Lemma 10. In the context of this lemma, we consider the sequence of random matrices $\{x_\tau x_\tau^\top/|E_\ell|\}_{\tau\in[E_\ell]}$, and note that $X^\top X/|E_\ell|=\sum_{\tau\in E_\ell}(x_\tau x_\tau^\top/|E_\ell|)$. We first upper bound the maximum eigenvalue of $x_\tau x_\tau^\top/|E_\ell|$, namely $\lambda_{\max}\left(x_\tau x_\tau^\top/|E_\ell|\right)$ for any $\tau\in E_\ell$ by

$$\lambda_{\max}\left(\frac{x_\tau x_\tau^\top}{|E_\ell|}\right)=\max_{\|z\|_2=1}z^\top\frac{x_\tau x_\tau^\top}{|E_\ell|}z\le\frac{1}{|E_\ell|}\max_{\|z\|_2=1}(x^\top z)^2\le\frac{x_{\max}^2}{|E_\ell|}\,.$$

This allows us to apply the matrix Chernoff bound in Lemma 10 (setting $\bar{\gamma}=1/2$ in the lemma) and get

$$\begin{aligned}\mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right)\ge\frac{\lambda_0^2}{2}\right)\;&\ge\;\mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right)\ge\;\frac{1}{2}\lambda_{\min}\left(\mathbb{E}\left[\frac{X^\top X}{|E_\ell|}\right]\right)\right)\\ &\ge\;1-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\,,\end{aligned}\tag{60}$$

where the first inequality follows from the fact that $\lambda_{\min}\left(\mathbb{E}[X^\top X/|E_\ell|]\right) \geq \lambda_0^2$.

Putting everything together, we get

$$\begin{aligned}
&\mathbb{P}\left(\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \gamma + \frac{\sqrt{d}\left(NL_\ell a_{\max} + 1\right)x_{\max}}{|E_\ell|\lambda_0^2}\right)\\
\geq\ &\mathbb{P}\left(\|\widehat{\beta}_{\ell+1} - \beta\|_2 \leq \gamma + \frac{\sqrt{d}\left(NL_\ell a_{\max} + 1\right)x_{\max}}{|E_\ell|\lambda_0^2}\right)\\
\geq\ &\mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\left(\|X^\top \bar{\mathcal{E}}\|_2 + \|X^\top A\|_2\right) \leq \gamma + \frac{\sqrt{d}\left(NL_\ell a_{\max} + 1\right)x_{\max}}{|E_\ell|\lambda_0^2}\right\}\bigcap\left\{\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \geq \frac{\lambda_0^2}{2}\right\}\right)\\
\geq\ &\mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\|X^\top \bar{\mathcal{E}}\|_2 \leq \gamma\right\} \bigcap \left(\bigcap_{i\in[N]}\mathcal{G}_{i,\ell}\right) \bigcap \left\{\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \geq \frac{\lambda_0^2}{2}\right\}\right)\\
\geq\ &1 - \mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\|X^\top \bar{\mathcal{E}}\|_2 > \gamma\right\}\right) - \sum_{i\in[N]}\mathbb{P}\left(\mathcal{G}_{i,\ell}^c\right) - \mathbb{P}\left(\left\{\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \leq \frac{\lambda_0^2}{2}\right\}\right)\\
\geq\ &1 - 2d\exp\left(-\frac{N\gamma^2\lambda_0^4|E_\ell|}{2{\epsilon_{\max}}^2 x_{\max}^2 d}\right) - \frac{N}{|E_\ell|} - d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right).
\end{aligned}$$

The first inequality follows from the fact that $\|z\|_1 \leq \|z\|_2$ for any vector z; the second inequality follows from Equation (52); the third inequality follows from Equation (59) when the event $\cap_{i\in[N]}\mathcal{G}_{i,\ell}$ occurs; the fourth inequality applies a simple union bound; and the final inequality follows from Equations (55), (60) and Lemma 1. □

Lemma 5 (Bounding Estimation Error in F^- and F^+)**.** *Define $\tilde{\sigma}_t$ to be the sigma algebra generated by all $\{x_\tau, a_{i,\tau}, \epsilon_{i,\tau}\}_{i\in[N],\tau\in[t]}$. Then, for any $\tilde{\sigma}_t$-measurable random variable z and $\gamma > 0$, we have*

