Incentive-aware Contextual Pricing with Non-parametric Market Noise

Negin Golrezaei
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139 golrezae@mit.edu

Patrick Jaillet
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 jaillet@mit.edu

Jason Cheuk Nam Liang
Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139 jcnliang@mit.edu

We consider a dynamic pricing problem for repeated contextual second-price auctions with strategic buyers whose goals are to maximize their long-term time discounted utility. The seller has very limited information about buyers' overall demand curves, which depends on d-dimensional context vectors characterizing auctioned items, and a non-parametric market noise distribution that captures buyers' idiosyncratic tastes. The noise distribution and the relationship between the context vectors and buyers' demand curves are both unknown to the seller. We focus on designing the seller's learning policy to set contextual reserve prices where the seller's goal is to minimize his regret for revenue. We first propose a pricing policy when buyers are truthful and show that it achieves a T-period regret bound of $\tilde{\mathcal{O}}(\sqrt{dT})$ against a clairvoyant policy that has full information of the buyers' demand. Next, under the setting where buyers bid strategically to maximize their long-term discounted utility, we develop a variant of our first policy that is robust to strategic (corrupted) bids. This policy 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. Because of this nice property, our robust policy enjoys a T-period regret of $\tilde{\mathcal{O}}(\sqrt{dT})$, matching that under the truthful setting up to a constant factor that depends on the utility discount factor.

Key words: repeated auctions, learning with strategic agents, incentive-aware learning, pricing

1. Introduction

We study the fundamental problem of designing pricing policies for highly heterogeneous items. This study is inspired by the availability of the massive amount of real-time data in online platforms and in particular, internet advertisement markets, where the seller has access to detailed information about items 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. 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 for both truthful and strategic buyers. The pricing policies should overcome the challenge of having limited information about buyers' willingness-to-pay, address the general learning problem for non-parametric demand curves, and resolve the issues that arise along with buyers' strategic behavior.

Formally, we consider a setting wherein any period t, the seller sells one item to buyers via running a second price auction with a reserve price. Note that in the second price auction, which is a prevalent selling mechanism in online advertising markets, the item is allocated to a buyer with the highest submitted bid as long as her bid is greater than or equal to the reserve price, and the winner pays the maximum value between the reserve price and the second-highest bid.

The item is characterized by a d-dimensional feature vector x_t, observed by the seller and buyers. We consider an interdependent contextual valuation model in which the buyer's valuation for the item is the sum of common and private components. The common component, which is the same across all the buyers, is a function of the feature vector; and the private component, which captures buyers' idiosyncratic tastes, is independently drawn from an unknown *non-parametric* noise distribution F. The common component determines the expected willingness-to-pay of the buyers and is the inner product of the feature vector and a fixed scaling factor β, which we call the "mean vector". We note that such a linear valuation model is very common in the literature of dynamic pricing; see, for example, Golrezaei et al. (2018a), Javanmard and Nazerzadeh (2016), Kanoria and Nazerzadeh (2017) and Javanmard (2017).

Under this interdependent and contextual valuation model, we study two settings, each of which characterizes different types of buyer behavior. In the first setting, called the *truthful setting*, buyers always bid their true valuations, and our main focus lies in obtaining accurate estimates of the unknown and non-parametric demand through buyers' truthful bids in order to develop revenue-maximizing pricing strategies. In the second setting, called the *strategic setting*, buyers may submit *corrupted*, i.e., untruthful, bids in order to trick the seller to lower future reserve prices. In this setting, our goal is to effectively learn buyers' demand in the presence of such corrupted data (bids). In particular, we associate this type of strategic bidding with the setup that buyers aim to maximize their long-term discounted utility. We point out that this discounted utility model has recently been used in the literature, for instance, see Amin et al. (2013, 2014), Golrezaei et al. (2018a), and Liu et al. (2018).

The main problem of interest is to design policies that dynamically learn/optimize contextual reserve prices against both truthful and strategic bidding behaviors. We take the perspective of a typical seller who is not aware of the mean vector β and noise distribution F, and aims to minimize his regret which is the revenue loss against a clairvoyant policy that knows both β and F. In fact, having full knowledge of β and F corresponds to knowing the optimal mapping between the observed context vector and the revenue-maximizing reserve prices. Hence, this benchmark will always set the optimal contextual reserve price, which eliminates buyers' incentive to submit corrupted bids. Yet, for our problem of interest, when β and F are unknown, in order to learn this optimal mapping the seller needs to learn the buyers' demand curves, which is associated with three main challenges: (i) the demand curve is constantly shifting due to the variations of the feature vector changes 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*[1] assumption on F, which is a common assumption in related literature, but has been shown to be unrealistic in online advertising markets (Celis et al. (2014) and Golrezaei et al. (2017)); (iii) in the strategic setting, buyers take advantage of the seller's lack of knowledge in demand through submitting corrupted bids to manipulate future reserve prices.

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

As our main contribution in this work, in light of the three aforementioned challenges, we present two effective dynamic contextual pricing policies with sublinear T-period regret of $\widetilde{\mathcal{O}}(\sqrt{T})$,[2] where the first policy lends itself to the truthful setting, and the second policy is tailored to the strategic setting. In the truthful setting, the seller encounters the challenges regarding both constant shifts and unknown shape of the demand curve; while in the strategic setting, the seller faces the additional challenge to learn buyers' demand (or valuation distribution) using corrupted bids. Designing a pricing policy against strategic buyers is more challenging than that against truthful buyers. However, by first studying the pricing problem in the truthful setting, we shed light on some key techniques that enable us to attain accurate estimates of the mean vector and the market noise distribution. In the strategic setting, we then build on these techniques to resolve issues due to intentionally corrupted bids and buyers' strategic behavior. We refer to our policy under the truthful setting as the NPAC-T policy, where "NPAC" stands for Non-Parametric Contextual and "T" refers to the truthful setting. Similarly, we refer to our policy under the strategic setting as NPAC-S, where "S" stands for the strategic setting. In the remaining part of the introduction, we briefly discuss these two policies.

NPAC-T Policy: The NPAC-T policy is designed for the truthful setting. In this policy, we take advantage of the fact that the optimal reserve price in truthful environments can be written as a function of the distributions of the highest and second-highest valuations; see Proposition 1. In light of this observation, to learn the optimal reserve prices, instead of the market noise distribution F, we estimate these distributions using all historical bids/valuations. We notice that the distribution of the highest and second-highest bids shift according to the time-varying feature vector x_t and unknown mean vector β. Thus, to estimate these distributions, we first estimate the mean vector by applying an ordinary least squares (OLS) estimator using all historical bids, and then take advantage of our most up-to-date estimate of β to estimate the distributions. We stress that because the noise distribution F is not parametric, joint estimation of the mean vector

[2] $\widetilde{\mathcal{O}}(\cdot)$ hides logarithmic factors.

β and distributions using a maximum likelihood estimator is not possible. Hence, we construct empirical distributions for the distributions of interest by carefully controlling for the errors of β that may propagate into our estimates. Finally, we use these estimates as surrogates of their ground-truth counterparts to set reserve prices, and show that our policy enjoys a low regret of the order $\tilde{\mathcal{O}}\left(d\sqrt{T}\right)$. This regret bound is presented in Theorem 1.

NPAC-S Policy: The NPAC-S policy is designed for the strategic setting and is robust against corrupted bids. It is based on two modifications of the NPAC-T policy. First, the NPAC-S policy partitions the entire time horizon into phases of increasing length, and instead of using all historical bids to update its estimates for the mean vector and the distributions of the second-highest and highest valuations, it only uses the data in the previous phase. This will reduce 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[3]; 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. Put differently, other buyers who are not chosen have a zero chance of winning the item regardless of their submitted bids. For the "isolated" buyer, the reserve price is chosen uniformly at random. These randomized isolation periods exploit the fact that buyers aim to maximize their long-term discounted utility, and motivate the buyers to not corrupt their bids significantly for a large number of periods. In the isolation periods, when the valuation of the isolated buyer is greater than the reserve price, underbidding can lead to utility loss. Similarly, when the valuation of the isolated buyer is smaller than the reserve price, overbidding can also result in utility loss.

We point out that the estimation techniques used in NPAC-S policy are the same as those in NPAC-T policy, even though the OLS estimator and empirical distributions are both vulnerable to large outliers/corruptions. The NPAC-S policy can still use OLS estimators and empirical distributions because, in virtue of our isolation periods, the number of past periods with large corruptions

[3] Buyers are aware of the randomized isolation periods as the seller announces and commits to his pricing policy.

is $\mathcal{O}(\log(t))$ for any period t with high probability. Put differently, the number of outliers in the submitted bids is small and as a result, there is no need to redesign the estimation techniques used in NPAC-T for the strategic setting. We highlight that prior to this work, the best regret bound for the same setting is $\mathcal{O}(dT^{2/3})$, given by Golrezaei et al. (2018a) who also proposes an algorithm that attains a regret of $\tilde{\mathcal{O}}(d\sqrt{T})$ for a parametric noise distribution F (via jointly estimating β and F using a maximum likelihood estimator). In Theorem 2, we show that the NPAC-S policy achieves a regret of $\tilde{\mathcal{O}}\left(d\sqrt{T}\right)$ for general non-parametric distributions F.

The rest of the paper is organized as followed. In Section 2, we review the literature related to our work, and in Section 3 we formally define the interdependent valuation model, as well as both the truthful and strategic settings. Section 4 defines the clairvoyant benchmark policy. In Section 5 and 6, we present and analyze NPAC-T and NPAC-S policies, respectively. Finally, we conclude in Section 7. Appendix 8 provides a proof regarding the benchmark policy, and Appendices 9 and 10 provide the proofs of the regret bounds for NPAC-T and NPAC-S, respectively.

2. Related Work

Our work lies in the intersection of dynamic pricing and online learning, which is an area that has attracted increasing interest in recent years according to an extensive survey by den Boer (2015), partially due to the booming activities in online marketplaces.[4]

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.

[4] Our work also lies in the domain of behavior-based pricing, where the seller learns from past behavior of buyers and adaptively updates pricing decisions. Behavior-based pricing, as well as price discrimination, have been studied extensively under different settings; see Esteves et al. (2009). Two relevant works include Fudenberg and Villas-Boas (2006) and Bikhchandani and McCardle (2012), each of which considers two-period interactions (non-contextual) between a seller and a strategic buyer, while our work focuses on multi-period interactions, associated with varying contexts, between a seller and multiple strategic buyers.

den Boer and Zwart (2013), Besbes and Zeevi (2009), and 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. 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 an 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), and 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. They present a tree-based non-parametric learning policy that adaptively splits the feature space into smaller bins (hyper-rectangles), and eventually learns near-optimal decisions in each bin. However, since their methodology is designed to handle very general objectives and not specifically tailored to pricing problems, within the context of dynamic pricing, 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. Shah et al. (2019) studied a dynamic pricing problem in repeated posted price mechanisms. They considered a model 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).

We now review some related work that studies dynamic pricing facing non-myopic buyers who may behave strategically to enjoy lower prices in the future.[5] We first highlight the significance of such research areas as there exists empirical evidence that shows buyers in online advertising markets indeed exhibit strategic behaviour; see Edelman and Ostrovsky (2007), Golrezaei et al. (2018b). Amin et al. (2013) studies the dynamic pricing problem in a posted price auction with a single buyer under the non-contextual setting where buyer valuation is fixed, or drawn from some fixed but unknown distribution. They introduce the notion of *strategic regret*, which is the revenue loss measured against a truthful buyer, and show that no sublinear strategic regret is achievable when buyer future utility is not time-discounted. Amin et al. (2014) proposes pricing policies for both a single truthful buyer and a strategic buyer using a contextual valuation model with no market noise disturbance. In the strategic-buyer setting, their algorithm achieves a regret of $\widetilde{\mathcal{O}}(T^{2/3})$, in contrast with our regret of $\widetilde{\mathcal{O}}(\sqrt{T})$ using the NPAC-S policy presented in Section 6. We point out that this is because, apart from the difference in buyers' valuation model, their posted price mechanism is associated with bandit feedback, which means the seller can only observe the outcome of the auction, while in our setting we assume that seller can examine all submitted bids.[6] In our work, we can handle the scenario of multiple buyers, and unlike Amin et al. (2013, 2014), we also need to deal with the issue of learning a non-parametric distribution function.

[5] 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) and Feng et al. (2019).

[6] This assumption is justified when with a small positive probability, the seller posts a reserve price of zero, incentivizing the buyers to submit positive bids in all the auctions.

Kanoria and Nazerzadeh (2017) also considers a buyer valuation model similar to ours and studies the incentive compatible properties of the second-price auction with personalized reserve prices in a setting where (i) the seller has access to a training dataset of submitted bids with a given size, (ii) buyers do not discount their future utilities, and (iii) the market noise distribution is MHR. The authors design a static (one-shot) pricing scheme whose revenue is ϵ away from the optimal revenue when the size of the training dataset is $\mathcal{O}(\frac{\log(1/\epsilon)}{\epsilon^2})$ whereas we design dynamic pricing schemes that minimize the seller's cumulative regret over a finite time horizon, compelling us to take into account the lasting effects of previous pricing decisions. Nonetheless, if one assumes that the training data are obtained from pure exploration periods, the static pricing scheme in Kanoria and Nazerzadeh (2017) can lead to a dynamic pricing scheme whose regret is in the order of $\mathcal{O}(T^{2/3})$.

Golrezaei et al. (2018a) also study a linear valuation model with unknown market noise and utility-maximizing buyers. Our work distinguishes itself from Golrezaei et al. (2018a) in two major ways. First, they focus on a setting where the market noise distribution can vary arbitrarily over time within a known class of distributions, and take a robust approach to design a pricing policy that works for any distribution within the class. In contrast, our work relaxes this constraint and does not require the seller to have any prior knowledge on the time-invariant noise distribution F. Second, in their setting, the seller only utilizes the outcome of the auctions to learn buyer demand and results in a regret of $\widetilde{\mathcal{O}}(T^{2/3})$. 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 a regret of $\widetilde{\mathcal{O}}(\sqrt{T})$ against a clairvoyant benchmark policy. We note that very recently Deng et al. (2019) build on the result of Golrezaei et al. (2018a) 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. (2018a).

Finally, we briefly discuss some recent literature within the domain of mechanism design and online learning that adopt methodologies from differential privacy to deal with strategic agents. McSherry and Talwar (2007) first establish the connections between differential privacy and mechanism design with the presence of strategic agents. Inspired by differential privacy assumptions, the authors assume that agents have limited manipulating power on the outcome of the mechanism, and with this assumption, they design mechanisms under which the agents have limited incentive to be untruthful. Yet, unlike our setting, their focus lies in designing (approximately) incentive compatible static mechanisms. Other works include: Mahdian et al. (2017) which studies a static non-contextual setting and aims to select feasible mechanisms (based on a fixed training dataset) that incentivize buyers to behave truthfully in the first place; and Liu et al. (2018) which considers an adversarial and non-contextual buyers' valuation model while we consider the stochastic setting in which buyer valuations are generated randomly according to a non-parametric and unknown distribution.

3. Model and Preliminaries

Notation: For any $a \in \mathbb{N}^+$, denote $[a] = \{1, 2, \ldots, a\}$. For two vectors $x, y \in \mathbb{R}^d$, we 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 $x_t \in \mathcal{X} \subset \{x \in \mathbb{R}^d : \|x\|_\infty \le x_{\max}\}$ 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 $\mathcal{D}$.[7] We assume that Σ is positive definite and unknown to the seller.

We focus on an interdependent valuation model in which buyers' valuation has a common value component and a private value component. Precisely, valuation of buyer $i \in [N]$ at time $t \in [T]$,

[7] 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 $\mu = \mathbb{E}_{x\sim\mathcal{P}}[x]$.

denoted by $v_{i,t}$, is given by $v_{i,t} = \langle \beta, x_t \rangle + \epsilon_{i,t}$, where β, called the *mean vector*, is fixed over the entire time horizon and unknown to the seller. The term $\langle \beta, x_t \rangle$ is the common value component that represents the expected (mean) buyers' valuation and $\epsilon_{i,t}$ is some mean zero market noise that captures the idiosyncratic tastes of buyers. The noise is drawn from some time-invariant distribution F with probability density function f, both unknown to the seller. We assume that market noise variables $\{\epsilon_{i,t}\}_{t\in[T],i\in[N]}$ are drawn independently from the distribution F. 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 by the seller. We further assume that for any $i \in [N]$ and $t \in [T]$, valuation $v_{i,t}$ is upper-bounded by $v_{\max}$, where the seller and buyers know $v_{\max}$, i.e., the maximum possible revenue/reward of the seller in an auction. We note that this type of boundedness assumption is common in the related literature, where the bound for maximum reward is typically set to 1; see, for example, Amin et al. (2013) and Cesa-Bianchi et al. (2015).

We highlight that our setting is general in the sense that we do not enforce distribution F to be parametric nor to satisfy the MHR assumption. We highlight that via analyzing real auction data sets, it has been shown that MHR does not necessarily hold in online advertising markets (Celis et al. 2014, Golrezaei et al. 2017). Thus, by intentionally not enforcing this assumption, our setting can better model realistic and practical environments in online advertisement markets.

Repeated Second Price Auctions with Reserve: As stated earlier, in each period $t \in [T]$, the seller runs a second price auction with anonymous reserve price r_t, where r_t is a function of the context vector x_t and the history set $\mathcal{H}_{t-1}$ that includes all the seller's observation in periods $1, 2, \ldots, t-1$:

$$\mathcal{H}_{t-1} := \{(r_1, b_1, x_1), (r_2, b_2, x_2), \ldots, (r_{t-1}, b_{t-1}, x_{t-1})\},$$

where we define $b_\tau := (b_{1,\tau}, b_{2,\tau}, \ldots, b_{N,\tau})$ to be the vector of all bids in period $\tau \in [t]$. Without loss of generality, we assume that all bids and valuations are never equal to one another. We now proceed to describe the second price auctions with reserve when the number of buyers $N \geq 2$. In any period $t \geq 1$:

(i) The seller observes the context vector $x_t \sim \mathcal{D}$ and reveals it to the buyers.

(ii) The seller computes reserve price r_t as a function of the context vector x_t and history set $\mathcal{H}_{t-1}$.

(iii) Each buyer $i \in [N]$, forms individual valuations $v_{i,t}$ and submit a bid $b_{i,t}$ to the seller.

(iv) Let $i^\star = \arg\max_{i\in[N]} b_{i,t}$.[8] If $b_{i^\star,t} \ge r_t$, the item is allocated to buyer $i^\star$, i.e., the buyer with the highest submitted bid, and he is charged the maximum between the reserve price and second highest bid; that is, $\max\{r_t, \max_{i\neq i^\star} b_{i,t}\}$. For any other buyer $i \neq i^\star$, payment is zero. If $b_{i^\star,t} < r_t$, the item is not allocated and all the payments are zero.

For the special case when there is only $N = 1$ buyer, the auction only differs in step (iv) of the multiple-buyer mechanism: the item is allocated to the buyer if her bid b_t is greater than or equal to the reserve r_t, and her payment is r_t if the item is allocated and 0 otherwise. In this paper, we mainly focus on the case of $N \ge 2$ buyers as analyzing the multi-buyer case is more challenging, especially when bidders are strategic. Yet we point out that our proposed algorithms can be easily generalized to the single-buyer second price auction.

Buyers' Bidding Behavior: With respect to buyers' bidding behavior, we propose pricing policies under two different settings: Truthful and strategic settings. In the truthful setting, buyers submit their true valuations, i.e., $b_{i,t} = v_{i,t}$ for any $i \in [N]$ and $t \in [T]$, and in the strategic setting, the submitted bids and true valuations of a buyer may not necessarily be equal. Under 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, $w_{i,t} := \mathbb{I}\{i \text{ wins the item in period } t\}$ indicates whether buyer i wins the item in period t according to the allocation rule of the mechanism, $p_{i,t}$ is the payment of buyer i in period t according the payment rule of the mechanism, and the expectation is taken with

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

respect to the randomness due to the noise distribution F, the context distribution $\mathcal{D}$, and buyers' bidding behavior. 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. (2018a), 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 an opportunity of showing their ads to targeted users. 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.

In the strategic setting, buyers may from time to time submit "corrupted bids", i.e., underbid (shade their bid) or overbid with respect to their true valuations, sacrificing current utility with the aim to lower future reserve prices and increase their future long-term utility. We assume that seller announces his pricing policy to all buyers so that buyers have full knowledge of the seller's learning and pricing algorithm and has the freedom to adopt any bidding strategy to maximize their long term utility.

We now describe the scope of feasible buyer bidding behavior in the strategic setting. Recall that the maximum possible valuation $v_{\max}$ is known to both buyers and the seller. Thus, buyers have no incentive to submit a bid greater than $v_{\max}$, i.e., $b_{i,t} \leq v_{\max}$ for all $i \in [N], t \in [T]$ as the seller only sets reserve prices less than or equal to $v_{\max}$. Furthermore, buyers submit bids satisfying the following relationship:

$$b_{i,t} = v_{i,t} - a_{i,t} \,,$$

where $a_{i,t}$ 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. Since the seller observes buyers' bids instead of their true valuations in the strategic setting, corrupted bids may deteriorate the estimation accuracy of buyers' demand, and as a result negatively impact pricing decisions in future periods. There are no restrictions on the degree of corruption $a_{i,t}$ for a buyer i in period t other than it is bounded, i.e.,

$$|a_{i,t}| \le a_{\max}, \quad i \in [N], t \in [T]. \tag{2}$$

Such a bound for $a_{i,t}$ is natural as buyers always submit non negative bids and all bids are bounded by $v_{\max}$.

We compare our policies in both truthful and stratgic settings to the same benchmark policy, detailed in the next section, that has full knowledge of the mean vector β and noise distribution F. We will further discuss the truthful and strategic settings in Section 5 and Section 6, respectively.

4. Benchmark and Seller's Regret

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

$$\mathsf{rev}_t(r_t) := \mathbb{E}\Big[\sum_{i\in[N]} p_{i,t}(r_t, b_t) \;\Big|\; x_t, r_t\Big], \tag{3}$$

where $p_{i,t}(r_t, b_t)$ is the payment of buyer i in period t, $b_t = (b_{1,t}, b_{2,t}, \ldots, b_{N,t})$ is the vector of all bids in period t, and the expectation in Equation (3) is taken with respect to the noise distribution in period t and any randomness in the bids submitted by buyers in period t (as buyers' bidding strategies may be randomized). In the truthful setting, we have $b_{i,t} = v_{i,t}$ for all $i \in [N]$ and $t \in [T]$. Note that the reserve price r_t may also be random and depend on the history set $\mathcal{H}_{t-1}$ as well as x_t. In the rest of the paper, for simplicity, we will use the following shorthand notation to represent payments: $p_{i,t} := p_{i,t}(r_t, b_t)$.

As discussed in the previous section, in a second price auction, the highest bidder wins the item and is charged a payment that is the maximum of the second-highest bid and the reserve price set by the seller, while no transaction occurs if the highest bid is less that the reserve price. Hence, buyer i's payment conditioned on the context $x_t \in \mathcal{X}$ for setting reserve r_t is given by

$$p_{i,t} \;=\; \max\{b_t^-, r_t\}\mathbb{I}\{b_{i,t} \geq \max\{b_t^+, r_t\}\}\,. \tag{4}$$

Here, b_t^- and b_t^+ are the second-highest and highest bids in period t, respectively. When the number of buyers $N = 1$, we set b_t^- to zero.

Maximizing seller's expected revenue is equivalent to minimizing his regret, where the regret is computed against a benchmark policy that knows the mean vector β as well as the non-parametric noise distribution F. As stated earlier, we use this benchmark to evaluate our policies in both the truthful and strategic settings. This is, in fact, a rather strong benchmark because it corresponds to the seller with full knowledge of buyer demand, which in turn knows the optimal contextual reserve price that maximizes expected revenue with respect to the current context. Hence, this seller will always set the optimal reserve price, and there will be no incentive for buyers' to corrupt their bids. To formally define this benchmark, 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 expected revenue is given by*

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

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 8. 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 $t \in [T]$. That is, ϵ_t^+ and ϵ_t^- are the N^{th} and $(N-1)^{th}$ order statistics of N independent random samples from distribution F. Note that this proposition also provides us with an alternative, and rather convenient expression for the seller's expected revenue in the truthful setting as opposed to that given in Equation (3). However, we point out that this expression is not equivalent to the seller's expected revenue when buyers are bidding untruthfully, and in that case, we must return to analyzing Equation (3).

Given Proposition 1, we define the benchmark 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

$$r^\star(x) = \arg\max_{y \geq 0} \left\{ \int_0^y F^-(z - \langle \beta, x \rangle) dz - y \left(F^+(y - \langle \beta, x \rangle) \right) \right\} . \tag{5}$$

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 given by

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

We observe that this benchmark sets the reserve price that maximizes the expected revenue under truthful buyer behavior illustrated in Proposition 1. In fact, the benchmark provides an optimal mapping from the feature vector x_t to reserve price $r^\star(x_t)$, and this mapping remains unchanged over the entire horizon because the mean vector β and noise distribution F are time-invariant. This echoes our earlier point that, in our consideration, a seller who possesses full knowledge of buyer demand will not deviate from setting the optimal contextual reserve. We highlight that this benchmark offers contextual prices and as a result, obtaining a low regret with respect to this strong benchmark is challenging. In non-contextual environments, regret is measured against a policy that offers a fixed reserve price so it is sufficient to learn a single optimal value, whereas, in contextual environments, the regret is measured against optimal prices that are context-dependent which requires learning the entire optimal mapping.

Here, we make several remarks. When distribution F satisfies the MHR assumption, the objective function of the optimization problem in Equation (5) is unimodal in the decision variable y, and according to Golrezaei et al. (2018a), $r^\star(x)$ can be simplified as follows:

$r^\star(x) = \arg\max_{y\geq 0}\{y(1-F(y-\langle\beta,x\rangle))\}$. However, as discussed in the introduction, we do not assume that the market noise distribution is MHR as in various related literature; see Roughgarden and Schrijvers (2016), Kanoria and Nazerzadeh (2017), and Golrezaei et al. (2018a). 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). Hence, although we are unable to simplify the expression for the optimal mapping which imposes additional challenges to the analysis, our model offers greater generality.

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, where r_t^π is a function of the context vector x_t and may or may not depend on the history $\mathcal{H}_{t-1}$. Then, the regret of policy π in period t and its cumulative T-period regret are defined as:

$$\mathsf{Regret}_t^\pi = \mathbb{E}\left[\mathrm{REV}_t^\star - \mathsf{rev}_t(r_t^\pi)\right] \quad \text{and} \quad \mathsf{Regret}^\pi(T) = \sum_{t\in[T]} \mathsf{Regret}_t^\pi\,, \tag{6}$$

where the optimal revenue $\mathrm{REV}_t^\star$ 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. We point out that in the calculations of $\mathsf{rev}_t(r_t^\pi)$ for any policy π, the buyers may not necessarily behave truthfully. Our goal is to design a policy that obtains a low regret for any mean vector β, noise distribution F, and feature distribution $\mathcal{D}$.

5. Truthful Setting: NPAC-T Policy

In this section, we study the problem of learning contextual reserve prices in the setting where all buyers are truthful and bid their true valuations, i.e. $b_{i,t} = v_{i,t}$ for any $i\in[N]$ and $t\in[T]$. Hence, here we will use the terms "bids" and "valuations" interchangeably. Recall that the seller observes bids of all the N buyers in each period. To design a pricing policy under this setting, we need to learn the mean vector β and the noise distribution F simultaneously. For the purpose of motivating our policy design, suppose that the mean vector β is known to the seller. Recall that, conditioned on x_t, F^+ and F^- are respectively the distributions of $\epsilon_t^+ = v_t^+ - \langle\beta,x_t\rangle$ and $\epsilon_t^- = v_t^- - \langle\beta,x_t\rangle$.

Therefore, one can naturally estimate F^+ and F^- via constructing their corresponding empirical distributions using the highest and second-highest bids and knowledge of the mean vector β. More specifically, in period $t \in [T]$, one can construct the most up-to-date empirical distributions

$$\hat{F}_t^+(z) = \frac{1}{t-1} \cdot \sum_{\tau \in [t-1]} \mathbb{I}\{v_t^+ - \langle \beta, x_t \rangle \leq z\} \quad \text{and} \quad \hat{F}_t^-(z) = \frac{1}{t-1} \cdot \sum_{\tau \in [t-1]} \mathbb{I}\{v_t^- - \langle \beta, x_t \rangle \leq z\},$$

and use these estimates as surrogates for F^+ and F^- respectively in Equation (5) to obtain an approximation of the optimal reserve price. However, one of the main challenges is that the mean vector β is unknown, and as a result, to get any estimate of F^+ and F^-, we would need to replace β in the empirical distributions (particularly within the indicator functions) with some estimate $\hat{\beta}$, whose estimation error could then potentially propagate into the estimation errors in distributions F^+ and F^-. We will later discuss how we resolve this error propagation issue. We highlight that because distribution F is non-parametric, a joint estimation of the distributions and mean vector via a maximum likelihood estimator is not possible. Yet, despite the challenges that arise from embedding the estimate for β in the estimates for F^+ and F^-, this intuition provides a rather straightforward direction to design our estimation and pricing policy.

In light of the previous discussion, we now describe our policy, called NPAC-T. The policy sequentially estimates β, F^-, and F^+ as more data points are observed, and uses its most recent estimate of β and F^-, and F^+ to set reserve prices. Let $\hat{\beta}_t$ be the estimate of the mean vector β and $\hat{F}_t^-(\cdot)$, $\hat{F}_t^+(\cdot)$ be the estimates of $F^+(\cdot)$ and $F^-(\cdot)$ respectively, at the beginning of period t. (We will define these estimates later.) Then, inspired by the optimal reserve price stated in Equation (5), the policy sets the reserve price r_t by replacing β, F^-, and F^+ with their estimates, as shown in Equation (7).

We now proceed to describe our estimation procedure. In any period $t \in [T]$, we use the average of submitted bids across all the previous periods, i.e., $\bar{v}_\tau = (\sum_{i \in [N]} v_{i,\tau})/N$ for any $\tau \in [t-1]$, to estimate the mean vector using an Ordinary Least Squares (OLS) estimator. Our estimate of β at the beginning of period t—before the auction is run—is denoted by $\hat{\beta}_t$ and is given in Equation (8). In Equation (8), $A^\dagger$ represents the pseudo inverse of a matrix A, so if A is invertible, we have

$A^\dagger = A^{-1}$.[9] We then use the most recent estimate of β, i.e., $\hat{\beta}_t$, to update our estimate of F^+ and F^-; see Equation (9). We also note that the policy does not require any pure exploration periods. The reason is that at the end of each auction, all bids are revealed to the seller and as a result, the seller can evaluate the revenue that he could have obtained had he posted any other reserve price other than r_t.

Algorithm 1 NPAC-T Policy: Non-Parametric Contextual Policy against Truthful Buyers

Initialize: $\hat{\beta}_1 = 0$, and $\hat{F}_1^-(z) = \hat{F}_1^+(z) = 0$ for $\forall z \in \mathbb{R}$.

For $t \geq 1$, observe $x_t \sim \mathcal{D}$. Then,

- **Set reserve price:**

$$r_t = \arg\max_{y \in [0, v_{\max}]} \left\{ \int_0^y \hat{F}_t^- \left(z - \langle \hat{\beta}_t, x_t \rangle \right) dz - y \left(\hat{F}_t^+ (y - \langle \hat{\beta}_t, x_t \rangle) \right) \right\} . \tag{7}$$

- **Update the estimate of the mean vector:**

$$\hat{\beta}_{t+1} = \Big(\sum_{\tau \in [t]} x_\tau x_\tau^\top \Big)^\dagger \Big(\sum_{\tau \in [t]} x_\tau \bar{v}_\tau \Big) \quad \text{where} \quad \bar{v}_\tau = \frac{1}{N} \sum_{i \in [N]} v_{i,\tau} \, . \tag{8}$$

- **Update the estimate of F^+ and F^-:** For any $z \in \mathbb{R}$, we have

$$\hat{F}_{t+1}^+(z) = \frac{1}{t} \sum_{\tau \in [t]} \mathbb{I}(v_\tau^+ - \langle \hat{\beta}_{t+1}, x_\tau \rangle \leq z) \quad \text{and} \quad \hat{F}_{t+1}^-(z) = \frac{1}{t} \sum_{\tau \in [t]} \mathbb{I}(v_\tau^- - \langle \hat{\beta}_{t+1}, x_\tau \rangle \leq z) \, . \tag{9}$$

Here, we briefly address how one can solve the optimization problem in Equation (7). The key observation is that for any period t, $\hat{F}_t^+(\cdot)$ is a step function with jumps at points in the finite set $B_t^+ := \{ v_\tau^+ - \langle \hat{\beta}_t, x_\tau \rangle \}_{\tau \in [t]}$. Furthermore, $\hat{F}_t^+(z) = 0$ for any $z < \min B_t^+$ and $\hat{F}_t^+(z) = 1$ for any $z \geq \max B_t^+$. Similarly, we have a corresponding finite set $B_t^- := \{ v_\tau^- - \langle \hat{\beta}_t, x_\tau \rangle \}_{\tau \in [t]}$ which includes all jump points for $\hat{F}_t^-(\cdot)$. This implies that in order to solve for the optimal reserve price r_t in Equation (7), it suffices to conduct a grid search for $\forall y \in B_t^+ \cup B_t^-$. More specifically, we let $\{ z^{(0)}, z^{(1)}, \ldots z^{(M)} \}$ be the ordered list (in increasing order) of all elements in $B_t^+ \cup B_t^- \cup \{0\}$, where

[9] In Lemma 3, we show that with high probability, $\sum_{\tau \in [t-1]} x_\tau x_\tau^\top$ in Equation (8) is positive definite, and hence invertible.

$z^{(0)} := 0$ and $M := |B_t^+ \cup B_t^-|$ (here, we assumed that $0 \notin B_t^+ \cup B_t^-$ without loss of generality). Hence,

$$r_t \;=\; \arg\max_{m\in[M]} \left\{ \sum_{j=1}^{m} \hat{F}_t^- \left(z^{(j)} - \langle \hat{\beta}_t, x_t \rangle \right) \cdot (z^{(j)} - z^{(j-1)}) - z^{(m)} \left(\hat{F}_t^+ (z^{(m)} - \langle \hat{\beta}_t, x_t \rangle) \right) \right\} .$$

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).

Finally, the following theorem is the main result of this section and upper bounds the regret of the NPAC-T policy in the truthful setting.

THEOREM 1 **(Regret of NPAC-T Policy)**. *Consider the case of $N \geq 2$ buyers. Then, in the truthful setting, the T-period regret of the NPAC-T policy is in the order of $\mathcal{O}\left(c_f \sqrt{dN^3 T \log(T)}\right)$, where the regret is computed against the benchmark in Definition 1 that knows the mean vector β and noise distribution F. Here, $c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z) > 0$ where f is the the probability density function of the market noise distribution F.*

Observe that through constant c_f, the regret bound provided in Theorem 1 depends on the particular instance of F. Here, we point out that the regret bound deteriorates as the market noise instance is more difficult to learn, i.e., when the probability density is rather concentrated in various regions on the real line. The dependency on N (i.e. the factor $\mathcal{O}\left(N^{3/2}\right)$) is due to the Lipschitz properties of F, F^- and F^+ which are induced by the boundedness of the pdf f, as illustrated in Lemma 1. That is, if it is directly assumed that distributions F, F^- and F^+ satisfy some Lipschitz conditions (that do not depend on N), we can improve the regret of the NPAC-T policy to $\mathcal{O}\left(c_f \sqrt{dT\log(T)}/\sqrt{N}\right)$.

The detailed proof of Theorem 1 is presented in Appendix 9, and here, we provide an outline for our analysis to upper bound the regret of the NPAC-T policy.

Proof sketch of Theorem 1 We first decompose the single period regret into the revenue loss due to estimation errors for β, F^- and F^+, and hence, the proof boils down to bounding these errors.

In Lemma 1, we first show Lipschitz properties for the distributions F, F^- and F^+. Next, as stated earlier, one of the main challenges is that estimation errors for β will further propagate into the estimation errors in F^- and F^+. Thus, in Lemma 2 we bound the estimation error $|F^-(z) - \hat{F}_t^-(z)|$ and $|F^+(z) - \hat{F}_t^+(z)|$ for any $z \in \mathbb{R}$ by controlling for the estimation errors in β. That is, we bound the estimation error in the distributions by assuming that our estimation error for β is small, namely $\|\hat{\beta}_t - \beta\|_1 \le \delta_t / x_{\max}$ where $\delta_t = \mathcal{O}(\sqrt{\log(t)/t})$. Here, we highlight that one cannot naively apply cdf-based concentration inequalities (e.g. the Dvoretzky-Kiefer-Wolfowitz Inequality) to bound the distribution estimation errors. This is because the estimates $\hat{F}_t^+(\cdot)$ and $\hat{F}_t^-(\cdot)$ evaluated at any point $z \in \mathbb{R}$ are biased. To see why note that because of the estimation error in the mean vector β, $v_\tau^+ - \langle \hat{\beta}_{t+1}, x_\tau \rangle$ and $v_\tau^- - \langle \hat{\beta}_{t+1}, x_\tau \rangle$ in the expressions of $\hat{F}_t^+(\cdot)$ and $\hat{F}_t^-(\cdot)$ in Equation (9) are not realizations of ϵ_τ^- and ϵ_τ^+, respectively. In particular, $\mathbb{E}\left[\hat{F}_t^-\big(z - \langle \hat{\beta}_t, x_t \rangle\big)\right] \neq F\big(z - \langle \hat{\beta}_t, x_t \rangle\big)$. This also sheds light on the more subtle and challenging issue: the estimates $\hat{F}_t^+(\cdot)$ and $\hat{F}_t^-(\cdot)$ are evaluated at points which may be random variables that depend on all information up to time t. Observe that the expression $F^-\big(z - \langle \hat{\beta}_t, x_t \rangle\big)$ is itself a random variable since $\hat{\beta}_t$ depends on the entire past history up to the current period. The same issue occurs when bounding the error for $\hat{F}_t^+(\cdot)$.