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{F}^-_{\ell+1}(z) - F^-(z)\right| \leq 2N^2 z_\ell\right) \ &\geq\ 1 - 4\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{4(d+N)}{|E_\ell|} - 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\\
\mathbb{P}\left(\left|\widehat{F}^+_{\ell+1}(z) - F^+(z)\right| \leq N z_\ell\right) \ &\geq\ 1 - 4\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{4(d+N)}{|E_\ell|} - 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right),
\end{aligned}$$

where $z_\ell := \gamma + c_f\delta_\ell + (c_f + L_\ell)/|E_\ell|$, $c_f = \sup_{\tilde{z}\in[-\epsilon_{\max},\epsilon_{\max}]} f(\tilde{z})$, δ_ℓ is defined in Equation (18), and $L_\ell = \log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)/\log(1/\eta)$.

Proof. Proof of Lemma 5. We first bound the error in the estimate of F, namely $\left|\widehat{F}_{\ell+1}(z) - F(z)\right|$. Then, we use the relationship $F^-(z) = NF^{N-1}(z) - (N-1)F^N(z)$ and $F^+(z) = F^N(z)$, as well as the definition of $\widehat{F}^-_{\ell+1}(z)$ and $\widehat{F}^+_{\ell+1}(z)$ in Equation (9) to show the desired probability bounds.

We first upper and lower bound $\widehat{F}^-_{\ell+1}(z)$ for any $z \in \mathbb{R}$. Recall the event $\mathcal{S}_{i,\ell} = \{t \in E_\ell : |a_{i,t}| \geq 1/|E_\ell|\}$ and in Lemma 1 we showed that $\mathbb{P}\left(|\mathcal{S}_{i,\ell}| > L_\ell\right) \leq 1/|E_\ell|$. Hence, for any $i \in [N]$, we have $|a_{i,t}| \leq 1/|E_\ell|$ for all periods $\tau \in E_\ell/\mathcal{S}_{i,\ell}$, so

$$\begin{aligned}
&\sum_{\tau\in E_\ell}\mathbb{I}\left\{b_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\}\\
=\ &\left(\sum_{\tau\in E_\ell/\mathcal{S}_{i,\ell}}\mathbb{I}\left\{b_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\} + \sum_{\tau\in\mathcal{S}_{i,\ell}}\mathbb{I}\left\{v_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\}\right)\\
&+\left(\sum_{\tau\in\mathcal{S}_{i,\ell}}\mathbb{I}\left\{b_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\} - \sum_{\tau\in\mathcal{S}_{i,\ell}}\mathbb{I}\left\{v_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\}\right). \qquad (61)
\end{aligned}$$

Consider the sum in first the parenthesis of Equation (61) and note that $b_{i,\tau} = v_{i,\tau} - a_{i,\tau} = \langle\beta, x_\tau\rangle + \epsilon_{i,\tau} - a_{i,\tau}$. Since $|a_{i,\tau}| \leq 1/|E_\ell|$ for any $i \in [N]$ and $\tau \in E_\ell/\mathcal{S}_{i,\ell}$,

$$\langle\beta, x_\tau\rangle + \epsilon_{i,\tau} - \frac{1}{|E_\ell|} \ \leq\ b_{i,\tau} \ \leq\ \langle\beta, x_\tau\rangle + \epsilon_{i,\tau} + \frac{1}{|E_\ell|}, \quad \forall\tau \in E_\ell/\mathcal{S}_{i,\ell}\,. \qquad (62)$$

Now, assume that the event $\xi_{\ell+1} = \left\{ \|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \delta_\ell / x_{\max} \right\}$ holds. Therefore, we can upper bound the sum in first the parenthesis of Equation (61) as

$$
\begin{aligned}
& \sum_{\tau \in E_\ell / \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z \right\} + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ v_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z \right\} \\
\leq & \sum_{\tau \in E_\ell / \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ \epsilon_{i,\tau} \leq z + \langle \widehat{\beta}_{\ell+1} - \beta, x_\tau \rangle + \frac{1}{|E_\ell|} \right\} + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ \epsilon_{i,\tau} \leq z + \langle \widehat{\beta}_{\ell+1} - \beta, x_\tau \rangle + \frac{1}{|E_\ell|} \right\} \\
= & \sum_{\tau \in E_\ell} \mathbb{I}\left\{ \epsilon_{i,\tau} \leq z + \langle \widehat{\beta}_{\ell+1} - \beta, x_\tau \rangle + \frac{1}{|E_\ell|} \right\} \\
\leq & \sum_{\tau \in E_\ell} \mathbb{I}\left\{ \epsilon_{i,\tau} \leq z + \delta_\ell + \frac{1}{|E_\ell|} \right\} ,
\end{aligned} \tag{63}
$$