The proof of Lemma 2 provides a more detailed discussion as well as a solution to such issues, and, by applying the Lipschitz properties of F^- and F^+, shows that the errors in estimating these distributions are bounded by $\gamma_t + 2c_f N^2 \delta_t$ and $\gamma_t + c_f N \delta_t$ with high probability, respectively, where $\gamma_t = \mathcal{O}(\sqrt{\log(t)/t})$ and $c_f = \sup_{z \in [-\epsilon_{\max}, \epsilon_{\max}]} f(z)$. The term $\delta_t = \mathcal{O}(\sqrt{\log(t)/t})$ can be viewed as the estimation error for β, which echoes our earlier point that estimation error in the mean vector propagates into the inaccuracy of the distributions' estimates. Moreover, the dependency on c_f shows that the estimation accuracy also depends on the specific problem instance with respect to the market noise distribution F. Furthermore, Lemma 3 shows that $\|\hat{\beta}_t - \beta\|_1 \le \delta_t / x_{\max}$ with high probability, which indicates the estimation error in β decreases as more data points are collected. Finally, combining these results, we sum up the cumulative single period revenue losses to attain the cumulative expected regret for the NPAC-T policy.

6. Strategic Setting: NPAC-S Policy

The previous section studied the dynamic pricing problem in the truthful setting, and in this section, we consider the same pricing problem when buyers behave strategically. It is a well-known result that during a single-shot second price auction, bidding truthfully is a buyer's weakly dominant action. However, this salient feature no longer exists when the second price auctions are run repeatedly over time. This is because the repeated seller-buyer interactions provide buyers with the opportunity to take advantage of the seller's lack of knowledge in buyers' valuations by submitting corrupted bids instead of their true valuations: since the seller does not know buyers' demand curves and aims to learn them using submitted bids, buyers are incentivized to leverage their private information and "game the system" by submitting (corrupted) bids in an untruthful and strategic manner so that they can manipulate future reserve prices.

With the aforementioned considerations, the problem of interest is, again, to develop a seller strategy that extracts as much revenue as possible from all the buyers, and equivalently minimize cumulative regret against the "truthful" benchmark described in Definition 1. Recall that under this benchmark, the seller has full information of β and F and given this knowledge, sets the optimal contextual reserve price defined in Equation (5).

To maximize seller's expected revenue in the strategic setting, we propose a policy called NPAC-S that builds on the NPAC-T policy described in the previous section. NPAC-S differs from NPAC-T in two main aspects. First, NPAC-S is a phased algorithm, where each phase $\ell \geq 1$, denoted as E_ℓ, has length $T^{1-2^{-\ell}}$ which implies $|E_1| = \sqrt{T}$ and $|E_\ell|/\sqrt{|E_{\ell-1}|} = \sqrt{T}$. At the end of each phase, estimates for β, F^-, and F^+ are calculated in a similar fashion as Equations (8) and (9) in NPAC-T, but NPAC-S only uses data in the previous phase instead of all historical data. We will illustrate the significance of this phased structure later in this section. Here, we point out that because $|E_{\ell+1}| \geq |E_\ell|$ for all $\ell \geq 1$ and for some $\tilde{\ell} \geq \log_2(\log_2(T))$ we have $|E_{\tilde{\ell}}| = T^{1-2^{-\tilde{\ell}}} \geq T/2$, we know that the total number of phases can be upper bounded by $\lceil \log_2(\log_2(T)) \rceil + 1$. The second difference is that in NPAC-S, during each period in phase ℓ, with probability $1/|E_\ell|$, the seller "isolates" one of the N buyers uniformly at random, and sets reserve $r_t = r_t^u \sim \text{Uniform}(0, v_{\max})$, where $v_{\max}$ is the maximum possible buyer valuation. The seller discloses this random isolation protocol to all buyers before the sequence of repeated auctions begins. 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. Other buyers who are not isolated are not eligible to participate in the auction. Hence, one can think of a randomized isolation period in NPAC-S as running a second price auction with a single buyer. On the other hand, with probability $1 - 1/|E_\ell|$, the seller sets the reserve price r_t based on estimates of β, F^-, and F^+ calculated at the end of the previous phase $\ell - 1$ using the data from $E_{\ell-1}$. The detailed policy is shown in Algorithm 2. We point out that the seller's pricing policy is announced to all buyers so that buyers examine the policy and have the freedom to adopt any bidding strategy that would maximize their long term utility.

We now highlight some important features of NPAC-S policy and address the intuition behind such a design. Similar to NPAC-T policy, NPAC-S policy uses an OLS estimator and empirical distributions to estimate the mean vector and distributions F^- and F^+. One may question the validity of using OLS and empirical distributions under the strategic behaviour of buyers since both estimation techniques are extremely vulnerable to corrupted data (outliers). This, in turn, incentivizes buyers to corrupt their bids to manipulate seller's estimates and reserve prices. Because of such undesirable properties of OLS and empirical distributions, naively adopting these procedures in a learning algorithm can lead to a fragile policy that is largely subject to buyers manipulation. Yet, we will now illustrate how the additional features of dividing the time horizon into different phases and also leveraging randomized isolation in NPAC-S will make our learning policy robust to the strategic behaviour.

Due to the phased structure of the algorithm, our estimates for β and 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

Algorithm 2 NPAC-S Policy: Non-Parametric Contextual Policy against Strategic Buyers

Initialize: $\hat{\beta}_1 = 0$, and $\hat{F}_1^-(z) = \hat{F}_1^+(z) = 0$ for $\forall z \in \mathbb{R}$.

For phase $\ell \geq 1$,

- **Set reserve price:** For $t \in E_\ell$, observe $x_t \sim \mathcal{D}$:

  - Isolation: With probability $1/|E_\ell|$, choose one of the N buyers uniformly at random and offer him the item at price of

$$r_t^u \sim \text{Uniform}(0, v_{\max})\,. \tag{10}$$

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

$$\hat{r}_t = \arg \max_{y \in [0, v_{\max}]} \left\{ \int_0^y \hat{F}_\ell^-(z - \langle \hat{\beta}_\ell, x_t \rangle) dz - y\Big(\hat{F}_\ell^+(y - \langle \hat{\beta}_\ell, x_t \rangle)\Big) \right\}. \tag{11}$$

  - Observe all bids $\{b_{i,t}\}_{i \in [N]}$

- **Update the estimate of the mean vector:**

$$\hat{\beta}_{\ell+1} = \Big(\sum_{\tau \in E_\ell} x_\tau x_\tau^\top \Big)^\dagger \Big(\sum_{\tau \in E_\ell} x_\tau \bar{b}_\tau \Big) \quad \text{where} \quad \bar{b}_\tau = \frac{1}{N} \sum_{i \in [N]} b_{i,\tau}\,. \tag{12}$$

- **Update the estimate of F^+ and F^-:**

$$\hat{F}_{\ell+1}(z) = \frac{1}{N|E_\ell|} \sum_{\tau \in E_\ell} \sum_{i \in [N]} \mathbb{I}(b_{i,\tau} - \langle \hat{\beta}_{\ell+1}, x_\tau \rangle \leq z), \quad \text{then} \tag{13}$$

$$\hat{F}_{\ell+1}^-(z) = N \hat{F}_{\ell+1}^{N-1}(z) - (N-1)\hat{F}_{\ell+1}^N(z) \quad \text{and} \quad \hat{F}_{\ell+1}^+(z) = \hat{F}_{\ell+1}^N(z)\,. \tag{14}$$

potentially reduce buyers' manipulating power on our estimates and reserve prices. Furthermore, as all buyers are aware of the randomized isolation protocol, the presence of isolation periods restricts buyers from significantly corrupting their bids, and the reason is as follows. If no isolation occurs, a buyer may submit a bid that is far from her true valuation but face no consequences since her bid may not change the outcome nor payment of the auction. An example may be a buyer having the lowest valuation among all buyers, and submits a bid by adding a large corruption to her valuation, but still ending up not being the second highest or highest bidder. Assuming, for simplicity, 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. Since we use all bids for our estimations, such largely corrupted bids may hurt our estimates. In words, when no isolation occurs, buyers may be able to distort the seller's revenue without losing anything. However, during an isolation period when a buyer is chosen at random, she is isolated from the influence of other buyers, and corrupting her bid may perhaps yield a 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 buyers to reduce the extent of corruption in their bids. Finally, it is worth noting that as long as buyers are aware of seller's commitment to the randomized isolation protocol, it is not necessary for the seller to reveal, during an isolation period, which buyer is being isolated.

We now provide the key result for this section which provides an upper bound for the expected cumulative regret of the NPAC-S policy.

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

Here, we compare the cumulative expected regret of the NPAC-S policy with that of the NPAC-T policy in the truthful setting in Theorem 1, namely $\mathcal{O}\left(c_f\sqrt{dN^3T\log(T)}\right)$. We observe that when $0<\eta\ll 1$, the regret upper bound of the NPAC-S policy is approximately $\widetilde{\mathcal{O}}(\sqrt{dT})$, which matches the regret upper bound of the NPAC-T policy. For general η, we notice that compared to the regret of NPAC-T policy, the NPAC-S regret has extra factors which consist of three components, namely $\mathcal{O}\left(\frac{\sqrt{\log(T)}T^{1/4}}{\log(1/\eta)}\right)$, $\mathcal{O}\left(N^{3/2}\right)$, and $\mathcal{O}\left(\log(\log(T))\right)$. The extra factor $\mathcal{O}\left(\frac{\sqrt{\log(T)}T^{1/4}}{\log(1/\eta)}\right)$ in the regret bound of the NPAC-S policy 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 utilities. Thus, in the presence of randomized isolation periods, impatient buyers are more unlikely to significantly corrupt their bids, which translates into a lower regret. Furthermore, the additional factor of $\mathcal{O}\left(N^{3/2}\right)$ in the regret bound of the NPAC-S policy is due to two aspects: First, in the worst case all N buyers are strategic and corrupt there bids, resulting in a multiplicative factor of N; second, this extra N factor corresponds to additional error in our estimations for β due to the strategic behavior of buyers, which further propagates into the estimation errors in F, F^- and F^+. The additional $\mathcal{O}(\log(\log(T)))$ factor of the NPAC-S regret, compared to the NPAC-T regret, 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. Finally, we note that although the NPAC-S policy is developed to cope with corrupted data (bids) submitted by strategic buyers, it is also applicable to the truthful setting and hence can be viewed as a general policy under different buyer behaviours.

We now provide an outline of the proof for Theorem 2.

Proof sketch of Theorem 2 The regret of Theorem 2 can be decomposed into two parts: (i) the expected revenue loss due to the discrepancy between the posted reserve price r_t and the optimal reserve price $r_t^\star$ defined in Equation (5), which is caused by estimation errors in β, F^- and F^+; and (ii) the revenue loss due to buyers' strategic bidding behaviour.

First, we observe that the reserve price set by the seller depends on the estimates of β, F^- and F^+ in Equations (12) and (13) of Algorithm 2. These estimates, however, are biased by buyers' corruptions $\{a_{i,t}\}_{i\in[N],t\in[T]}$. Regarding corrupted bids, we distinguish between two cases of buyer behavior, namely a buyer submits a bid that is "slightly" corrupted, and a buyer submits a significantly corrupted bid. Intuitively, small corruptions will have less impact on the overall accuracy of estimates for β, F^- and F^+. Hence, the major focus should be placed on bids that are largely corrupted. Fortunately, utility-maximizing buyers do not have an incentive to significantly corrupt their bids for a large amount of periods, as by doing so, they may suffer from substantial utility loss, especially during isolation periods when they are chosen. Hence, one key aspect in proving Theorem 2 is to take advantage of this utility maximizing behaviour and show that the number of times a buyer significantly corrupts his bids is small. Specifically, we define

$$\mathcal{S}_{i,\ell} := \left\{ t \in E_\ell : |a_{i,t}| \geq \frac{1}{|E_\ell|} \right\} . \tag{15}$$

The set $\mathcal{S}_{i,\ell}$ for any $i \in [N]$ includes all periods in phase E_ℓ during which buyer i significantly corrupts his bids. In Lemma 5, we show that $|\mathcal{S}_{i,\ell}| = \mathcal{O}\left(\frac{\log(|E_\ell|)}{\log(1/\eta)}\right)$ with high probability. Now, based on this result, we are able to show that our estimate of β, F^- and F^+ are relatively accurate since the number of significantly corrupted bids is small, while the rest of the bids are only slightly corrupted and would not have a large impact on estimates. In Lemma 7, we show that

$$\|\hat{\beta}_{\ell+1} - \beta\|_1 = \mathcal{O}\left(\frac{1}{\sqrt{|E_\ell|}} + \frac{\log(|E_\ell|)}{\log(1/\eta)|E_\ell|}\right)$$

with high probability. On the other hand, in Lemma 8, we show that

$$\left|\hat{F}^-_{\ell+1}(z) - F^-(z)\right| = \mathcal{O}\left(\frac{N^2}{\sqrt{|E_\ell|}} + \frac{N^2\log(|E_\ell|)}{\log(1/\eta)|E_\ell|}\right) \text{ and } \left|\hat{F}^+_{\ell+1}(z) - F^+(z)\right| = \mathcal{O}\left(\frac{N}{\sqrt{|E_\ell|}} + \frac{N\log(|E_\ell|)}{\log(1/\eta)|E_\ell|}\right)$$

for any $z \in \mathbb{R}$ with high probability. Finally, Lemma 9 combines Lemmas 7 and 8 to bound the total impact of estimation errors on seller's revenue, which enables us to show that the revenue loss due to estimation errors is $\widetilde{\mathcal{O}}(\sqrt{|E_\ell|})$ in phase ℓ.

Next, we analyze the revenue loss due to buyers' strategic behavior. We note that there are several reasons for why bidding untruthfully may lead to a reduction in revenue. 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 light of this example, we refer to the situation where a buyer could have changed the allocation outcome had she bid untruthfully as an "allocation mismatch" for that buyer. Utilizing the fact that $|\mathcal{S}_{i,\ell}| = \mathcal{O}\left(\frac{\log(|E_\ell|)}{\log(1/\eta)}\right)$ with high probability, we show in Lemma 6 that the number of periods during which an allocation mismatch for any buyer occurs in phase E_ℓ is in the order of $\mathcal{O}\left(c_f + \frac{\log(|E_\ell|)}{\log(1/\eta)}\right)$. Moreover, we show that when a mismatch does not occur, the revenue loss due to buyers' strategic behavior can be bounded by $\sum_{i\in[N]}\sum_{t\in E_\ell}|a_{i,t}|$. Again, by employing the order of $|\mathcal{S}_{i,\ell}|$ for any $i \in [N]$, we have

$$\sum_{t\in E_\ell}\sum_{i\in[N]}|a_{i,t}| \le \sum_{t\in E_\ell/(\cup_{i\in[N]}\mathcal{S}_{i,\ell})}\sum_{i\in[N]}|a_{i,t}| + \mathcal{O}\left(\frac{N^2\log(|E_\ell|)}{\log(1/\eta)}\right) = \mathcal{O}\left(\frac{N^2\log(|E_\ell|)}{\log(1/\eta)}\right),$$

since $|a_{i,t}| \leq 1/|E_\ell|$ for all $t \in E_\ell/\mathcal{S}_{i,\ell}$. Eventually, this allows us to show the revenue loss in phase E_ℓ due to buyers' strategic behavior is in the order of $\mathcal{O}\left(\frac{N^2 \log(|E_\ell|)}{\log(1/\eta)}\right)$.

Finally, combining the two aforementioned constituents that cause a revenue loss, and utilizing the fact that there are $\mathcal{O}(\log\log(T))$ phases in total, we can show the cumulative regret over the entire horizon T as stated in Theorem 2.

7. Concluding Remarks

The problem of designing data-driven optimal auctions has drawn much attention lately; see, for example, Roughgarden and Wang (2016), Golrezaei et al. (2017), Derakhshan et al. (2019), and Javanmard and Nazerzadeh (2016). There, the goal is to deviate from classical Bayesian settings and take advantage of available data to optimize auction design either in an offline or online fashion. Although this approach is very natural, it can lead to designs that are vulnerable to buyers' manipulations as the data may be generated by strategic agents. In this work, we show how one can make data-driven auction design robust to buyers' manipulations even in a setting with a contextual and non-parametric demand model. To robustify our data-driven auction design, we (i) restart our policy over time in a systematic way and (ii) incorporate randomized isolation periods. While the former reduces the impact of corrupted (manipulative) bids, the latter makes manipulations costly for the buyers and incentivizes them to well-behave.

References

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

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

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

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

Bikhchandani S, McCardle K (2012) Behavior-based price discrimination by a patient seller. *The BE Journal of Theoretical Economics* 12(1).

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

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

Celis LE, Lewis G, Mobius M, 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, Mansour Y (2015) Regret minimization for reserve prices in second-price auctions. *IEEE Transactions on Information Theory* 61(1):549–564.

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

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

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

den Boer AV (2015) Dynamic pricing and learning: historical origins, current research, and new directions. *Surveys in operations research and management science* 20(1):1–18.

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

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

Derakhshan M, Golrezaei N, Leme RP (2019) Lp-based approximation for personalized reserve prices. *arXiv preprint arXiv:1905.01526* .

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.

Edelman B, Ostrovsky M (2007) Strategic bidder behavior in sponsored search auctions. *Decision support systems* 43(1):192–198.

Esteves RB, et al. (2009) A survey on the economics of behaviour-based price discrimination. Technical report, NIPE-Universidade do Minho.

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

Fudenberg D, Villas-Boas JM (2006) Behavior-based price discrimination and customer recognition. *Handbook on economics and information systems* 1:377–436.

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

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

Golrezaei N, Lobel I, Paes Leme R (2018b) Auction design for roi-constrained buyers. *Available at SSRN 3124929* .

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

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

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

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

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

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

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

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

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

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

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

Roughgarden T, Schrijvers O (2016) Ironing in the dark. *Proceedings of the 2016 ACM Conference on Economics and Computation*, 1–18 (ACM).

Roughgarden T, Wang JR (2016) Minimizing regret with multiple reserves. *Proceedings of the 2016 ACM Conference on Economics and Computation*, 601–616 (ACM).