where the first equality follows from $v_{i,\tau} = \langle \beta, x_\tau \rangle + \epsilon_{i,\tau}$ and $b_{i,\tau} = v_{i,\tau} - a_{i,\tau}$; the first inequality follows Equation (62); and the final inequality is due to the occurrence of the event $\xi_{\ell+1} = \left\{ \|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \delta_\ell / x_{\max} \right\}$. Similarly, we can also lower bound the sum in the first parenthesis of Equation (61):

$$
\sum_{\tau \in E_\ell / \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z \right\} + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z \right\} \;\geq\; \sum_{\tau \in E_\ell} \mathbb{I}\left\{ \epsilon_{i,\tau} \leq z - \delta_\ell - \frac{1}{|E_\ell|} \right\} . \tag{64}
$$

Furthermore, assuming events $\mathcal{G}_{i,\ell} = \{|\mathcal{S}_{i,\ell}| \leq L_\ell\}$ hold for all $i \in [N]$, we can simply upper bound and lower bound the expression in the second parenthesis of Equation (61):

$$
-L_\ell \;\leq\; \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z \right\} - \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{ v_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z \right\} \;\leq\; L_\ell . \tag{65}
$$

Combining Equations (61), (63), (64), (65), and using the definition

$$
\widehat{F}_{\ell+1}(z) = \frac{1}{N|E_\ell|} \sum_{i \in [N]} \sum_{\tau \in E_\ell} \mathbb{I}\left\{ b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z \right\},
$$

under the occurrence of events $\xi_{\ell+1}$, and $\mathcal{G}_{i,\ell}$ for all $i \in [N]$, we have

$$
\begin{aligned}
& \frac{1}{N|E_\ell|} \sum_{i \in [N]} \sum_{\tau \in E_\ell} \mathbb{I}\left\{ \epsilon_{i,\tau} \leq z - \delta_\ell - \frac{1}{|E_\ell|} \right\} - \frac{L_\ell}{|E_\ell|} \;\leq\; \widehat{F}_{\ell+1}(z) \;\text{ and} \\
& \widehat{F}_{\ell+1}(z) \;\leq\; \frac{1}{N|E_\ell|} \sum_{i \in [N]} \sum_{\tau \in E_\ell} \mathbb{I}\left\{ \epsilon_{i,\tau} \;\leq\; z + \delta_\ell + \frac{1}{|E_\ell|} \right\} + \frac{L_\ell}{|E_\ell|} .
\end{aligned} \tag{66}
$$

Now, for any $\gamma > 0$,

$$
\begin{aligned}
&\mathbb{P}\left(F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right)-\widehat{F}_{\ell+1}(z)\leq \gamma+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &\mathbb{P}\left(\left\{F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right)-\widehat{F}_{\ell+1}(z)\leq \gamma+\frac{L_\ell}{|E_\ell|}\right\}\bigcap \xi_{\ell+1}\bigcap\left(\bigcap_{i\in[N]}\mathcal{G}_{i,\ell}\right)\right)\\
\geq\ &\mathbb{P}\left(\left\{F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right)-\frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\leq z-\delta_\ell-\frac{1}{|E_\ell|}\right\}\leq\gamma\right\}\bigcap \xi_{\ell+1}\bigcap\left(\bigcap_{i\in[N]}\mathcal{G}_{i,\ell}\right)\right)\\
\geq\ &\mathbb{P}\left(\left\{\sup_{\tilde z\in\mathbb{R}}\left|F(\tilde z)-\frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\leq \tilde z\right\}\right|\leq\gamma\right\}\bigcap \xi_{\ell+1}\bigcap\left(\bigcap_{i\in[N]}\mathcal{G}_{i,\ell}\right)\right)\\
\geq\ &1-\mathbb{P}\left(\left\{\sup_{\tilde z\in\mathbb{R}}\left|F(\tilde z)-\frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\leq \tilde z\right\}\right|>\gamma\right\}\right)-\mathbb{P}\left(\xi_{\ell+1}^c\right)-\sum_{i\in[N]}\mathbb{P}\left(\mathcal{G}_{i,\ell}^c\right)\\
\geq\ &1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\left(\frac{2d+N}{|E_\ell|}+d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\right)-\frac{N}{|E_\ell|}\\
=\ &1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{2(d+N)}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right),
\end{aligned}
\tag{67}
$$