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

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

8. Appendix for Section 4: 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 (3) and (4), the seller's expected revenue conditioned on x_t by setting reserve price r_t is:

$$\begin{aligned}\mathsf{rev}_t(r_t) &= \mathbb{E}\left[\max\{r_t, v_t^-\}\mathbb{I}\{v_t^+ \geq r_t\} \mid x_t, r_t\right] \\ &= \mathbb{E}\left[r_t\mathbb{I}\{v_t^+ \geq r_t \geq v_t^-\} + v_t^-\mathbb{I}\{v_t^+ \geq v_t^- \geq r_t\} \mid x_t, r_t\right], \end{aligned} \tag{16}$$

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

$$\mathbb{E}\left[r_t\mathbb{I}\{v_t^+ \geq r_t \geq v_t^-\} \mid x_t, r_t\right] = r_t N \left[Q_t(r_t)\right]^{N-1}\left[1 - Q_t(r_t)\right], \tag{17}$$

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 (16) is

$$\begin{aligned}\mathbb{E}\left[v_t^-\mathbb{I}\{v_t^+ \geq v_t^- \geq r_t\} \mid x_t, r_t\right] &= \mathbb{E}\left[v_t^-\mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] \\ &= \mathbb{E}\left[(v_t^- - r_t)\mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] + r_t\mathbb{E}\left[\mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] \\ &= \int_0^\infty \mathbb{P}\left(v_t^- - r_t \geq z\right) dz + r_t\left[1 - Q_t^-(r_t)\right] \\ &= \int_{r_t}^\infty \left[1 - Q_t^-(z)\right] dz + r_t\left[1 - Q_t^-(r_t)\right] \\ &= \mathbb{E}\left[v_t^- \mid x_t, r_t\right] - \int_0^{r_t}\left[1 - Q_t^-(z)\right] dz + r_t\left[1 - Q_t^-(r_t)\right] \\ &= \mathbb{E}\left[v_t^- \mid x_t\right] + \int_0^{r_t} Q_t^-(z)dz - r_t Q_t^-(r_t). \end{aligned} \tag{18}$$

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

$$Q_t^-(r_t) = N\left[Q_t(r_t)\right]^{N-1}\left[1 - Q_t(r_t)\right] + \left[Q_t(r_t)\right]^N. \tag{19}$$

Hence, combining Equations (16), (17), (18), and (19), we have

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

9. Appendix for Section 5: Proof of Theorem 1

To show that the T-period regret of Algorithm 1 is upper bounded as $\mathsf{Regret}(T) = \mathcal{O}\left(\sqrt{dN^3T\log(T)}\right)$, we will makes use of the following events throughout our proof:

$$\xi_t := \left\{\|\hat{\beta}_t - \beta\|_1 \le \frac{\delta_t}{x_{\max}}\right\} \quad \text{where} \quad \delta_t = \frac{4\sqrt{d\log(t-1)}\epsilon_{\max}x_{\max}^2}{\lambda_0^2\sqrt{N(t-1)}} \tag{20}$$

$$\xi_t^- := \left\{\left|F^-(z) - \hat{F}_t^-(z)\right| \le \gamma_t + 2c_f N^2\delta_t, \quad \forall z\in\mathbb{R}\right\} \tag{21}$$

$$\xi_t^+ := \left\{\left|F^+(z) - \hat{F}_t^+(z)\right| \le \gamma_t + c_f N\delta_t, \quad \forall z\in\mathbb{R}\right\}, \tag{22}$$

where $\gamma_t := \sqrt{2\log(t)}/\sqrt{t}$, λ_0^2 is the minimum eigenvalue of covariance matrix Σ, and $c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z)$. Under event ξ_t, the estimation error in the mean vector is small, and under event ξ_t^- and ξ_t^+, the estimation errors in F^- and F^+ are small respectively. Furthermore, we note that under these events, our estimates for β, F^-, and F^+ become more accurate as we obtain more data points over time. We now define a threshold time period T_0, starting from which our estimates are "sufficiently accurate":

$$\mathsf{T}_0 = \left\lceil \max\left\{\sqrt{dT}, \frac{16x_{\max}^2\log(T)}{\lambda_0^2}, 2\right\}\right\rceil + 1 = \mathcal{O}(\sqrt{dT}). \tag{23}$$

We will later discuss the significance of this construction for T_0.

For simplicity, we let $y_t := \langle\beta, x_t\rangle$, and $\hat{y}_t := \langle\hat{\beta}_t, x_t\rangle$. Then, assuming that buyers are always truthful, the regret in period t is given by

$$\begin{aligned}\mathsf{Regret}_t &= \mathbb{E}\left[\mathrm{REV}_t^\star - \mathsf{rev}_t(r_t)\right] \\ &= \mathbb{E}\left[\int_0^{r_t^\star} F^-(z-y_t)dz - r_t^\star\left[F^+(r_t^\star-y_t)\right] - \int_0^{r_t} F^-(z-y_t)dz + r_t\left[F^+(r_t-y_t)\right]\right],\end{aligned} \tag{24}$$

where the expectation is taken with respect to $x_t \sim \mathcal{D}$ and r_t; the second equality follows from Proposition 1. Define

$$\mathcal{R}_t := \int_0^{r_t^\star} F^-(z-y_t)dz - r_t^\star\left[F^+(r_t^\star - y_t)\right] - \int_0^{r_t} F^-(z-y_t)dz + r_t\left[F^+(r_t - y_t)\right] \quad \text{and} \tag{25}$$

$$\rho_t(r, y, F^{(1)}, F^{(2)}) := \int_0^r F^{(2)}(z-y)dz - r\left[F^{(1)}(r-y)\right]\,. \tag{26}$$

Then, we have

$$\begin{aligned}
\mathcal{R}_t \;&=\; \rho_t(r_t^\star, y_t, F^-, F^+) - \rho_t(r_t, y_t, F^-, F^+) \\
&=\; \rho_t(r_t^\star, y_t, F^-, F^+) - \rho_t(r_t^\star, \hat{y}_t, F^-, F^+) \\
&\quad + \rho_t(r_t^\star, \hat{y}_t, F^-, F^+) - \rho_t(r_t^\star, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) \\
&\quad + \rho_t(r_t^\star, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) - \rho_t(r_t, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) \\
&\quad + \rho_t(r_t, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) - \rho_t(r_t, \hat{y}_t, F^-, F^+) \\
&\quad + \rho_t(r_t, \hat{y}_t, F^-, F^+) - \rho_t(r_t, y_t, F^-, F^+)\,.
\end{aligned} \tag{27}$$

We note that the second equation follows from adding and subtracting terms. Observe the first and the last terms of the second equation capture the impact of estimation error in the mean vector β. Further, the second and forth terms capture the impact of the estimation errors in distributions F^- and F^+ while the third term captures the errors in reserve price with respect to all estimations. We now invoke Lemma 4, where we show that when events ξ_t, ξ_t^-, and ξ_t^+ happen, for $r \in \{r_t^\star, r_t\}$ we have

(i) $|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \hat{y}_t, F^-, F^+)| \;\le\; 3rc_f N^2\delta_t$ a.s.

(ii) $\left|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+)\right| \;\le\; r(3c_f N^2\delta_t + 2\gamma_t)$ a.s.

This result relies on the Lipschitz properties of F^- and F^+ shown in Lemma 1, and allows us to bound the regret using estimation errors without imposing the MHR assumption. 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 (27), we get

$$\mathcal{R}_t \cdot \mathbb{I}\{\xi_t \cap \xi_t^- \cap \xi_t^+\}$$

$$\leq 3(r_t^\star + r_t)c_f N^2\delta_t + \left(\rho_t(r_t^\star, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) - \rho_t(r_t, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+)\right) + (r_t^\star + r_t)\cdot(3c_f N^2\delta_t + 2\gamma_t)$$

$$= 2(r_t^\star + r_t)\cdot(3c_f N^2\delta_t + \gamma_t) + \left(\rho_t(r_t^\star, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) - \rho_t(r_t, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+)\right). \tag{28}$$

We recall that the seller's pricing decision r_t in period t is defined in Equation (7), and observe that $r_t = \arg\max_{r\in[0,v_{\max}]} \rho_t(r, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+)$. Since $r_t^\star \in (0, v_{\max})$, we obtain the fact that $\rho_t(r_t^\star, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) - \rho_t(r_t, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+) \leq 0$. Plugging this into Equation (28), we get

$$\mathcal{R}_t \cdot \mathbb{I}\{\xi_t \cap \xi_t^- \cap \xi_t^+\} \leq 2(r_t^\star + r_t)\cdot(3c_f N^2\delta_t + \gamma_t) \leq 4v_{\max}(3c_f N^2\delta_t + \gamma_t). \tag{29}$$

So far, we have bounded the single period regret for some period $t \in [T]$ assuming that events ξ_t, ξ_t^-, ξ_t^+ all hold. But, before we sum this regret up across all periods to get the cumulative regret, we first turn to upper bound the probability that not all of the events $\{\xi_t\}_{t\geq \mathsf{T}_0}, \{\xi_t^-\}_{t\geq \mathsf{T}_0}, \{\xi_t^+\}_{t\geq \mathsf{T}_0}$ occur, where $\mathsf{T}_0 = \mathcal{O}(\sqrt{dT})$ is defined in Equation (23). We denote the complement of a set $\mathcal{A}$ to be $\mathcal{A}^c$. Hence, the probability that not all of the events $\{\xi_t\}_{t\geq \mathsf{T}_0}, \{\xi_t^-\}_{t\geq \mathsf{T}_0}, \{\xi_t^+\}_{t\geq \mathsf{T}_0}$ occur is

$$\mathbb{P}\left(\bigcup_{t=\mathsf{T}_0}^{T} \left(\{\xi_t\}^c \cup \{\xi_t^-\}^c \cup \{\xi_t^+\}^c\right)\right) \leq \sum_{t=\mathsf{T}_0}^{T} \left[\mathbb{P}\left(\{\xi_t\}^c\right) + \mathbb{P}\left(\{\xi_t^-\}^c\right) + \mathbb{P}\left(\{\xi_t^+\}^c\right)\right].$$

According to Lemma 3, the probability $\mathbb{P}(\{\xi_t\}^c)$ is bounded by

$$\mathbb{P}(\{\xi_t\}^c) \leq \frac{2d}{(t-1)^2} + d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right) \leq \frac{2d}{(t-1)^2} + \frac{d}{T^2}, \tag{30}$$

where the second inequality is due to our construction of T_0, such that for $t \geq \mathsf{T}_0$ we have $t-1 \geq 16x_{\max}^2 \log(T)/\lambda_0^2$. On the other hand, by taking $\gamma = \gamma_t = \sqrt{2\log(t)}/\sqrt{t}$ in Lemma 2, the probability $\mathbb{P}(\{\xi_t^-\}^c)$ is bounded as:

$$\begin{aligned}
\mathbb{P}\left(\{\xi_t^-\}^c\right) &\leq 4\exp\left(-t\gamma^2\right) + \frac{4d}{(t-1)^2} + 2d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right)\\
&\leq 4\exp\left(-t\cdot\left(\frac{\sqrt{2\log(t)}}{\sqrt{t}}\right)^2\right) + \frac{4d}{(t-1)^2} + \frac{2d}{T^2}\\
&\leq \frac{4}{t^2} + \frac{4d}{(t-1)^2} + \frac{2d}{T^2}.
\end{aligned}$$

Similarly, we have $\mathbb{P}\left(\left\{\xi_t^+\right\}^c\right) \leq \frac{4}{t^2}+\frac{4d}{(t-1)^2}+\frac{2d}{T^2}$. Therefore,

$$\mathbb{P}\left(\bigcup_{t=\mathsf{T}_0}^{\top}\left(\{\xi_t\}^c\cup\left\{\xi_t^-\right\}^c\cup\left\{\xi_t^+\right\}^c\right)\right) \leq \sum_{t=\mathsf{T}_0}^{\top}\left(\frac{5d}{T^2}+\frac{10d}{(t-1)^2}+\frac{8}{t^2}\right)$$
$$\leq \int_{\sqrt{dT}}^{\infty}\frac{15d+8}{\tau^2}d\tau \leq \frac{23d}{\sqrt{dT}}=\frac{23\sqrt{d}}{\sqrt{T}}. \tag{31}$$

The second inequality is due to $\mathsf{T}_0 \geq \sqrt{dT}$ according to its definition, and the subsequent inequalities use the fact that $d \geq 1$.

Finally, we break down the cumulative expected regret into three parts: the trivial bound $v_{\max}$ for the regret from period $t=1$ to T_0-1, the cumulative expected regret from period $t=\mathsf{T}_0$ to T given the occurrence of events $\{\xi_t\}_{t\geq\mathsf{T}_0}, \{\xi_t^-\}_{t\geq\mathsf{T}_0}, \{\xi_t^+\}_{t\geq\mathsf{T}_0}$, and the trivial bound from period $t=\mathsf{T}_0$ to T if not all these events occur. Hence, the final cumulative regret is

$$\begin{aligned}
&\mathsf{Regret}(T)\\
\leq\ & v_{\max}(\mathsf{T}_0-1)+\sum_{t=\mathsf{T}_0}^{T}\mathbb{E}\left[\mathcal{R}_t\right]\\
=\ & v_{\max}(\mathsf{T}_0-1)+\sum_{t=\mathsf{T}_0}^{T}\mathbb{E}\left[\mathcal{R}_t\cdot\mathbb{I}\{\xi_t\cap\xi_t^-\cap\xi_t^+\}\right]+\sum_{t=\mathsf{T}_0}^{T}\mathbb{E}\left[\mathcal{R}_t\cdot\left(\mathbb{I}\left\{(\xi_t)^c\cup\left(\xi_t^-\right)^c\cup\left(\xi_t^+\right)^c\right\}\right)\right]\\
\leq\ & v_{\max}(\mathsf{T}_0-1)+\sum_{t=\mathsf{T}_0}^{T}\mathbb{E}\left[\mathcal{R}_t\cdot\left(\mathbb{I}\{\xi_t\cap\xi_t^-\cap\xi_t^+\}\right)\right]+Tv_{\max}\mathbb{P}\left(\bigcup_{t=\mathsf{T}_0}^{T}\left((\xi_t)^c\cup\left(\xi_t^-\right)^c\cup\left(\xi_t^+\right)^c\right)\right)\\
\leq\ & v_{\max}\mathsf{T}_0+\sum_{t=\mathsf{T}_0}^{T}4v_{\max}\left(3c_fN^2\delta_t+\gamma_t\right)+Tv_{\max}\cdot\mathbb{P}\left(\bigcup_{t=\mathsf{T}_0}^{T}\left((\xi_t)^c\cup\left(\xi_t^-\right)^c\cup\left(\xi_t^+\right)^c\right)\right)\\
\leq\ & v_{\max}\mathsf{T}_0+4v_{\max}\sum_{t=\mathsf{T}_0}^{T}\left(3c_fN^2\delta_t+\gamma_t\right)+Tv_{\max}\cdot\frac{23\sqrt{d}}{\sqrt{T}}
\end{aligned}$$

The first inequality applies Equation (24) and uses the trivial bound $v_{\max}$ for each of the first T_0 periods; the second inequality is because $\mathcal{R}_t \leq v_{\max}$; the third inequality follows from Equations (29), and the final inequality follows from (31). Then, by plugging in the definition of δ_t and γ_t,

$$\begin{aligned}
&\mathsf{Regret}(T)\\
\leq\ & v_{\max}\mathsf{T}_0+4v_{\max}\int_0^T\left(3c_fN^2\cdot\frac{4\sqrt{d\log(\tau)}\epsilon_{\max}x_{\max}^2}{\lambda_0^2\sqrt{N\tau}}+\frac{\sqrt{2\log(\tau)}}{\sqrt{\tau}}\right)d\tau+23v_{\max}\sqrt{dT}\\
\leq\ & v_{\max}\mathsf{T}_0+4v_{\max}\left(\frac{24c_f\epsilon_{\max}x_{\max}^2\sqrt{dN^3T\log(T)}}{\lambda_0^2}+2\sqrt{2T\log(T)}\right)+23v_{\max}\sqrt{dT}\\
=\ & \mathcal{O}\left(c_f\sqrt{dN^3T\log(T)}\right),
\end{aligned}$$

where the first inequality follows from the fact that a summation of monotonically decreasing functions can be upper bounded by its integral.

9.1. Lemmas for proving Theorem 1

LEMMA 1 **(Lipschitz Property for F, F^- and F^+).** *The following hold for any $z_1, z_2 \in \mathbb{R}$:*

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

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

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

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

Proof of Lemma 1. 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:

Case 1: $(z_1 < z_2 \le -\epsilon_{\max} \;\text{ or }\; \epsilon_{\max} \le z_1 < z_2)$: $|F(z_2) - F(z_1)| = 0 \le c_f |z_2 - z_1|$.

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 $\tilde{z} \in (z_1, z_2)$.

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

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

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

$$
\begin{aligned}
& |F^-(z_2) - F^-(z_1)| \\
= \; & \left| N F^{N-1}(z_2) - (N-1) F^N(z_2) - \left(N F^{N-1}(z_1) - (N-1) F^N(z_1) \right) \right| \\
\le \; & N \left| F^{N-1}(z_2) - F^{N-1}(z_1) \right| + (N-1) \left| F^N(z_2) - F^N(z_1) \right|
\end{aligned}
$$

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

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

LEMMA 2 **(Bounding Estimation Errors in F^- and F^+)**. *Define σ_t to be the sigma algebra generated by all $\{x_\tau\}_{\tau\in[t]}$ and $\{\epsilon_{i,\tau}\}_{i\in[N],\tau\in[t]}$. Then, for any σ_t-measurable random variable z and any $\gamma > 0$, we have*

$$
\mathbb{P}\left(\left|F^-(z) - \hat{F}_t^-(z)\right| \leq \gamma + 2c_f N^2 \delta_t\right) \;\geq\; 1 - 4\exp\left(-t\gamma^2\right) - \frac{4d}{(t-1)^2} - 2d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right) \quad \textit{and}
$$

$$
\mathbb{P}\left(\left|F^+(z) - \hat{F}_t^+(z)\right| \leq \gamma + c_f N \delta_t\right) \;\geq\; 1 - 4\exp\left(-t\gamma^2\right) - \frac{4d}{(t-1)^2} - 2d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right),
$$

where $\delta_t = \frac{4\sqrt{d\log(t-1)}\epsilon_{\max}x_{\max}^2}{\lambda_0^2\sqrt{N(t-1)}}$ and $c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z)$.

Proof of Lemma 2. Our goal here is to show that $\left|F^-(z) - \hat{F}_t^-(z)\right|$ is small with high probability. According to our estimate of F^- in NPAC-T policy,

$$
\begin{aligned}
\hat{F}_t^-(z) &= \frac{1}{t-1}\sum_{\tau\in[t-1]} \mathbb{I}\left\{v_\tau^- - \langle\hat{\beta}_t, x_\tau\rangle \leq z\right\} \\
&= \frac{1}{t-1}\sum_{\tau\in[t-1]} \mathbb{I}\left\{\epsilon_\tau^- \leq z + \langle\hat{\beta}_t - \beta, x_\tau\rangle\right\}.
\end{aligned}
$$

We highlight that $\mathbb{E}\left[\hat{F}_t^-(z)\right] \neq \frac{1}{t-1}\sum_{\tau\in[t-1]} F^-\left(z + \langle\hat{\beta}_t - \beta, x_\tau\rangle\right)$ because both z and $\hat{\beta}_t$ are σ_t-measurable. Hence, one cannot naively apply concentration inequalities to bound $\left|F^-(z) - \hat{F}_t^-(z)\right|$.