where the second inequality follows from Equation (66), the fourth inequality uses the union bound, and the final inequality follows from the DKW inequality (Theorem 9), Lemma 4, and Lemma 1. We note that we can apply the DKW inequality because $\{\epsilon_{i,\tau}\}_{\tau\in E_\ell, i\in[N]}$ are $N|E_\ell|$ i.i.d. realizations of noise variables. According to the Lipschitz property of F shown in Lemma 8, $|F(z-\delta_\ell-1/|E_\ell|)-F(z)|\leq c_f(\delta_\ell+1/|E_\ell|)$ for $\forall z\in\mathbb{R}$. Hence, combining this with Equation (67), yields

$$
\begin{aligned}
&\mathbb{P}\left(F(z)-\widehat{F}_{\ell+1}(z)\ \leq\ \gamma+c_f\left(\delta_\ell+\frac{1}{|E_\ell|}\right)+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &\mathbb{P}\left(F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right)-\widehat{F}_{\ell+1}(z)\ \leq\ \gamma+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{2(d+N)}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right).
\end{aligned}
\tag{68}
$$

Similarly, $|F(z+\delta_\ell+1/|E_\ell|)-F(z)|\leq c_f(\delta_\ell+1/|E_\ell|)$ for $\forall z\in\mathbb{R}$, so we can show

$$
\begin{aligned}
&\mathbb{P}\left(\widehat{F}_{\ell+1}(z)-F(z)\ \leq\ \gamma+c_f\left(\delta_\ell+\frac{1}{|E_\ell|}\right)+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &\mathbb{P}\left(\widehat{F}_{\ell+1}(z)-F\left(z+\delta_\ell+\frac{1}{|E_\ell|}\right)\ \leq\ \gamma+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{2(d+N)}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right).
\end{aligned}
\tag{69}
$$

Combining Equations (68) and (69) using a union bound yields

$$
\begin{aligned}
&\mathbb{P}\left(\left|\widehat{F}_{\ell+1}(z)-F(z)\right|\leq\gamma+c_f\delta_\ell+\frac{c_f+L_\ell}{|E_\ell|}\right)\\
\geq\ &1-4\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{4(d+N)}{|E_\ell|}-2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right).
\end{aligned}
\tag{70}
$$

Finally, we now bound $|\widehat{F}_t^-(z)-F^-(z)|$ and $|\widehat{F}_t^+(z)-F^+(z)|$ using the fact that $F^-(z)=NF^{N-1}(z)-(N-1)F^N(z)$

and $F^+(z) = F^N(z)$.

$$
\begin{aligned}
|\widehat{F}^-_{\ell+1}(z) - F^-(z)| &= \left| N\widehat{F}^{N-1}_{\ell+1}(z) - (N-1)\widehat{F}^N_{\ell+1}(z) - \left(NF^{N-1}(z) - (N-1)F^N(z)\right)\right| \\
&\le N\left|\widehat{F}^{N-1}_{\ell+1}(z) - F^{N-1}(z)\right| + (N-1)\left|\widehat{F}^N_{\ell+1}(z) - F^N(z)\right| \\
&= N\left|\left(\widehat{F}_{\ell+1}(z) - F(z)\right)\left(\sum_{n=1}^{N-1}\left(\widehat{F}_{\ell+1}(z)\right)^{n-1}(F(z))^{N-1-n}\right)\right| \\
&\quad + (N-1)\left|\left(\widehat{F}_{\ell+1}(z) - F(z)\right)\left(\sum_{n=1}^{N}\left(\widehat{F}_{\ell+1}(z)\right)^{n-1}(F(z))^{N-n}\right)\right| \\
&\le N(N-1)\left|\widehat{F}_{\ell+1}(z) - F(z)\right| + (N-1)N\left|\widehat{F}_{\ell+1}(z) - F(z)\right| \\
&< 2N^2\left|\widehat{F}_{\ell+1}(z) - F(z)\right| . \qquad (71)
\end{aligned}
$$