Therefore, our approach is to first construct upper and lower bounds for $\hat{F}_t^-(z)$ that are not functions of $\hat{\beta}_t$, and then apply concentration inequalities on these upper and lower bounds respectively.

Recall the definition of the event $\xi_t = \left\{\|\hat{\beta}_t - \beta\|_1 \leq \delta_t / x_{\max}\right\}$. Under this event, we have

$$\frac{1}{t-1}\sum_{\tau\in[t-1]} \mathbb{I}\left\{\epsilon_\tau^- \leq z - \delta_t\right\} \ \leq\ \hat{F}_t^-(z) \ \leq\ \frac{1}{t-1}\sum_{\tau\in[t-1]} \mathbb{I}\left\{\epsilon_\tau^- \leq z + \delta_t\right\} . \tag{32}$$

Now, for any $\gamma > 0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{F}_t^-(z) - F^-(z+\delta_t) \leq \gamma\right)\\
\geq\ &\mathbb{P}\left(\left\{\hat{F}_t^-(z) - F^-(z+\delta_t) \leq \gamma\right\} \bigcap \xi_t\right)\\
\geq\ &\mathbb{P}\left(\left\{\frac{\sum_{\tau\in[t-1]} \mathbb{I}\left\{\epsilon_\tau^- \leq z+\delta_t\right\}}{t-1} - F^-(z+\delta_t) \leq \gamma\right\} \bigcap \xi_t\right)\\
\geq\ &\mathbb{P}\left(\left\{\sup_{\tilde{z}\in\mathbb{R}}\left|\frac{1}{t-1}\sum_{\tau\in[t-1]} \mathbb{I}\left\{\epsilon_\tau^- \leq \tilde{z}\right\} - F^-(\tilde{z})\right| \leq \gamma\right\} \bigcap \xi_t\right)\\
\geq\ &1 - \mathbb{P}\left(\sup_{\tilde{z}\in\mathbb{R}}\left|\frac{1}{t-1}\sum_{\tau\in[t-1]} \mathbb{I}\left\{\epsilon_\tau^- \leq \tilde{z}\right\} - F^-(\tilde{z})\right| > \gamma\right) - \mathbb{P}\left(\xi_t^c\right)\\
\geq\ &1 - 2\exp\left(-2(t-1)\gamma^2\right) - \frac{2d}{(t-1)^2} - d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right)\\
\geq\ &1 - 2\exp\left(-t\gamma^2\right) - \frac{2d}{(t-1)^2} - d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right),
\end{aligned}$$

where the second inequality follows from (32), the fourth inequality follows from the union bound, and the second last inequality follows from the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality (Theorem 11) and Lemma 3. We note that we can apply the DKW inequality because $\{\epsilon_\tau^-\}_{\tau\in[t-1]}$ are $t-1$ i.i.d. realizations of the $(N-1)^{th}$ order statistic of N i.i.d. noise variables. The last inequality holds because $t \geq 2$ and as a result, $t-1 \geq \frac{t}{2}$. Furthermore, invoking the Lipschitz properties of F^- shown in Lemma 1, we have $|F^-(z+\delta_t) - F^-(z)| \leq 2c_f N^2 \delta_t$. Therefore,

$$\begin{aligned}
&\mathbb{P}\left(\hat{F}_t^-(z) - F^-(z) \ \leq\ \gamma + 2c_f N^2 \delta_t\right)\\
\geq\ &\mathbb{P}\left(\hat{F}_t^-(z) - F^-(z+\delta_t) \leq \gamma\right)\\
\geq\ &1 - 2\exp\left(-t\gamma^2\right) - \frac{2d}{(t-1)^2} - d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right) .
\end{aligned} \tag{33}$$

Similarly, we have $|F^-(z)-F^-(z-\delta_t)|\le 2c_fN^2\delta_t$ by Lemma 1. Hence, following the same arguments as in the case of $\mathbb{P}\left(\hat{F}_t^-(z)-F^-(z)\ \le\ \gamma+2c_fN^2\delta_t\right)$, we have

$$\begin{aligned}\mathbb{P}\left(F^-(z)-\hat{F}_t^-(z)\le\gamma+2c_fN^2\delta_t\right)\ &\ge\ \mathbb{P}\left(F^-(z-\delta_t)-\hat{F}_t^-(z)\le\gamma\right)\\ &\ge\ 1-2\exp\left(-t\gamma^2\right)-\frac{2d}{(t-1)^2}-d\exp\left(-\frac{(t-1)\lambda_0^2}{8x_{\max}^2}\right). \qquad (34)\end{aligned}$$

Finally, applying a union bound on Equations (33) and (34) will yield the result in the statement of the lemma. We can show a similar result for $F^+(z)-\hat{F}_t^+(z)$ by following the same reasoning. □

Lemma 3 **(Bounding Estimation Errors in β).** *For any $\gamma>0$,*

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

where λ_0^2 is the minimum eigenvalue of covariance matrix Σ and the estimate $\hat{\beta}_{t+1}$ is defined in Equation (8). Furthermore, setting $\gamma=\frac{4\sqrt{d\log(t)}\epsilon_{\max}x_{\max}}{\lambda_0^2\sqrt{Nt}}$ and denoting $\delta_{t+1}=\gamma x_{\max}$, we have

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

Proof of Lemma 3. The proof of Lemma 3 is inspired by Lemma EC.7.2 in Bastani and Bayati (2015). 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 $\mathbb{E}[x_tx_t^\top]=\Sigma+\mathbb{E}[x]\mathbb{E}[x]^\top$, we know that the smallest eigenvalue of $\mathbb{E}[x_tx_t^\top]$ is at least $\lambda_0^2>0$. We denote the design matrix of all the features up to time t as X where $X\in\mathbb{R}^{t\times d}$, and $\bar{\epsilon}_\tau=\frac{\sum_{i\in[N]}\epsilon_{i,\tau}}{N}$ for $\forall\tau\in[t]$.

We first consider the case where the smallest eigenvalue of the second moment matrix $\lambda_{\min}(X^\top X/t)\ge\lambda_0^2/2$, which implies that $(X^\top X)^{-1}$ exists and $(X^\top X)^{-1}=(X^\top X)^\dagger$. Later, we show that with high probability, $\lambda_{\min}(X^\top X/t)\ge\lambda_0^2/2$. By the definition of $\hat{\beta}_{t+1}$ and $\bar{v}_t$ in Equations (8),

$$\hat{\beta}_{t+1}=\left(X^\top X\right)^{-1}X^\top\begin{pmatrix}\bar{v}_1\\ \vdots\\ \bar{v}_t\end{pmatrix}=\left(X^\top X\right)^{-1}X^\top\begin{pmatrix}\frac{\sum_{i\in[N]}v_{i,1}}{N}\\ \vdots\\ \frac{\sum_{i\in[N]}v_{i,t}}{N}\end{pmatrix}$$

$$
= \beta + \left(X^\top X\right)^{-1} X^\top \begin{pmatrix} \frac{\sum_{i\in[N]} \epsilon_{i,1}}{N} \\ \vdots \\ \frac{\sum_{i\in[N]} \epsilon_{i,t}}{N} \end{pmatrix}
$$

$$
= \beta + \left(X^\top X\right)^{-1} X^\top \bar{\mathcal{E}},
$$

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

$$
\|\hat{\beta}_{t+1} - \beta\|_2 = \|\left(X^\top X\right)^{-1} X^\top \bar{\mathcal{E}}\|_2 \le \frac{2}{t\lambda_0^2} \cdot \|X^\top \bar{\mathcal{E}}\|_2, \tag{35}
$$

since we assumed $\lambda_{\min}(X^\top X/t) \ge \lambda_0^2/2$. Denote X^j as the jth column of X, i.e., the jth row of $X^\top$, for $j = 1, 2 \dots d$. Since $\|X^\top \bar{\mathcal{E}}\|_2^2 = \sum_{j\in[d]} \left|\bar{\mathcal{E}}^\top X^j\right|^2$, for any $\gamma > 0$ we have

$$
\bigcap_{j\in[d]} \left\{ \left|\bar{\mathcal{E}}^\top X^j\right| \ \le\ \frac{t\lambda_0^2\gamma}{2\sqrt{d}} \right\} \subseteq \left\{ \frac{2}{t\lambda_0^2} \cdot \|X^\top \bar{\mathcal{E}}\|_2 \le \gamma \right\}. \tag{36}
$$

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 [10] random variables. Therefore by Hoeffding's inequality, for any $\tilde{\gamma} > 0$

$$
\mathbb{P}\left(\left|N\bar{\mathcal{E}}^\top X^j\right| \le \tilde{\gamma}\right) \ge 1 - 2\exp\left(-\frac{\tilde{\gamma}^2}{2\epsilon_{\max}^2 x_{max}^2 tN}\right). \tag{37}
$$

Hence,

$$
\begin{aligned}
\mathbb{P}\left(\frac{2}{t\lambda_0^2} \cdot \|X^\top \bar{\mathcal{E}}\|_2 \le \gamma\right) \ &\ge\ \mathbb{P}\left(\bigcap_{j\in[d]} \left\{ \left|\bar{\mathcal{E}}^\top X^j\right| \le \frac{t\lambda_0^2\gamma}{2\sqrt{d}} \right\}\right) \\
&\ge\ 1 - \sum_{j\in[d]} \mathbb{P}\left(\left|\bar{\mathcal{E}}^\top X^j\right| > \frac{t\lambda_0^2\gamma}{2\sqrt{d}}\right) \\
&\ge\ 1 - 2d\exp\left(-\frac{N\gamma^2\lambda_0^4 t}{8\epsilon_{\max}^2 x_{\max}^2 d}\right),
\end{aligned} \tag{38}
$$

where the first inequality follows from Equation (36), the second inequality applies the union bound, and the last inequality follows from Equation (37) by replacing $\tilde{\gamma}$ with $Nt\lambda_0^2\gamma/(2\sqrt{d})$.

[10] A random variable Z is σ-subgaussian if for $\forall \gamma \in \mathbb{R}$, $\mathbb{E}[\exp(\gamma Z)] \le \exp(\gamma^2\sigma^2/2)$.

Now it only remains to show $\lambda_{\min}(X^\top X/t) \geq \lambda_0^2/2$ with high probability, which can be achieved by applying Lemma 12. In the context of this lemma, we consider the sequence of random matrices $\{x_\tau x_\tau^\top/t\}_{\tau\in[t]}$, and note that $X^\top X/t = \sum_{\tau\in[t]}(x_\tau x_\tau^\top/t)$. We first upper bound the maximum eigenvalue of $x_\tau x_\tau^\top/t$, namely $\lambda_{\max}(x_\tau x_\tau^\top/t)$ for any $\tau \in [t]$ by

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

This allows us to apply Lemma 12 (setting $\bar{\gamma} = 1/2$ in the lemma) and get

$$\begin{aligned} \mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{t}\right) \geq \frac{\lambda_0^2}{2}\right) &\geq \mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{t}\right) \geq \frac{\lambda_{\min}\left(\mathbb{E}[X^\top X/t]\right)^2}{2}\right) \\ &\geq 1 - d\exp\left(-\frac{t\lambda_0^2}{8x_{max}^2}\right), \end{aligned} \tag{39}$$

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

Therefore,

$$\begin{aligned} \mathbb{P}\left(\|\hat{\beta}_{t+1} - \beta\|_1 \leq \gamma\right) &\geq \mathbb{P}\left(\|\hat{\beta}_{t+1} - \beta\|_2 \leq \gamma\right) \\ &\geq \mathbb{P}\left(\left\{\frac{2}{t\lambda_0^2}\cdot\|X^\top\bar{\mathcal{E}}\|_2 \leq \gamma\right\}\bigcap\left\{\lambda_{\min}\left(\frac{X^\top X}{t}\right) \geq \frac{\lambda_0^2}{2}\right\}\right) \\ &\geq 1 - \mathbb{P}\left(\frac{2}{t\lambda_0^2}\cdot\|X^\top\bar{\mathcal{E}}\|_2 > \gamma\right) - \mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{t}\right) < \frac{\lambda_0^2}{2}\right) \\ &\geq 1 - 2d\exp\left(-\frac{N\gamma^2\lambda_0^4 t}{8\epsilon_{\max}^2 x_{max}^2 d}\right) - d\exp\left(-\frac{t\lambda_0^2}{8x_{max}^2}\right). \end{aligned}$$

The first inequality follows from the fact that $\|z\|_1 \leq \|z\|_2$ for any vector z; the second inequality follow from Equation (35); the fourth inequality applies a simple union bound; and the final inequality follows from Equations (38) and (39). □

The following Lemma utilizes the Lipschitz properties of F^- and F^+ shown in Lemma 2 to bound the seller's regret in terms of the estimation errors of both β, F^- and F^+.

LEMMA 4 **(Bounding the Impact of Estimation Errors on Revenue)**. *Assume that the events* $\xi_t = \left\{\|\hat{\beta}_t - \beta\|_1 \leq \delta_t/x_{max}\right\}$, $\xi_t^- = \left\{\left|F^-(z) - \hat{F}_t^-(z)\right| \leq \gamma_t + 2c_f N^2\delta_t \text{ for } \forall z\in\mathbb{R}\right\}$ *and* $\xi_t^+ = \left\{\left|F^+(z) - \hat{F}_t^+(z)\right| \leq \gamma_t + c_f N\delta_t \text{ for } \forall z\in\mathbb{R}\right\}$ *occur with* $\gamma_t = \sqrt{2\log(t)}/\sqrt{t}$, *and* $\delta_t = \frac{4\sqrt{d\log(t-1)}\epsilon_{\max}x_{\max}^2}{\lambda_0^2\sqrt{N(t-1)}}$. *Then, for* $r \in \{r_t^\star, r_t\}$ *we have the following:*

(i) $|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \hat{y}_t, F^-, F^+)| \le 3rc_f N^2 \delta_t$ *a.s.*

(ii) $\left|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+)\right| \le r(3c_f N^2 \delta_t + 2\gamma_t)$ *a.s.*

where $y_t = \langle \beta, x_t \rangle$, $\hat{y}_t = \langle \hat{\beta}_t, x_t \rangle$, $\hat{\beta}_t, \hat{F}_t^-, \hat{F}_t^+$ *are defined in Equations (8) and (9). The function* ρ_t *is defined in Equation (26).*

Proof of Lemma 4. ***Part (i)*** *We consider the following:*

$$
\begin{aligned}
& \left|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \hat{y}_t, F^-, F^+)\right| \\
= & \left|\int_0^r \left[F^-(z - y_t) - F^-(z - \hat{y}_t)\right] dz - r\left[F^+(r - y_t) - F^+(r - \hat{y}_t)\right]\right| \\
\le & \int_0^r \left|F^-(z - y_t) - F^-(z - \hat{y}_t)\right| dz + r\left|F^+(r - y_t) - F^+(r - \hat{y}_t)\right| \\
\le & \int_0^r 2c_f N^2 |y_t - \hat{y}_t| dz + rc_f N |y_t - \hat{y}_t| \\
\le & \int_0^r 2c_f N^2 \left(\|\hat{\beta}_t - \beta\|_1 x_{\max}\right) dz + rc_f N \|\hat{\beta}_t - \beta\|_1 x_{\max} \\
< & \ 3rc_f N^2 \delta_t \,.
\end{aligned}
$$

The first equality follows from definition of ρ_t *in Equation 26; the second inequality follows from Lemma 1, the third inequality follows from Cauchy's inequality:* $|y_t - \hat{y}_t| = |\langle \hat{\beta}_{\ell+1} - \beta, x_t \rangle| \le \|\hat{\beta}_{\ell+1} - \beta\|_1 x_{\max}$, *and the last inequality follows from the occurrence of the occurrence of* ξ_t *and* $N \ge 1$.

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

$$
\begin{aligned}
& \left|\rho_t(r, \hat{y}_t, F^-, F^+) - \rho_t(r, \hat{y}_t, \hat{F}_t^-, \hat{F}_t^+)\right| \\
= & \left|\int_0^r \left[F^-(z - \hat{y}_t) - \hat{F}_t^-(z - \hat{y}_t)\right] dz - r\left[F^+(r - \hat{y}_t) - \hat{F}_t^+(r - \hat{y}_t)\right]\right| \\
\le & \int_0^r \left|F^-(z - \hat{y}_t) - \hat{F}_t^-(z - \hat{y}_t)\right| dz + r\left|F^+(r - \hat{y}_t) - \hat{F}_t^+(r - \hat{y}_t)\right| \\
< & \ r(3c_f N^2 \delta_t + 2\gamma_t) \,.
\end{aligned}
$$

where the last inequality follows from the occurrence of events ξ_t^- *and* ξ_t^+, *and* $N \ge 1$. □

10. Appendix for Section 6: Proof of Theorem 2

We first introduce some definitions that we will extensively rely on throughout our proof of Theorem 2. We start off with the "good" events $\xi_{\ell+1}$, $\xi_{\ell+1}^-$ and $\xi_{\ell+1}^+$ for $\ell \ge 1$ in which the estimates of β, F^-

and F^+ are accurate:

$$\xi_{\ell+1} = \left\{ \|\hat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}} \right\} \tag{40}$$

$$\text{where } \delta_\ell := \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max}x_{\max}^2}{\lambda_0^2\sqrt{N|E_\ell|}} + \frac{\sqrt{d}\,(NL_\ell a_{\max}+1)\,x_{\max}^2}{|E_\ell|\lambda_0^2}, \tag{41}$$

$$\xi_{\ell+1}^- = \left\{ \left|\hat{F}_{\ell+1}^-(z) - F^-(z)\right| \ \leq\ 2N^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + NL_\ell}{|E_\ell|}\right) \right\}, \tag{42}$$

$$\xi_{\ell+1}^+ = \left\{ \left|\hat{F}_{\ell+1}^+(z) - F^+(z)\right| \ \leq\ N\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + NL_\ell}{|E_\ell|}\right) \right\}, \tag{43}$$

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:

$$\mathcal{G}_{i,\ell} := \{|\mathcal{S}_{i,\ell}| \leq L_\ell\}. \tag{44}$$

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 2 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 buyers' strategic bidding behaviour. 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 5, 7, 8, and 9.

(iii) Bound $\mathcal{R}_t^{(2)}$ using Lemmas 5 and 6.