The second equality uses $a^m - b^m = (a-b)\left(\sum_{n=1}^m a^{n-1}b^{m-n}\right)$ for any integer $m \ge 2$. The second inequality follows from $\widehat{F}_{\ell+1}(z), F(z) \in [0,1]$ for $\forall z \in \mathbb{R}$. Combining Equations (70) and (71), we get

$$
\begin{aligned}
&\mathbb{P}\left(\left|\widehat{F}^-_{\ell+1}(z) - F^-(z)\right| \le 2N^2\left(\gamma + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right)\right) \\
&\ge 1 - 4\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{4(d+N)}{|E_\ell|} - 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x^2_{\max}}\right) .
\end{aligned}
$$

The probability bound for $\left|\widehat{F}^-_{\ell+1}(z) - F^-(z)\right|$ can be shown in a similar fashion by noting that similar to Equation (71) we can show $|\widehat{F}^+_{\ell+1}(z) - F^+(z)| < N\left|\widehat{F}_{\ell+1}(z) - F(z)\right|$. □

Lemma 6 (Bounding the Impact of Estimation Errors on Revenue)**.** *We assume that the events* $\xi_{\ell+1} = \left\{\|\widehat{\beta}_{\ell+1} - \beta\|_1 \le \frac{\delta_\ell}{x_{\max}}\right\}$, $\xi^-_{\ell+1} = \left\{\left|\widehat{F}^-_{\ell+1}(z) - F^-(z)\right| \le 2N^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f+L_\ell}{|E_\ell|}\right)\right\}$ *and* $\xi^+_{\ell+1} = \left\{\left|\widehat{F}^+_{\ell+1}(z) - F^+(z)\right| \le N\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f+L_\ell}{|E_\ell|}\right)\right\}$ *occur for some phase* $\ell \ge 1$*, where* $z \in \mathbb{R}$*,* $\gamma_\ell = \sqrt{\log(|E_\ell|)}/\sqrt{2N|E_\ell|}$*, and* δ_ℓ *is defined in Equation (18). Hence for any* $r \in \{r^\star_t,\ r_t\}$ *where* $t \in E_{\ell+1}$ *we have the following:*

(i) $|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, F^-, F^+)| \le 3rc_fN^2\delta_\ell$ *a.s.*

(ii) $\left|\rho_t(r, \widehat{y}_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, \widehat{F}^-_{\ell+1}, \widehat{F}^+_{\ell+1})\right| \le 3rN^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f+L_\ell}{|E_\ell|}\right)$ *a.s.*

where $y_t = \langle\beta, x_t\rangle$*,* $\widehat{y}_t = \langle\widehat{\beta}_{\ell+1}, x_t\rangle$*,* $\widehat{\beta}_{\ell+1}, \widehat{F}^-_{\ell+1}, \widehat{F}^+_{\ell+1}$ *are defined in Equations (8) and (9). The function* ρ_t *is defined in Equation (27).*

Proof. Proof of Lemma 6. **Part (i)** We consider the following:

$$
\begin{aligned}
&\left|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, F^-, F^+)\right| \\
=\ &\left|\int_0^r \left[F^-(z - y_t) - F^-(z - \widehat{y}_t)\right] dz - r\left[F^+(r - y_t) - F^+(r - \widehat{y}_t)\right]\right| \\
\le\ &\int_0^r \left|F^-(z - y_t) - F^-(z - \widehat{y}_t)\right| dz + r\left|F^+(r - y_t) - F^+(r - \widehat{y}_t)\right| \\
\le\ &\int_0^r 2c_fN^2|y_t - \widehat{y}_t| dz + rc_fN|y_t - \widehat{y}_t| \\
\le\ &\int_0^r 2c_fN^2\left(\|\widehat{\beta}_{\ell+1} - \beta\|_1 x_{\max}\right) dz + rc_fN\|\widehat{\beta}_{\ell+1} - \beta\|_1 x_{\max} \\
\le\ &3rc_fN^2\delta_\ell .
\end{aligned}
$$

The first equality follows from definition of ρ_t in Equation (27), and the second inequality applies the Lipschitz property of F^- and F^+ using Lemma 8. The third inequality follows from Cauchy's inequality: $|y_t - \widehat{y}_t| = |\langle \widehat{\beta}_{\ell+1} - \beta, x_t\rangle| \leq \|\widehat{\beta}_{\ell+1} - \beta\|_1 x_{\max}$, and the last inequality follows from the occurrence of $\xi_{\ell+1}$ and $N \geq 1$.