(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 2, the expected revenue in period t is given by

$$\mathsf{rev}_t(r_t) \; = \; \mathbb{E}\left[\max\{b_t^-, \hat{r}_t\}\mathbb{I}\{b_t^+ > \hat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} + \sum_{i\in[N]} r_t^u \mathbb{I}\{b_{i,t} > r_t^u\}\mathbb{I}\{i \text{ is isolated}\} \;\mid\; x_t, r_t\right], \tag{45}$$

where the expectation is taken with respect to $\{(x_\tau, \epsilon_{i,\tau}, a_{i,\tau})\}_{\tau\in[t], i\in[N]}$ and $\hat{r}_t, r_t^u$ are defined in Equations (10) and (11) respectively. Hence, the regret is given by

$$\begin{aligned}
\mathsf{Regret}_t \; &= \; \mathbb{E}\left[\mathrm{REV}_t^\star - \mathsf{rev}_t(r_t)\right] \\
&= \; \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \mathsf{rev}_t(r_t)\right] \\
&= \; \left(\mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right]\right) \\
&\quad + \left(\mathbb{E}\left[\max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} - \mathsf{rev}_t(r_t)\right]\right) \\
&:= \; \mathcal{R}_t^{(1)} + \mathcal{R}_t^{(2)},
\end{aligned} \tag{46}$$

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 (5)) if the seller has full knowledge of F and β; and we defined:

$$\begin{aligned}
\mathcal{R}_t^{(1)} &:= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
\mathcal{R}_t^{(2)} &:= \mathbb{E}\left[\max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} - \mathsf{rev}_t(r_t)\right]
\end{aligned} \tag{47}$$

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

$$\begin{aligned}
\mathcal{R}_t^{(1)} \; &= \; \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
&= \; \mathbb{E}\left[\left(\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\right)\mathbb{I}\{\text{no isolation in } t\}\right] \\
&\quad + \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\left(1 - \mathbb{I}\{\text{no isolation in } t\}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\right]\left(1 - \frac{1}{|E_\ell|}\right) \\
&\quad + \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] \cdot \frac{1}{|E_\ell|} \\
&\le \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}\right] + \frac{v_{\max}}{|E_\ell|}, \qquad (48)
\end{aligned}
$$

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 $\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} \le v_{\max}$.

For simplicity, we define

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

so Equation (48) yields

$$
\mathcal{R}_t^{(1)} \le \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)}\right] + \frac{v_{\max}}{|E_\ell|}, \qquad (49)
$$

where the expectation is taken with respect to the context x_t and reserve price $\hat{r}_t$. Notice that $\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\}$ 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, similar to the Equations (24) and (27) in the proof of Theorem 1, by defining $y_t := \langle \beta, x_t\rangle$, $\hat{y}_t := \langle \hat{\beta}_\ell, x_t\rangle$ and $\rho_t(r, y, F^{(1)}, F^{(2)}) := \int_0^r F^{(2)}(z-y)dz - r\left[F^{(1)}(r-y)\right]$ (Equation (26)), we can apply Proposition 1 and obtain

$$
\begin{aligned}
\widetilde{\mathcal{R}}_t^{(1)} &= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \hat{r}_t\}\mathbb{I}\{v_t^+ > \hat{r}_t\} \;\middle|\; x_t, \hat{r}_t\right] \\
&= \rho_t(r_t^\star, y_t, F^-, F^+) - \rho_t(r_t^\star, \hat{y}_t, F^-, F^+) \\
&\quad + \rho_t(r_t^\star, \hat{y}_t, F^-, F^+) - \rho_t(r_t^\star, \hat{y}_t, \hat{F}_{\ell+1}^-, \hat{F}_{\ell+1}^+) \\
&\quad + \rho_t(r_t^\star, \hat{y}_t, \hat{F}_{\ell+1}^-, \hat{F}_{\ell+1}^+) - \rho_t(\hat{r}_t, \hat{y}_t, \hat{F}_{\ell+1}^-, \hat{F}_{\ell+1}^+) \\
&\quad + \rho_t(\hat{r}_t, \hat{y}_t, \hat{F}_{\ell+1}^-, \hat{F}_{\ell+1}^+) - \rho_t(\hat{r}_t, \hat{y}_t, F^-, F^+) \\
&\quad + \rho_t(\hat{r}_t, \hat{y}_t, F^-, F^+) - \rho_t(\hat{r}_t, y_t, F^-, F^+). \qquad (50)
\end{aligned}
$$

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

We now invoke Lemma 9, where we show that when events $\xi_{\ell+1}$, $\xi^-_{\ell+1}$ and $\xi^+_{\ell+1}$ happen for some phase $\ell \geq 1$, we have for $r \in \{r^\star_t, \hat{r}_t\}$,

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

(ii) $\left|\rho_t(r, \hat{y}_t, F^-, F^+) - \rho_t(r, \hat{y}_t, \hat{F}^-_{\ell+1}, \hat{F}^+_{\ell+1})\right| \;\leq\; 3rN^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right)$ 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 (50), we get

$$
\begin{aligned}
\widetilde{\mathcal{R}}^{(1)}_t \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi^-_{\ell+1} \cap \xi^+_{\ell+1}\right\} \;\leq\; & 3(r^\star_t + \hat{r}_t)c_f N^2 \delta_\ell \\
& + 3(r^\star_t + \hat{r}_t) N^2 \left(\gamma_\ell + c_f \delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right) \\
& + \rho_t(r^\star_t, \hat{y}_t, \hat{F}^-_{\ell+1}, \hat{F}^+_{\ell+1}) - \rho_t(\hat{r}_t, \hat{y}_t, \hat{F}^-_{\ell+1}, \hat{F}^+_{\ell+1}) \,. \qquad (51)
\end{aligned}
$$

We recall that the seller's pricing decision $\hat{r}_t$ when no isolation occurs is defined in Equation (11), and realize that in fact $\hat{r}_t = \arg\max_{r \in (0, v_{\max}]} \rho_t(r, \hat{y}_t, \hat{F}^-_{\ell+1}, \hat{F}^+_{\ell+1})$. So, by the optimality of $\hat{r}_t$ and $r^\star_t \leq v_{\max}$, we obtain the fact that $\rho_t(r^\star_t, \hat{y}_t, \hat{F}^-_{\ell+1}, \hat{F}^+_{\ell+1}) - \rho_t(\hat{r}_t, \hat{y}_t, \hat{F}^-_{\ell+1}, \hat{F}^+_{\ell+1}) \leq 0$. Using this inequality in (51), we get

$$
\begin{aligned}
& \widetilde{\mathcal{R}}^{(1)}_t \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi^-_{\ell+1} \cap \xi^+_{\ell+1}\right\} \\
\leq\; & 6v_{\max} c_f N^2 \delta_\ell + 6 v_{\max} N^2 \left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right) \\
=\; & 12 v_{\max} c_f N^2 \delta_\ell + 6 v_{\max} N^2 \left(\frac{\sqrt{\log(|E_\ell|)}}{\sqrt{2N|E_\ell|}} + \frac{c_f + L_\ell}{|E_\ell|}\right) \\
=\; & 12 v_{\max} c_f N^2 \delta_\ell + \frac{6 v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{6 v_{\max} N^2 (c_f + L_\ell)}{|E_\ell|} \,, \qquad (52)
\end{aligned}
$$

where we used the fact that $r^\star_t, \hat{r}_t \leq v_{\max}$ 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 $\mathcal{R}_t^{(1)}$ in period $t \in E_{\ell+1}$, we continue to bound Equation (49):

$$
\begin{aligned}
\mathcal{R}_t^{(1)} &\leq \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)}\right] + \frac{v_{\max}}{|E_\ell|} \\
&= \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi_{\ell+1}^- \cap \xi_{\ell+1}^+\right\}\right] + \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1}^c \cup \left(\xi_{\ell+1}^-\right)^c \cup \left(\xi_{\ell+1}^+\right)^c\right\}\right] + \frac{v_{\max}}{|E_\ell|} \\
&\leq \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi_{\ell+1}^- \cap \xi_{\ell+1}^+\right\}\right] + v_{\max}\mathbb{P}\left(\xi_{\ell+1}^c \cup \left(\xi_{\ell+1}^-\right)^c \cup \left(\xi_{\ell+1}^+\right)^c\right) + \frac{v_{\max}}{|E_\ell|} \\
&\leq 12 v_{\max} c_f N^2 \delta_\ell + \frac{6 v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f + L_\ell) + 9N + 15d + 9\right)}{|E_\ell|}\,, \qquad (53)
\end{aligned}
$$

where the second inequality follows from a simple observation that $\widetilde{\mathcal{R}}_t^{(1)} \leq v_{\max}$ almost surely, and the third inequality uses Equation (52) and Lemma 10, which shows $\mathbb{P}\left(\xi_{\ell+1}^c \cup \left(\xi_{\ell+1}^-\right)^c \cup \left(\xi_{\ell+1}^+\right)^c\right) \leq (9N + 15d + 8)/|E_\ell|$,

(iii) Bounding $\mathcal{R}_t^{(2)}$: So far, we have bounded $\mathcal{R}_t^{(1)}$ for $t \in E_{\ell+1}$ $(\ell \geq 1)$, and will move on to bound $\mathcal{R}_t^{(2)}$ defined in Equation (46) for $t \in E_\ell$ for any $\ell \geq 1$. We define

$$
b_{-i,t}^+ = \max_{j \neq i} b_{j,t} \quad \text{and} \quad v_{-i,t}^+ = \max_{j \neq i} v_{j,t}\,, \qquad (54)
$$

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

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

$$
\begin{aligned}
&\leq \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\hat{r}_t\}\mathbb{I}\{\max\{v_{-i,t}^+,\hat{r}_t\} < v_{i,t} < \max\{b_{-i,t}^+\hat{r}_t\}\}\right] \\
&\qquad + \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\hat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{b_{-i,t}^+\hat{r}_t\}\} - \max\{b_t^-,\hat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b_{-i,t}^+\hat{r}_t\}\}\right] \\
&\leq \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v_{-i,t}^+,\hat{r}_t\} < v_{i,t} < \max\{b_{-i,t}^+\hat{r}_t\}\}\right] \\
&\qquad + \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\hat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{b_{-i,t}^+\hat{r}_t\}\} - \max\{b_t^-,\hat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b_{-i,t}^+\hat{r}_t\}\}\right], \quad (55)
\end{aligned}
$$

where the first inequality follows from Equation (45); the third inequality is due to the fact that $\sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\hat{r}_t\}\mathbb{I}\{\max\{b_{-i,t}^+\hat{r}_t\} < v_{i,t} < \max\{v_{-i,t}^+,\hat{r}_t\}\}\right] \geq 0$; and the last inequality holds because $\max\{v_t^-,\hat{r}_t\} \leq v_{\max}$. To continue the bound for Equation (55), we use the definition of $\mathcal{B}_{i,\ell} := \mathcal{B}_{i,\ell}^s \cup \mathcal{B}_{i,\ell}^o$ in Lemma 6, where

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

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, 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. The "s" and "o" present represent shading and overbidding respectively. Hence, for any period $t \in E_\ell/\mathcal{B}_{i,\ell} := \left\{t \in E_\ell : \mathbb{I}\left\{v_{i,t} > \{b_{-i,t}^+,\hat{r}_t\}\right\} = \mathbb{I}\left\{b_{i,t} > \{b_{-i,t}^+,\hat{r}_t\}\right\}\right\}$ (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 $\mathbb{I}\{v_{i,t} > \max\{b_{-i,t}^+,\hat{r}_t\}\} = \mathbb{I}\{b_{i,t} > \max\{b_{-i,t}^+,\hat{r}_t\}\}$. Therefore, defining $\mathcal{B}_\ell := \cup_{i\in[N]}\mathcal{B}_{i,\ell}$, we have

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

$$
\begin{aligned}
&+\sum_{i\in[N]}\mathbb{E}\left[\left(\max\{v_t^-,\hat{r}_t\}-\max\{b_t^-,\hat{r}_t\}\right)\mathbb{I}\{b_{i,t}>\max\{b_{-i,t}^+,\hat{r}_t\}\}\right]\\
\leq\ &\sum_{i\in[N]}v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v_{-i,t}^+,\hat{r}_t\}<v_{i,t}<\max\{b_{-i,t}^+,\hat{r}_t\}\}\right]+\mathbb{E}\left[\max\{v_t^-,\hat{r}_t\}-\max\{b_t^-,\hat{r}_t\}\right]\\
\leq\ &\sum_{i\in[N]}v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v_{-i,t}^+,\hat{r}_t\}<v_{i,t}<\max\{b_{-i,t}^+,\hat{r}_t\}\}\right]+\mathbb{E}\left[\left(v_t^- - b_t^-\right)^+\right].
\end{aligned}
$$

The first inequality follows from Equation (55); the first equality follows from the fact that $t\in E_\ell/\mathcal{B}_\ell$; the second inequality holds because $\sum_{i\in[N]}\mathbb{I}\{b_{i,t}>\max\{b_{-i,t}^+\hat{r}_t\}\}\ \leq\ \sum_{i\in[N]}\mathbb{I}\{b_{i,t}>b_{-i,t}^+\}\}=1$; 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

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

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

$$
\mathcal{R}_t^{(2)}\mathbb{I}\{t\in E_\ell/\mathcal{B}_\ell\}\ \leq\ v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v_{-i^\star,t}^+,\hat{r}_t\}<v_{i^\star,t}<\max\{b_{-i^\star,t}^+,\hat{r}_t\}\}\right]+\mathbb{E}\left[\left(v_t^- - b_t^-\right)^+\right], \quad (56)
$$

To bound the first term in Equation (56), we again evoke the inequality $\max\{a,c\}-\max\{b,c\}=(a-b)^+$ for any $a,b,c\in\mathbb{R}$ and get $\max\{b_{-i^\star,t}^+,\hat{r}_t\}-\max\{v_{-i^\star,t}^+,\hat{r}_t\}\leq\left(b_{-i^\star,t}^+-v_{-i^\star,t}^+\right)^+$. Hence,

$$
\begin{aligned}
&\mathbb{E}\left[\mathbb{I}\{\max\{v_{-i^\star,t}^+,\hat{r}_t\}<v_{i^\star,t}<\max\{b_{-i^\star,t}^+,\hat{r}_t\}\}\right]\\
\leq\ &\mathbb{E}\left[\mathbb{I}\{\max\{b_{-i^\star,t}^+,\hat{r}_t\}-\left(b_{-i^\star,t}^+-v_{-i^\star,t}^+\right)^+<v_{i^\star,t}<\max\{b_{-i^\star,t}^+,\hat{r}_t\}\}\right]\\
=\ &\mathbb{E}\left[\mathbb{E}\left[\mathbb{I}\{\max\{b_{-i^\star,t}^+,\hat{r}_t\}-\left(b_{-i^\star,t}^+-v_{-i,t}^+\right)^+<v_{i^\star,t}<\max\{b_{-i^\star,t}^+,\hat{r}_t\}\}\ \Big|\ b_{-i^\star,t}^+,v_{-i^\star,t}^+\right]\right]\\
=\ &\mathbb{E}\left[\int_{\max\{b_{-i^\star,t}^+,\hat{r}_t\}-\left(b_{-i^\star,t}^+-v_{-i^\star,t}^+\right)^+-\langle\beta,x_t\rangle}^{\max\{b_{-i^\star,t}^+,\hat{r}_t\}-\langle\beta,x_t\rangle}f(z)dz\right]\\
\leq\ &c_f\mathbb{E}\left[\left(b_{-i^\star,t}^+-v_{-i^\star,t}^+\right)^+\right]. \qquad (57)
\end{aligned}
$$

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 (57), we have

$$\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t}, \hat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t}, \hat{r}_t\}\}\right] \;\le\; c_f(-a_{j,t})^+ \;\le\; c_f \sum_{i\in[N]} (-a_{i,t})^+ \,. \tag{58}$$

To bound the second term in Equation (56), namely $\mathbb{E}\left[\left(v_t^- - b_t^-\right)^+\right]$, without loss of generality assume $v_{1,t} \ge v_{2,t} \ge \cdots \ge v_{N,t}$. Hence $v_t^- = v_{2,t}$. If $b_{2,t} \le b_t^-$, we have $v_t^- - b_t^- \le v_{2,t} - b_{2,t} = a_{2,t}$. 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,

$$\mathbb{E}\left[\left(v_t^- - b_t^-\right)^+\right] \le \max_{j\in[N]} (a_{j,t})^+ \le \sum_{j\in[N]} (a_{j,t})^+ \,. \tag{59}$$

Finally, combining Equations (56), (58), and (59), we have for any $t \in E_\ell$ and $\ell \ge 1$

$$\mathcal{R}_t^{(2)} \mathbb{I}\{t \in E_\ell / \mathcal{B}_\ell\} \le v_{\max} c_f \sum_{i\in[N]} (-a_{i,t})^+ + \sum_{i\in[N]} (a_{i,t})^+ \le (v_{\max} c_f + 1) \sum_{i\in[N]} |a_{i,t}| \tag{60}$$

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.

$$\begin{aligned}
&\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(1)} \\
&\le \sum_{t\in E_{\ell+1}} \left(12 v_{\max} c_f N^2 \delta_\ell + \frac{6 v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f + L_\ell) + 9N + 15d + 9\right)}{|E_\ell|} \right) \\
&= |E_{\ell+1}| \left(12 v_{\max} c_f N^2 \delta_\ell + \frac{6 v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f + L_\ell) + 9N + 15d + 9\right)}{|E_\ell|} \right) \\
&= |E_{\ell+1}| \cdot \frac{3 v_{\max}\sqrt{2N^3 \log(|E_\ell|)}}{\sqrt{|E_\ell|}} \left(\frac{4 c_f \epsilon_{\max} x_{\max}^2 \sqrt{d}}{\lambda_0^2} + 1 \right) \\
&\quad + \frac{|E_{\ell+1}|}{|E_\ell|} \left(\frac{12 v_{\max} c_f N^2 \sqrt{d}\,(N L_\ell a_{\max} + 1)\, x_{\max}^2}{\lambda_0^2} + v_{\max}\left(6N^2(c_f + L_\ell) + 9N + 15d + 9\right) \right) \\
&\le c_1^1 c_f \sqrt{dTN^3 \log(|E_\ell|)} + c_2^2 c_f \sqrt{d} N^3 L_\ell T^{\frac{1}{4}} \\
&\le c_1 c_f \sqrt{dN^3 \log(|E_\ell|)} \left(\sqrt{T} + \frac{\sqrt{N^3 \log(|E_\ell|)} T^{\frac{1}{4}}}{\log(1/\eta)} \right),
\end{aligned} \tag{61}$$

for some absolute constants $c_1^1, c_1^2, c_1 > 0$. The first inequality follows from Equation (53). In the second equality, we then used the definition of $\delta_\ell = \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max} x_{\max}^2}{\lambda_0^2 \sqrt{N|E_\ell|}} + \frac{\sqrt{d}(N L_\ell a_{\max}+1) x_{\max}^2}{|E_\ell| \lambda_0^2}$, defined in Equation (41). In the second inequality, we relied on the construction of the length of phases in Algorithm 2, 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 $L_\ell = \log(v_{\max}^2 N|E_\ell|^4 - 1)/\log(\frac{1}{\eta})$.