Part (ii) Similar to part (i), we have

$$\begin{aligned}
&\left|\rho_t(r, \widehat{y}_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, \widehat{F}^-_{\ell+1}, \widehat{F}^+_{\ell+1})\right| \\
= &\left|\int_0^r \left[F^-(z - \widehat{y}_t) - \widehat{F}^-_{\ell+1}(z - \widehat{y}_t)\right] dz - r\left[F^+(r - \widehat{y}_t) - \widehat{F}^+_{\ell+1}(r - \widehat{y}_t)\right]\right| \\
\leq &\int_0^r \left|F^-(z - \widehat{y}_t) - \widehat{F}^-_{\ell+1}(z - \widehat{y}_t)\right| dz + r\left|F^+(r - \widehat{y}_t) - \widehat{F}^+_{\ell+1}(r - \widehat{y}_t)\right| \\
\leq &\ 3rN^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right),
\end{aligned}$$

where the last inequality follows from the occurrence of events $\xi^-_{\ell+1}$ and $\xi^+_{\ell+1}$ and $N \geq 1$. □

Lemma 7 (Bounding probabilities)**.** *The probability that not all events $\xi_{\ell+1}$, $\xi^-_{\ell+1}$ and $\xi^+_{\ell+1}$ occur for some phase $\ell \geq 1$ is bounded as*

$$\mathbb{P}\left(\xi^c_{\ell+1} \cup \left(\xi^-_{\ell+1}\right)^c \cup \left(\xi^+_{\ell+1}\right)^c\right) \ \leq\ \frac{9N + 15d + 8}{|E_\ell|},$$

where the events $\xi_{\ell+1}$, $\xi^-_{\ell+1}$ and $\xi^+_{\ell+1}$ are defined in Equations (17), (19), and (20) respectively.

Proof. Proof of Lemma 7.

We first bound the probability of $\xi^c_{\ell+1}$, and then proceed to bound the the probability of $\left(\xi^-_{\ell+1}\right)^c$ and $\left(\xi^+_{\ell+1}\right)^c$.

Recall that $\xi_{\ell+1} = \left\{\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}}\right\}$. Then,

$$\begin{aligned}
\mathbb{P}\left(\xi^c_{\ell+1}\right) &\leq \frac{2d + N}{|E_\ell|} + d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x^2_{\max}}\right) \\
&\leq \frac{2d + N}{|E_\ell|} + d\exp\left(-\frac{\log(|E_\ell|)T^{\frac{1}{4}}\lambda_0^2}{8x^2_{\max}}\right) \\
&\leq \frac{N + 3d}{|E_\ell|},
\end{aligned} \tag{72}$$

where the first inequality follows from Lemma 4 by taking $\gamma = \sqrt{2d\log(|E_\ell|)}\epsilon_{\max}x_{\max}/\left(\lambda_0^2\sqrt{N|E_\ell|}\right)$; the second inequality uses the fact that $|E_\ell| \geq |E_1| = \sqrt{T},\ \ T \geq \max\left\{\left(\frac{8x^2_{\max}}{\lambda_0^2}\right)^4, 9\right\}$, which implies $|E_\ell| \geq \log(|E_\ell|)\sqrt{|E_\ell|} \geq T^{\frac{1}{4}}\log(|E_\ell|)$. Note that here we used the fact that $\sqrt{x} \geq \log(x)$ for all $x \geq 9$.

We now bound the probability of $\left(\xi^-_{\ell+1}\right)^c$:

$$\begin{aligned}
\mathbb{P}\left(\left(\xi^-_{\ell+1}\right)^c\right) &\leq 4\exp\left(-2N|E_\ell|\cdot\left(\frac{\sqrt{\log(|E_\ell|)}}{\sqrt{2N|E_\ell|}}\right)^2\right) + \frac{4(d + N)}{|E_\ell|} + 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x^2_{\max}}\right) \\
&\leq \frac{2(2N + 3d + 2)}{|E_\ell|},
\end{aligned} \tag{73}$$

where the first inequality follows from Lemma 5 by taking $\gamma = \gamma_\ell = \sqrt{\log(|E_\ell|)}/\sqrt{2N|E_\ell|}$, and the last inequality again uses the fact that $|E_\ell| \geq \log(|E_\ell|)\sqrt{|E_\ell|} \geq T^{\frac{1}{4}}\log(|E_\ell|)$ when $T \geq \max\left\{\left(\frac{8x^2_{\max}}{\lambda_0^2}\right)^4, 9\right\}$.

Similarly, we can bound the probability of $\left(\xi^+_{\ell+1}\right)^c$:

$$\mathbb{P}\left(\left(\xi^+_{\ell+1}\right)^c\right) \ \leq \ \frac{2(2N+3d+2)}{|E_\ell|}\,, \tag{74}$$

Finally, combining Equations (72), (73) and (74), we have

$$\mathbb{P}\left(\xi^c_{\ell+1} \cup \left(\xi^-_{\ell+1}\right)^c \cup \left(\xi^+_{\ell+1}\right)^c\right) \ \leq \ \mathbb{P}\left(\xi^c_{\ell+1}\right) + \mathbb{P}\left(\left(\xi^-_{\ell+1}\right)^c\right) + \mathbb{P}\left(\left(\xi^+_{\ell+1}\right)^c\right) \ \leq \ \frac{9N+15d+8}{|E_\ell|}\,.$$

□

Lemma 8 (Lipschitz Property for F, F^- and F^+)**.** *The following hold for any $z_1, z_2 \in \mathbb{R}$:*

(i) $|F(z_1) - F(z_2)| \leq c_f |z_1 - z_2|$.

(ii) $|F^-(z_1) - F^-(z_2)| \leq 2c_f N^2 |z_1 - z_2|$.

(iii) $|F^+(z_1) - F^+(z_2)| \leq c_f N |z_1 - z_2|$.

Here, $0 < c_f = \sup_{z \in [-\epsilon_{\max}, \epsilon_{\max}]} f(z)$.

Proof. Proof of Lemma 8. Without loss of generality, we assume $z_1 < z_2$. Note that $F(z) = 0$ for $\forall z \in (-\infty, -\epsilon_{\max}]$, and $F(z) = 1$ for $\forall z \in [\epsilon_{\max}, \infty)$.

Part (i) We consider the following cases:

1. **Case 1:** ($z_1 < z_2 \leq -\epsilon_{\max}$ or $\epsilon_{\max} \leq z_1 < z_2$): $|F(z_2) - F(z_1)| = 0 \leq c_f |z_2 - z_1|$.
2. **Case 2:** ($-\epsilon_{\max} < z_1 < z_2 < \epsilon_{\max}$): By the mean value theorem, $|F(z_2) - F(z_1)| = f(\tilde{z})|z_2 - z_1| < c_f |z_2 - z_1|$, where $\tilde{z} \in (z_1, z_2)$.
3. **Case 3:** ($z_1 \leq -\epsilon_{\max} < z_2 < \epsilon_{\max}$): We have $|z_2 - (-\epsilon_{\max})| = z_2 - (-\epsilon_{\max}) \leq z_2 - z_1$ and $F(z_1) = F(-\epsilon_{\max}) = 0$. Hence $|F(z_2) - F(z_1)| = |F(z_2) - F(-\epsilon_{\max})| = f(\tilde{z})|z_2 - (-\epsilon_{\max})| \leq c_f |z_2 - z_1|$, where $\tilde{z} \in (-\epsilon_{\max}, z_2)$ by the mean value theorem.
4. **Case 4:** ($-\epsilon_{\max} < z_1 < \epsilon_{\max} \leq z_2$): We have $|\epsilon_{\max} - z_1| = \epsilon_{\max} - z_1 \leq z_2 - z_1$ and $F(z_2) = F(\epsilon_{\max}) = 1$. Hence $|F(z_2) - F(z_1)| = |F(\epsilon_{\max}) - F(z_1)| = f(\tilde{z})|\epsilon_{\max} - z_1| \leq c_f |z_2 - z_1|$, where $\tilde{z} \in (z_1, \epsilon_{\max})$ by the mean value theorem.