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 (70) of Lemma 6. Denote $K_{\ell+1} := 2L_{\ell+1} + 4c_f + 8\log(|E_{\ell+1}|)$. Then, we have

$$\begin{aligned}
\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(2)} &= \mathbb{E}\left[\sum_{t\in\mathcal{B}_{\ell+1}} \mathcal{R}_t^{(2)}\right] + \mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}} \mathcal{R}_t^{(2)}\right] \\
&\le v_{\max}\mathbb{E}\left[|\mathcal{B}_{\ell+1}|\cdot \mathbb{I}\{|\mathcal{B}_{\ell+1}| \le NK_{\ell+1}\right] + v_{\max}\mathbb{E}\left[|\mathcal{B}_{\ell+1}|\cdot \mathbb{I}\{|\mathcal{B}_{\ell+1}| > NK_{\ell+1}\}\right] \\
&\quad + (v_{\max}c_f + 1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}} \sum_{i\in[N]} |a_{i,t}|\right] \\
&\le v_{\max}NK_{\ell+1} + v_{\max}|E_{\ell+1}|\cdot \mathbb{P}\left(|\mathcal{B}_{\ell+1}| > NK_{\ell+1}\right) + (v_{\max}c_f+1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}} \sum_{i\in[N]} |a_{i,t}|\right] \\
&\le v_{\max}NK_{\ell+1} + 4v_{\max}N + (v_{\max}c_f+1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}} \sum_{i\in[N]} |a_{i,t}|\right] \\
&\le v_{\max}N(K_{\ell+1}+4) + (v_{\max}c_f+1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}} \sum_{i\in[N]} |a_{i,t}|\right],
\end{aligned} \tag{62}$$

where the first inequality follows from Equation (60) 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 6 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 (15), and consider the following

$$\begin{aligned}
\mathbb{E}\left[\sum_{t\in E_{\ell+1}} \sum_{i\in[N]} |a_{i,t}|\right] &\le \mathbb{E}\left[\sum_{t\in\mathcal{S}_{\ell+1}} \sum_{i\in[N]} |a_{i,t}|\right] + \mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{S}_{\ell+1}} \sum_{i\in[N]} \frac{1}{|E_{\ell+1}|}\right] \\
&\le Na_{\max}\mathbb{E}\left[|\mathcal{S}_{\ell+1}|\right] + N \\
&= Na_{\max}\mathbb{E}\left[|\mathcal{S}_{\ell+1}|\cdot\left(\mathbb{I}\{|\mathcal{S}_{\ell+1}| \le NL_{\ell+1}\} + \mathbb{I}\{|\mathcal{S}_{\ell+1}| > NL_{\ell+1}\}\right)\right] + N \\
&\le Na_{\max}\left(NL_{\ell+1} + |E_{\ell+1}|\cdot\mathbb{P}\left(|\mathcal{S}_{\ell+1}| > NL_{\ell+1}\right)\right) + N \\
&\le N^2 a_{\max}\left(L_{\ell+1}+1\right) + N\,,
\end{aligned} \tag{63}$$

where the first inequality holds because $|a_{i,t}| \leq 1/|E_{\ell+1}|$ for all $t \in E_{\ell+1}/\mathcal{S}_{\ell+1}$ and the fourth inequality follows from Lemma 5 that shows $\mathbb{P}\left(|\mathcal{S}_{i,\ell+1}| > L_{\ell+1}\right) \leq 1/|E_{\ell+1}|$, which implies $\mathbb{P}\left(|\mathcal{S}_{\ell+1}| \leq NL_{\ell+1}\right) \geq \mathbb{P}\left(\cap_{i\in[N]}\left\{|\mathcal{S}_{i,\ell+1}| \leq L_{\ell+1}\right\}\right) \geq 1 - N/|E_{\ell+1}|$.

Hence, Equations (62) and (63) show that $\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(2)}$ is upper bounded as

$$\begin{aligned}\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(2)} \;&\leq\; v_{\max} N(K_\ell + 4) + (v_{\max} c_f + 1)\left(N^2 a_{\max}\left(L_{\ell+1} + 1\right) + N\right) \\ &\leq\; c_2 c_f N^2 \cdot \frac{\log(|E_{\ell+1}|)}{\log(1/\eta)}\,, \qquad (64)\end{aligned}$$

for some absolute constant $c_2 > 0$. Combining this with the upper bound $c_1 c_f \sqrt{dN^3 \log(|E_\ell|)}\left(\sqrt{T} + \frac{\sqrt{N^3\log(|E_\ell|)}T^{\frac{1}{4}}}{\log(1/\eta)}\right)$ shown in Equation (61), the expected cumulative regret in phase $E_{\ell+1}$ is

$$\sum_{t\in E_{\ell+1}} \mathsf{Regret}_t \;\leq\; c_3 c_f \sqrt{dN^3\log(T)}\left(\sqrt{T} + \frac{\sqrt{N^3\log(T)}T^{\frac{1}{4}}}{\log(1/\eta)}\right),$$

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

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

10.1. Lemmas for proving Theorem 2

LEMMA 5 **(Bounding number of lies)**. *Consider a buyer $i \in [N]$ and some phase $\ell \geq 1$. Then, the cardinality of $\mathcal{S}_{i,\ell} = \{t \in E_\ell : |a_{i,t}| \geq 1/|E_\ell|\}$ is bounded as*

$$\mathbb{P}\left(|\mathcal{S}_{i,\ell}| > L_\ell\right) \leq \frac{1}{|E_\ell|}\,,$$

where $L_\ell = \log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)/\log(1/\eta)$ and $v_{\max}$ is the maximum possible buyer valuation.

Proof of Lemma 5. According to the definitions of the cumulative discounted utility defined in Equation (1) and the NPAC-S policy in Algorithm 2, 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

$$u_{i,t}(b) = \begin{cases} \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right) \mathbb{I}\{b > \max\{r_t, b^+_{-i,t}\}\} & \text{no isolation} \\ (v_{i,t} - r_t)\, \mathbb{I}\{b > r_t\} & i \text{ is isolated} \\ 0 & j \neq i \text{ is isolated} \end{cases}, \tag{65}$$

where $b^+_{-i,t}$ is the highest bid excluding that of buyer i, and the reserve price $r_t = \hat{r}_t \mathbb{I}\{\text{no isolation in } t\} + r^u_t(1 - \mathbb{I}\{\text{no isolation in } t\})$ ($\hat{r}_t$ and r^u_t are defined in Equations (10) and (11) 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

$$\begin{aligned} u^-_{i,t}\mathbb{I}\{\text{no isolation in } t\} &= (u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t}))\, \mathbb{I}\{\text{no isolation in } t\} \\ &= \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right) \mathbb{I}\{v_{i,t} > \max\{r_t, b^+_{-i,t}\}\} \\ &\quad - \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right) \mathbb{I}\{b_{i,t} > \max\{r_t, b^+_{-i,t}\}\} \\ &= \left|v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right| \mathbb{I}\{v_{i,t} > \max\{r_t, b^+_{-i,t}\} > b_{i,t}\} \\ &\quad + \left|v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right| \mathbb{I}\{v_{i,t} < \max\{r_t, b^+_{-i,t}\} < b_{i,t}\} \\ &\geq 0\,. \end{aligned} \tag{66}$$

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,

$$u^-_{i,t}\mathbb{I}\{i \text{ is isolated}\} = (u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t}))\, \mathbb{I}\{i \text{ is isolated}\}$$

$$
\begin{aligned}
&= (v_{i,t}-r_t)\,\mathbb{I}\{v_{i,t}>r_t\}-(v_{i,t}-r_t)\,\mathbb{I}\{b_{i,t}>r_t\}\\
&= (v_{i,t}-r_t)\,\mathbb{I}\{v_{i,t}>r_t>b_{i,t}\}+(-v_{i,t}+r_t)\,\mathbb{I}\{v_{i,t}<r_t<b_{i,t}\}\,. \qquad (67)
\end{aligned}
$$

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

$$
\begin{aligned}
&\mathbb{E}[u_{i,t}^- \mid a_{i,t}]\\
=\ &\mathbb{E}[u_{i,t}^-\mathbb{I}\{i \text{ is isolated}\}+u_{i,t}^-\mathbb{I}\{\text{no isolation in } t\} \mid a_{i,t}]\\
\geq\ &\mathbb{E}[u_{i,t}^-\mathbb{I}\{i \text{ is isolated}\} \mid a_{i,t}]\\
=\ &\mathbb{E}\left[(v_{i,t}-r_t)\,\mathbb{I}\{v_{i,t}>r_t>b_t\}+(-v_{i,t}+r_t)\,\mathbb{I}\{b_t<r_t<v_{i,t}\} \mid a_{i,t}\right]\\
=\ &\frac{1}{v_{\max}N|E_\ell|}\mathbb{E}\left[\mathbb{E}\left[\int_{v_{i,t}-a_{i,t}}^{v_{i,t}}(v_{i,t}-r)dr+\int_{v_{i,t}}^{v_{i,t}+a_{i,t}}(-v_{i,t}+r)dr \;\middle|\; a_{i,t}, v_{i,t}\right] \;\middle|\; a_{i,t}\right]\\
=\ &\frac{(a_{i,t})^2}{v_{\max}N|E_\ell|}\,. \qquad (68)
\end{aligned}
$$

The first inequality follows from $u_{i,t}^-\mathbb{I}\{i \text{ is isolated}\} \geq 0$ as demonstrated in Equation (66). Now we lower bound the total expected utility loss in phase E_ℓ. First, by Equations (66) and (67), 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

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

$$\geq \mathbb{E}\left[\sum_{t=s_{\ell+1}-\tilde{z}}^{s_{\ell+1}-1} \frac{\eta^t}{v_{\max}N|E_\ell|^3} \cdot \mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right]$$

$$= \frac{\eta^{s_{\ell+1}}\left(1-\eta^{-\tilde{z}}\right)}{(1-\eta)v_{\max}N|E_\ell|^3}\mathbb{P}\left(|\mathcal{S}_{i,\ell}| \geq \tilde{z}\right), \tag{69}$$

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 (68) 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 $\ell+1, \ell+2, \ldots$, 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 (69), 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 $\tilde{z} = \log\left(v^2_{\max}N|E_\ell|^4 - 1\right) / \log(1/\eta)$ and by rearranging terms, the inequality above yields

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

□

Lemma 6 **(Bounding outcome changes for non-isolation periods)**. *Define the following two sets of time periods:*

$$\mathcal{B}^s_{i,\ell} = \left\{t \in E_\ell : \mathbb{I}\left\{v_{i,t} \geq \{b^+_{-i,t}, \hat{r}_t\}\right\} = 1\ ,\ \mathbb{I}\left\{b_{i,t} \geq \{b^+_{-i,t}, \hat{r}_t\}\right\} = 0\right\} \quad \text{and}$$

$$\mathcal{B}^o_{i,\ell} = \left\{t \in E_\ell : \mathbb{I}\left\{v_{i,t} \geq \{b^+_{-i,t}, \hat{r}_t\}\right\} = 0\ ,\ \mathbb{I}\left\{b_{i,t} \geq \{b^+_{-i,t}, \hat{r}_t\}\right\} = 1\right\}, \tag{70}$$

where $b^+_{-i,t}$ is the highest among all bids excluding that submitted by buyer i, and $\hat{r}_t$ is the reserve price offered to all buyers if no isolation occurs (defined in Equation (11)). Then, for $\mathcal{B}_{i,\ell} := \mathcal{B}^s_{i,\ell} \cup \mathcal{B}^o_{i,\ell}$, we have

$$\mathbb{P}\left(|\mathcal{B}_{i,\ell}| \le 2L_\ell + 4c_f + 8\log(|E_\ell|)\right) \;\ge\; 1 - \frac{4}{|E_\ell|}\,.$$

Here, the probability is taken with respect to the randomness in $\{(x_\tau, \epsilon_{i,\tau}, a_{i,\tau})\}_{\tau \in E_\ell, i \in [N]}$.

Proof of Lemma 6. We first provide a road-map for the proof of this lemma. Recall the definition of $\mathcal{S}_{i,\ell}$ in Equation (15), and hence $E_\ell/\mathcal{S}_{i,\ell} = \{t \in E_\ell : |a_{i,t}| \le 1/|E_\ell|\}$, where $a_{i,t} = v_{i,t} - b_{i,t}$. Note that $E_\ell/\mathcal{S}_{i,\ell}$ can be considered as the set of periods during which buyer i "slightly" corrupts her bids. We start with the case when a buyer shades her bids. For any give period $t \in E_\ell$, we provide an upper bound on the probability that a buyer "slightly" shades her bids but changes the outcome of the auction had she bid truthfully (i.e. $t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}$). Then, we translate these probabilities for every period in E_ℓ into a high probability bound for the number of such periods (i.e. a bound for $|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}|$), which further yields a bound for $|\mathcal{B}^s_{i,\ell}|$ since $|\mathcal{B}^s_{i,\ell}| = |\mathcal{S}_{i,\ell} \cap \mathcal{B}^s_{i,\ell}| + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| \le |\mathcal{S}_{i,\ell}| + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}|$, and $|\mathcal{S}_{i,\ell}|$ is bounded in virtue of Lemma 5. A symmetric argument results in a bound for $|\mathcal{B}^o_{i,\ell}|$ by considering the case where buyers overbid, and the final result will follow from the fact that $\mathcal{B}_{i,\ell} = \mathcal{B}^s_{i,\ell} \cup \mathcal{B}^o_{i,\ell}$.

In light of this roadmap, we now formally prove the lemma. Defining $\mathcal{H}_{i,t} := \{(b^+_{-i,\tau}, \hat{r}_\tau, x_\tau)\}_{\tau \in [t]}$, we have

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}\} \mid \mathcal{H}_{i,t}\right] \\
=\;& \mathbb{P}\left(t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell} \mid \mathcal{H}_{i,t}\right) \\
=\;& \mathbb{P}\left(v_{i,t} \ge \max\{b^+_{-i,t}, \hat{r}_t\}\;,\; b_{i,t} < \max\{b^+_{-i,t}, \hat{r}_t\}\;,\; a_{i,t} \in (0, 1/|E_\ell|) \mid \mathcal{H}_{i,t}\right) \\
=\;& \mathbb{P}\left(\max\{b^+_{-i,t}, \hat{r}_t\} - \langle x_t, \beta\rangle \le \epsilon_{i,t} \le \max\{b^+_{-i,t}, \hat{r}_t\} - \langle x_t, \beta\rangle + a_{i,t}\;,\; a_{i,t} \in (0, 1/|E_\ell|) \mid \mathcal{H}_{i,t}\right) \\
\le\;& \mathbb{P}\left(\max\{b^+_{-i,t}, \hat{r}_t\} - \langle x_t, \beta\rangle \le \epsilon_{i,t} \le \max\{b^+_{-i,t}, \hat{r}_t\} - \langle x_t, \beta\rangle + 1/|E_\ell| \mid \mathcal{H}_{i,t}\right) \\
=\;& \mathbb{E}\left[\int_{\max\{b^+_{-i,t}, \hat{r}_t\} - \langle x_t, \beta\rangle}^{\max\{b^+_{-i,t}, \hat{r}_t\} - \langle x_t, \beta\rangle + 1/|E_\ell|} f(z)dz \;\middle|\; \mathcal{H}_{i,t}\right] \\
\le\;& \frac{c_f}{|E_\ell|}\,.
\end{aligned} \tag{71}$$

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

Define $\zeta_t = \mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}\}$ and $\phi_t = \mathbb{E}\left[\mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}\} \mid \mathcal{H}_{i,t}\right]$. 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 13, by setting $z_{1,t} = \zeta_t$, $z_{2,t} = \phi_t$, $\tilde{\gamma} = 1/2$ and $A = 2\log(|E_\ell|)$ we have $|z_{1,t} - z_{2,t}| \le 1$

$$\mathbb{P}\left(\frac{1}{2}\sum_{t\in E_\ell}\zeta_t \ge \sum_{t\in E_\ell}\phi_t + 2\log(|E_\ell|)\right) \le \exp\left(-\log(|E_\ell|)\right). \tag{72}$$

Now, according to Equation (71), we have $\phi_t \le c_f/|E_\ell|$, so $\sum_{t\in E_\ell}\phi_t \le c_f$. Moreover, $|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| = \sum_{t\in E_\ell}\zeta_t$. Hence, following Equation (72), we have

$$\begin{aligned}\mathbb{P}\left(|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| \ge 2c_f + 4\log(|E_\ell|)\right) &\le \mathbb{P}\left(\frac{1}{2}\sum_{t\in E_\ell}\zeta_t \ge \sum_{t\in E_\ell}\phi_t + 2\log(|E_\ell|)\right)\\ &\le \exp\left(-\log(|E_\ell|)\right) = \frac{1}{|E_\ell|}.\end{aligned} \tag{73}$$

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

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

The second inequality follows from $|\mathcal{B}^s_{i,\ell}| \le L_\ell + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}|$ when the event $\mathcal{G}_{i,t}$ occurs; the third inequality applies the union bound, and the final inequality follows from Equation (73) and Lemma 5.

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

LEMMA 7 **(Bounding Estimation Errors in β).** *For any phase E_ℓ and $\gamma > 0$, we have*

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

where λ_0^2 is the minimum eigenvalue of the covariance matrix Σ, $\hat{\beta}_{\ell+1}$ is defined in Equation (8), and $L_\ell = \log(v_{\max}^2 N|E_\ell|^4 - 1)/\log(1/\eta)$. 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

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

Proof of Lemma 7. The proof of this lemma is based on several modifications to that of Lemma 3 to resolve the issues that arise when estimating β in the presence of corrupted bids submitted by buyers.

Recall that the smallest eigenvalue λ_0^2 of the covariance matrix Σ of $x \sim \mathcal{D}$ is greater than 0, and as argued in the proof of Lemma 3, we note 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 $\lambda_{\min}(X^\top X/|E_\ell|) \ge \lambda_0^2/2$, 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 (12) we have

$$\begin{aligned}
\hat{\beta}_{\ell+1} = \left(X^\top X\right)^{-1} X^\top \begin{pmatrix} \bar{b}_1 \\ \vdots \\ \bar{b}_t \end{pmatrix} &= \left(X^\top X\right)^{-1} X^\top \begin{pmatrix} \frac{\sum_{i\in[N]} v_{i,1} - a_{i,1}}{N} \\ \vdots \\ \frac{\sum_{i\in[N]} v_{i,t} - a_{i,t}}{N} \end{pmatrix} \\
&= \beta + \left(X^\top X\right)^{-1} X^\top \begin{pmatrix} \frac{\sum_{i\in[N]} \epsilon_{i,1} - a_{i,1}}{N} \\ \vdots \\ \frac{\sum_{i\in[N]} \epsilon_{i,t} - a_{i,t}}{N} \end{pmatrix} \\
&= \beta + \left(X^\top X\right)^{-1} X^\top \left(\bar{\mathcal{E}} - A\right)\,, \qquad (74)
\end{aligned}$$

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,

$$\begin{aligned}\|\hat{\beta}_{\ell+1}-\beta\|_2 &= \|\left(X^\top X\right)^{-1}X^\top\left(\bar{\mathcal{E}}-A\right)\|_2\\ &\le \frac{1}{|E_\ell|\lambda_0^2}\left(\|X^\top\bar{\mathcal{E}}\|_2+\|X^\top A\|_2\right). \end{aligned}\tag{75}$$

Denote X^j as the jth column of X, i.e. the jth row of $X^\top$ for $j=1,2\dots d$, 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 $\gamma>0$,

$$\bigcap_{j\in[d]}\left\{\left|\bar{\mathcal{E}}^\top X^j\right|\le\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right\}\subseteq\left\{\frac{1}{|E_\ell|\lambda_0^2}\cdot\|X^\top\bar{\mathcal{E}}\|_2\le\gamma\right\}. \tag{76}$$

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 $\tilde{\gamma}>0$

$$\mathbb{P}\left(\left|N\bar{\mathcal{E}}^\top X^j\right|\le\tilde{\gamma}\right)\ge 1-2\exp\left(-\frac{\tilde{\gamma}^2}{2{\epsilon_{\max}}^2x_{\max}^2|E_\ell|N}\right). \tag{77}$$

Replacing $\tilde{\gamma}$ with $N|E_\ell|\lambda_0^2\gamma/\sqrt{d}$ and using Equation (76) yields:

$$\begin{aligned}\mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\cdot\|X^\top\bar{\mathcal{E}}\|_2\le\gamma\right\}\right) &\ge \mathbb{P}\left(\bigcap_{j\in[d]}\left\{\left|\bar{\mathcal{E}}^\top X^j\right|\le\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right\}\right)\\ &\ge 1-\sum_{j\in[d]}\mathbb{P}\left(\left|\bar{\mathcal{E}}^\top X^j\right|>\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right)\\ &\ge 1-2d\exp\left(-\frac{N\gamma^2\lambda_0^4|E_\ell|}{2{\epsilon_{\max}}^2x_{\max}^2d}\right), \end{aligned}\tag{78}$$

where the first inequality follows from Equation (76), the second inequality applies the union bound, and the last inequality follows from Equation (77).