Part (ii) & (iii) We recall that $F^-(z) = NF^{N-1}(z) - (N-1)F^N(z)$ and $F^+(z) = F^N(z)$, so

$$\begin{aligned}
& |F^-(z_2) - F^-(z_1)| \\
= \ & \left|NF^{N-1}(z_2) - (N-1)F^N(z_2) - \left(NF^{N-1}(z_1) - (N-1)F^N(z_1)\right)\right| \\
\leq \ & N\left|F^{N-1}(z_2) - F^{N-1}(z_1)\right| + (N-1)\left|F^N(z_2) - F^N(z_1)\right| \\
= \ & N\left|(F(z_2) - F(z_1))\left(\sum_{n=1}^{N-1} (F(z_2))^{n-1}(F(z_1))^{N-1-n}\right)\right| \\
& \quad + (N-1)\left|(F(z_2) - F(z_1))\left(\sum_{n=1}^{N} (F(z_2))^{n-1}(F(z_1))^{N-n}\right)\right| \\
\leq \ & N(N-1)\left|F(z_2) - F(z_1)\right| + (N-1)N\left|F(z_2) - F(z_1)\right| \\
< \ & 2N^2 c_f |z_2 - z_1|\,.
\end{aligned}$$

The second equality uses $a^m - b^m = (a-b)\left(\sum_{n=1}^m a^{n-1}b^{m-n}\right)$ for any $a, b \in \mathbb{R}$ and integer $m \geq 2$. The second inequality follows from $F(z) \in [0,1]$ for $\forall z \in \mathbb{R}$. The final inequality follows from the Lipschitz property of F shown in part (i). Following the same arguments, we can also show that $|F^+(z_2) - F^+(z_1)| \leq c_f N |z_2 - z_1|$. □

D SUPPLEMENTARY LEMMAS

Lemma 9 (Dvoretzky-Kiefer-Wolfowitz Inequality (Dvoretzky et al. (1956)))**.** *Let $Z_1, Z_2, \dots Z_n$ be i.i.d. random variables with cumulative distribution function F, and denote the associated empirical distribution function as*

$$\widehat{F}(z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{Z_i \leq z\} \quad , z \in \mathbb{R} \,. \tag{75}$$

Then, for any $\bar{\gamma} > 0$,

$$\mathbb{P}\left(\sup_{z\in\mathbb{R}} \left|\widehat{F}(z) - F(z)\right| \leq \bar{\gamma}\right) \geq 1 - 2\exp\left(-2n\bar{\gamma}^2\right) \,. \tag{76}$$

Lemma 10 (Matrix Chernoff Bound (Tropp et al. (2015)))**.** *Consider a finite sequence of independent, random matrices $\{Z_k \in \mathbb{R}^d\}_{k\in[K]}$. Assume that $0 \leq \lambda_{\min}(Z_k)$ and $\lambda_{\max}(Z_k) \leq B$ for any k. Denote $Y = \sum_{k\in[K]} Z_k$, $\mu_{\min} = \lambda_{\min}(\mathbb{E}[Y])$, and $\mu_{\max} = \lambda_{\max}(\mathbb{E}[Y])$. Then for $\forall \bar{\gamma} \in (0,1)$,*

$$\mathbb{P}\left(\lambda_{\min}(Y) \leq \bar{\gamma}\mu_{\min}\right) \leq d\exp\left(-\frac{(1-\bar{\gamma})^2\mu_{\min}}{2B}\right) \,.$$

Lemma 11 (Multiplicative Azuma Inequality(Koufogiannakis and Young (2014)))**.** *Let $Z_1 = \sum_{\tau\in[\widetilde{T}]} z_{1,\tau}$ and $Z_2 = \sum_{\tau\in[\widetilde{T}]} z_{2,\tau}$ be sums of non-negative random variables, where $\widetilde{T}$ is a random stopping time with a finite expectation, and, for all $\tau \in [\widetilde{T}]$, $|z_{1,\tau} - z_{2,\tau}| \leq 1$ and $\mathbb{E}\left[(z_{1,\tau} - z_{2,\tau}) \mid \sum_{s<\tau} z_{1,s}, \sum_{s<\tau} z_{2,s}\right] \leq 0$. Let $\tilde{\gamma} \in [0,1]$ and $A \in \mathbb{R}$. Then,*

$$\mathbb{P}\left((1-\tilde{\gamma})Z_1 \geq Z_2 + A\right) \leq \exp\left(-\tilde{\gamma}A\right)$$