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 5 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.

$$\begin{aligned}\|X^\top A\|_2^2=\sum_{j\in[d]}\left|A^\top X^j\right|^2 &= \sum_{j\in[d]}\left(\frac{\sum_{\tau\in E_\ell}\sum_{i\in[N]}a_{i,\tau}X_{\tau j}}{N}\right)^2\\ &\le \sum_{j\in[d]}\left(\frac{\sum_{\tau\in E_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\right)^2. \end{aligned}\tag{79}$$

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

$$\frac{\sum_{\tau\in S_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N} \;\le\; \sum_{\tau\in S_\ell} a_{\max}x_{\max} \;\le\; NL_\ell a_{\max}x_{\max}\,, \tag{80}$$

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 $E_\ell/\mathcal{S}_\ell$,

$$\frac{\sum_{\tau\in E_\ell/\mathcal{S}_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N} \;\le\; \sum_{\tau\in E_\ell/\mathcal{S}_\ell}\frac{x_{\max}}{|E_\ell|} \;\le\; \sum_{\tau\in E_\ell}\frac{x_{\max}}{|E_\ell|} = x_{\max}\,. \tag{81}$$

Combining Equations (79), (80), and (81), we have

$$\|X^\top A\|_2 \;\le\; \sqrt{d\left(\frac{\sum_{\tau\in[t]}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\right)^2} \;\le\; \sqrt{d}\,(NL_\ell a_{\max}+1)\,x_{\max}\,. \tag{82}$$

Now, following the same arguments of Equation (39) in the proof of Lemma 3, but by replacing t with $|E_\ell|$, we have

$$\mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right)\ge\frac{\lambda_0^2}{2}\right) \;\ge\; 1-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\,. \tag{83}$$

Putting everything together, we get

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

The first inequality follows from the fact that $\|z\|_1\le\|z\|_2$ for any vector z; the second inequality follows from Equation (75); the third inequality follows from Equation (82) 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 (78), (83) and Lemma 5. □

LEMMA 8 **(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*

$$\mathbb{P}\left(\left|\hat{F}^-_{\ell+1}(z) - F^-(z)\right| \le 2N^2 z_\ell\right) \ \ge\ 1 - 4\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{4(d+N)}{|E_\ell|} - 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)$$
$$\mathbb{P}\left(\left|\hat{F}^+_{\ell+1}(z) - F^+(z)\right| \le N z_\ell\right) \ \ge\ 1 - 4\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{4(d+N)}{|E_\ell|} - 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right),$$

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 (41), and $L_\ell = \log(v_{\max}^2 N|E_\ell|^4 - 1)/\log(1/\eta)$.

Proof of Lemma 8. We first bound the error in the estimate of F, namely $\left|\hat{F}_{\ell+1}(z) - F(z)\right|$. 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 $\hat{F}^-_{\ell+1}(z)$ and $\hat{F}^+_{\ell+1}(z)$ in Equation (14) to show the desired probability bounds.

We first upper and lower bound $\hat{F}^-_{\ell+1}(z)$ for any $z \in \mathbb{R}$. Recall the event $\mathcal{S}_{i,\ell} = \{t \in E_\ell : |a_{i,t}| \ge 1/|E_\ell|\}$ and in Lemma 5 we showed that $\mathbb{P}(|\mathcal{S}_{i,\ell}| > L_\ell) \le 1/|E_\ell|$. Hence, for any $i \in [N]$, we have $|a_{i,t}| \le 1/|E_\ell|$ for all periods $\tau \in E_\ell/\mathcal{S}_{i,\ell}$, so

$$\begin{aligned}
&\sum_{\tau\in E_\ell} \mathbb{I}\left\{b_{i,\tau} - \langle\hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} \\
=\ &\left(\sum_{\tau\in E_\ell/\mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle\hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} + \sum_{\tau\in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle\hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\}\right) \\
&+ \left(\sum_{\tau\in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle\hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} - \sum_{\tau\in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle\hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\}\right).
\end{aligned} \tag{84}$$

Consider the sum in first the parenthesis of Equation (84) and note that $b_{i,\tau} = v_{i,\tau} - a_{i,\tau} = \langle\beta, x_\tau\rangle + \epsilon_{i,\tau} - a_{i,\tau}$. Since $|a_{i,\tau}| \le 1/|E_\ell|$ for any $i \in [N]$ and $\tau \in E_\ell/\mathcal{S}_{i,\ell}$,

$$\langle\beta, x_\tau\rangle + \epsilon_{i,\tau} - \frac{1}{|E_\ell|} \ \le\ b_{i,\tau} \ \le\ \langle\beta, x_\tau\rangle + \epsilon_{i,\tau} + \frac{1}{|E_\ell|}, \quad \forall \tau \in E_\ell/\mathcal{S}_{i,\ell}. \tag{85}$$

Now, assume that the event $\xi_{\ell+1} = \left\{\|\hat{\beta}_{\ell+1} - \beta\|_1 \le \delta_\ell/x_{\max}\right\}$ holds. Therefore, we can upper bound the sum in first the parenthesis of Equation (84) as

$$\sum_{\tau\in E_\ell/\mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle\hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} + \sum_{\tau\in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle\hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\}$$

$$
\begin{aligned}
&\le \sum_{\tau \in E_\ell / \mathcal{S}_{i,\ell}} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \langle \hat{\beta}_{\ell+1} - \beta, x_\tau\rangle + \frac{1}{|E_\ell|}\right\} + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \langle \hat{\beta}_{\ell+1} - \beta, x_\tau\rangle + \frac{1}{|E_\ell|}\right\} \\
&= \sum_{\tau \in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \langle \hat{\beta}_{\ell+1} - \beta, x_\tau\rangle + \frac{1}{|E_\ell|}\right\} \\
&\le \sum_{\tau \in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \delta_\ell + \frac{1}{|E_\ell|}\right\}, \qquad (86)
\end{aligned}
$$

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 (85); and the final inequality is due to the occurrence of the event $\xi_{\ell+1} = \left\{\|\hat{\beta}_{\ell+1} - \beta\|_1 \le \delta_\ell / x_{\max}\right\}$. Similarly, we can also lower bound the sum in the first parenthesis of Equation (84):

$$
\sum_{\tau \in E_\ell / \mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle \hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle \hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} \;\ge\; \sum_{\tau \in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \le z - \delta_\ell - \frac{1}{|E_\ell|}\right\}. \qquad (87)
$$

Furthermore, assuming events $\mathcal{G}_{i,\ell} = \{|\mathcal{S}_{i,\ell}| \le L_\ell\}$ hold for all $i \in [N]$, we can simply upper bound and lower bound the expression in the second parenthesis of Equation (84):

$$
-L_\ell \;\le\; \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle \hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} - \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle \hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} \;\le\; L_\ell\,. \qquad (88)
$$

Combining Equations (84), (86), (87), (88), and using the definition $\hat{F}_{\ell+1}(z) = \frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{b_{i,\tau} - \langle \hat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\}$, under the occurrence of events $\xi_{\ell+1}$, and $\mathcal{G}_{i,\ell}$ for all $i \in [N]$, we have

$$
\begin{aligned}
&\frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau} \le z - \delta_\ell - \frac{1}{|E_\ell|}\right\} - \frac{L_\ell}{|E_\ell|} \;\le\; \hat{F}_{\ell+1}(z) \;\text{ and} \\
&\hat{F}_{\ell+1}(z) \;\le\; \frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau} \;\le\; z + \delta_\ell + \frac{1}{|E_\ell|}\right\} + \frac{L_\ell}{|E_\ell|}\,. \qquad (89)
\end{aligned}
$$

Now, for any $\gamma > 0$,

$$
\begin{aligned}
&\mathbb{P}\left(F\left(z - \delta_\ell - \frac{1}{|E_\ell|}\right) - \hat{F}_{\ell+1}(z) \le \gamma + \frac{L_\ell}{|E_\ell|}\right) \\
\ge\; &\mathbb{P}\left(\left\{F\left(z - \delta_\ell - \frac{1}{|E_\ell|}\right) - \hat{F}_{\ell+1}(z) \le \gamma + \frac{L_\ell}{|E_\ell|}\right\} \bigcap \xi_{\ell+1} \bigcap \left(\bigcap_{i\in[N]} \mathcal{G}_{i,\ell}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&\geq \mathbb{P}\left(\left\{F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right)-\frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\leq z-\delta_\ell-\frac{1}{|E_\ell|}\right\}\leq\gamma\right\}\bigcap\xi_{\ell+1}\bigcap\left(\bigcap_{i\in[N]}\mathcal{G}_{i,\ell}\right)\right)\\
&\geq \mathbb{P}\left(\left\{\sup_{\tilde z\in\mathbb{R}}\left|F(\tilde z)-\frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\leq\tilde z\right\}\right|\leq\gamma\right\}\bigcap\xi_{\ell+1}\bigcap\left(\bigcap_{i\in[N]}\mathcal{G}_{i,\ell}\right)\right)\\
&\geq 1-\mathbb{P}\left(\left\{\sup_{\tilde z\in\mathbb{R}}\left|F(\tilde z)-\frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\leq\tilde z\right\}\right|>\gamma\right\}\right)-\mathbb{P}\left(\xi_{\ell+1}\right)-\sum_{i\in[N]}\mathbb{P}\left(\mathcal{G}_{i,\ell}\right)\\
&\geq 1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\left(\frac{2d+N}{|E_\ell|}+d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\right)-\frac{N}{|E_\ell|}\\
&= 1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{2(d+N)}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right), \qquad (90)
\end{aligned}
$$

where the second inequality follows from Equation (89), the fourth inequality uses the union bound, and the final inequality follows from the DKW inequality (Theorem 11), Lemma 7, and Lemma 5. 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 1, $|F(z-\delta_\ell-1/|E_\ell|)-F(z)|\leq c_f(\delta_\ell+1/|E_\ell|)$ for $\forall z\in\mathbb{R}$. Hence, combining this with Equation (90), yields

$$
\begin{aligned}
&\mathbb{P}\left(F(z)-\hat F_{\ell+1}(z)\ \leq\ \gamma+c_f\left(\delta_\ell+\frac{1}{|E_\ell|}\right)+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &\mathbb{P}\left(F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right)-\hat F_{\ell+1}(z)\ \leq\ \gamma+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{2(d+N)}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right). \qquad (91)
\end{aligned}
$$

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

$$
\begin{aligned}
&\mathbb{P}\left(\hat F_{\ell+1}(z)-F(z)\ \leq\ \gamma+c_f\left(\delta_\ell+\frac{1}{|E_\ell|}\right)+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &\mathbb{P}\left(\hat F_{\ell+1}(z)-F\left(z+\delta_\ell+\frac{1}{|E_\ell|}\right)\ \leq\ \gamma+\frac{L_\ell}{|E_\ell|}\right)\\
\geq\ &1-2\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{2(d+N)}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right). \qquad (92)
\end{aligned}
$$

Combining Equations (91) and (92) using a union bound yields

$$
\begin{aligned}
&\mathbb{P}\left(\left|\hat F_{\ell+1}(z)-F(z)\right|\leq\gamma+c_f\delta_\ell+\frac{c_f+L_\ell}{|E_\ell|}\right)\\
\geq\ &1-4\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{4(d+N)}{|E_\ell|}-2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right). \qquad (93)
\end{aligned}
$$

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

$$\begin{aligned}
|\hat{F}_{\ell+1}^-(z) - F^-(z)| &= \left|N\hat{F}_{\ell+1}^{N-1}(z) - (N-1)\hat{F}_{\ell+1}^N(z) - \left(NF^{N-1}(z) - (N-1)F^N(z)\right)\right| \\
&\le N\left|\hat{F}_{\ell+1}^{N-1}(z) - F^{N-1}(z)\right| + (N-1)\left|\hat{F}_{\ell+1}^N(z) - F^N(z)\right| \\
&= N\left|\left(\hat{F}_{\ell+1}(z) - F(z)\right)\left(\sum_{n=1}^{N-1}\left(\hat{F}_{\ell+1}(z)\right)^{n-1}(F(z))^{N-1-n}\right)\right| \\
&\quad + (N-1)\left|\left(\hat{F}_{\ell+1}(z) - F(z)\right)\left(\sum_{n=1}^{N}\left(\hat{F}_{\ell+1}(z)\right)^{n-1}(F(z))^{N-n}\right)\right| \\
&\le N(N-1)\left|\hat{F}_{\ell+1}(z) - F(z)\right| + (N-1)N\left|\hat{F}_{\ell+1}(z) - F(z)\right| \\
&< 2N^2\left|\hat{F}_{\ell+1}(z) - F(z)\right|. \qquad (94)
\end{aligned}$$

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 $\hat{F}_{\ell+1}(z), F(z) \in [0,1]$ for $\forall z \in \mathbb{R}$. Combining Equations (93) and (94), we get

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

The probability bound for $\left|\hat{F}_{\ell+1}^-(z) - F^-(z)\right|$ can be shown in a similar fashion by noting that similar to Equation (94) we can show $|\hat{F}_{\ell+1}^+(z) - F^+(z)| < N\left|\hat{F}_{\ell+1}(z) - F(z)\right|$. □

Lemma 9 **(Bounding the Impact of Estimation Errors on Revenue)**. *We assume that the events* $\xi_{\ell+1} = \left\{\|\hat{\beta}_{\ell+1} - \beta\|_1 \le \frac{\delta_\ell}{x_{\max}}\right\}$, $\xi_{\ell+1}^- = \left\{\left|\hat{F}_{\ell+1}^-(z) - F^-(z)\right| \le 2N^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f+L_\ell}{|E_\ell|}\right)\right\}$ *and* $\xi_{\ell+1}^+ = \left\{\left|\hat{F}_{\ell+1}^+(z) - F^+(z)\right| \le N\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f+L_\ell}{|E_\ell|}\right)\right\}$ *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 (41). Hence for any* $r \in \{r_t^\star,\ r_t\}$ *where* $t \in E_{\ell+1}$ *we have the following:*

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

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

where $y_t = \langle\beta, x_t\rangle$, $\hat{y}_t = \langle\hat{\beta}_{\ell+1}, x_t\rangle$, $\hat{\beta}_{\ell+1}, \hat{F}_{\ell+1}^-, \hat{F}_{\ell+1}^+$ *are defined in Equations (12) and (14). The function* ρ_t *is defined in Equation (26).*

Proof of Lemma 9. **Part (i)** We consider the following:

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

The first equality follows from definition of ρ_t in Equation (26), and the second inequality applies the Lipschitz property of F^- and F^+ using Lemma 1. The third inequality follows from Cauchy's inequality: $|y_t - \hat{y}_t| = |\langle \hat{\beta}_{\ell+1} - \beta, x_t\rangle| \leq \|\hat{\beta}_{\ell+1} - \beta\|_1 x_{\max}$, and the last inequality follows from the occurrence of $\xi_{\ell+1}$ and $N \geq 1$.

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

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

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

LEMMA 10 **(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 (40), (42), and (43) respectively.

Proof of Lemma 10. 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 $\xi_{\ell+1} = \left\{\|\hat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}}\right\}$. Then,

$$\begin{aligned}\mathbb{P}\left(\xi^c_{\ell+1}\right) &\leq \frac{2d+N}{|E_\ell|} + d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\\ &\leq \frac{2d+N}{|E_\ell|} + d\exp\left(-\frac{\log(|E_\ell|)T^{\frac{1}{4}}\lambda_0^2}{8x_{\max}^2}\right)\\ &\leq \frac{N+3d}{|E_\ell|}, \qquad (95)\end{aligned}$$

where the first inequality follows from Lemma 7 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 $|E_\ell| \geq |E_1| = \sqrt{T}$, $T \geq \max\left\{\left(\frac{8x_{\max}^2}{\lambda_0^2}\right)^4, 9\right\}$, 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 $\left(\xi^-_{\ell+1}\right)^c$:

$$\begin{aligned}\mathbb{P}\left(\left(\xi^-_{\ell+1}\right)^c\right) &\leq 4\exp\left(-2N|E_\ell|\cdot\left(\frac{\sqrt{\log(|E_\ell|)}}{\sqrt{2N|E_\ell|}}\right)^2\right) + \frac{4(d+N)}{|E_\ell|} + 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\\ &\leq \frac{2(2N+3d+2)}{|E_\ell|}, \qquad (96)\end{aligned}$$

where the first inequality follows from Lemma 8 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 $T \geq \max\left\{\left(\frac{8x_{\max}^2}{\lambda_0^2}\right)^4, 9\right\}$.

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

$$\mathbb{P}\left(\left(\xi^+_{\ell+1}\right)^c\right) \leq \frac{2(2N+3d+2)}{|E_\ell|}, \qquad (97)$$

Finally, combining Equations (95), (96) and (97), we have

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

□

11. Supplementary Lemmas

LEMMA 11 **(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*

$$\hat{F}(z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{Z_i \le z\} \quad , z \in \mathbb{R}. \tag{98}$$

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

$$\mathbb{P}\left(\sup_{z\in\mathbb{R}} \left|\hat{F}(z) - F(z)\right| \le \bar{\gamma}\right) \ge 1 - 2\exp\left(-2n\bar{\gamma}^2\right). \tag{99}$$

LEMMA 12 **(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 $0 \le \lambda_{\min}(Z_k)$ and $\lambda_{\max}(Z_k) \le B$ 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 $\forall \bar{\gamma} \in (0,1)$,*

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

LEMMA 13 **(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 $\tau \in [\widetilde{T}]$, $|z_{1,\tau} - z_{2,\tau}| \le 1$ and $\mathbb{E}\left[(z_{1,\tau} - z_{2,\tau}) \mid \sum_{s<\tau} z_{1,s}, \sum_{s<\tau} z_{2,s}\right] \le 0$. Let $\tilde{\gamma} \in [0,1]$ and $A \in \mathbb{R}$. Then,*

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