
Comparing Uniform Price and Discriminatory Multi-Unit Auctions through Regret Minimization

Anonymous Author(s)

Affiliation

Address

email

1 Repeated multi-unit auctions, where a seller allocates multiple identical items over many rounds, are
2 common mechanisms in electricity markets and treasury auctions. We compare the two predominant
3 formats: uniform-price and discriminatory auctions, focusing on the perspective of a single bidder
4 learning to bid against stochastic adversaries. We characterize the learning difficulty in each format,
5 showing that the regret scales similarly for both auction formats under both full-information and
6 bandit feedback, as $\tilde{O}(\sqrt{T})$ and $\tilde{O}(T^{2/3})$, respectively. However, analysis beyond worst-case regret
7 reveals structural differences: uniform-price auctions may admit faster learning rates, with regret
8 scaling as $\tilde{O}(\sqrt{T})$ in settings where discriminatory auctions remain at $\tilde{O}(T^{2/3})$. Finally, we provide
9 a specific analysis for auctions in which the other participants are symmetric and have unit-demand,
10 and show that in these instances a similar regret rate separation appears.

11 1 Introduction

12 Uniform price auctions and discriminatory auctions are widely used market mechanisms for the
13 allocation of resources. They have been widely implemented and studied for electricity markets
14 (Elmaghraby & Oren, 1999; Hortaçsu & Puller, 2008) and treasury auctions (Nyborg et al., 2002).
15 They are the most widespread simultaneous multi-unit auctions of identical goods (Krishna, 2009).
16 Uniform price and discriminatory auctions are both standard mechanisms that allocate items to
17 buyers willing to pay the most; their rules are identical except for the prices buyers pay. Since the two
18 mechanisms are so similar, they can easily replace each other and have been frequently compared
19 (De Keijzer et al., 2013). Their differences have been studied from several points of view: from their
20 efficiency in Ausubel et al., 2014, the revenue they generate for the auctioneer, to the expected social
21 welfare in different types of equilibria (Syrkanis & Tardos, 2013).

22
23 Participating in these auctions can be challenging from a bidder’s standpoint since the bidding
24 process is not truthful (i.e., the best strategy depends on other buyers’ behavior). However, when
25 auctions occur repeatedly, buyers can improve their strategies over time using online learning
26 techniques; this is referred to as *learning to bid* (Feng et al., 2018). The performance of these online
27 learning strategies is often measured by regret, which compares the cumulative rewards with that of
28 the best fixed strategy in hindsight (Lattimore & Szepesvári, 2020). We propose to leverage this
29 online learning framework to quantify the difficulty of learning to bid optimally, therefore providing
30 a new approach to compare uniform and discriminatory price auctions.

31
32 Learning to bid in repeated multiunit auctions has recently been studied, both in the uniform pricing
33 case Brânzei et al., 2023 and Potfer et al., 2024, and in the discriminatory pricing case by Galgana
34 and Golrezaei, 2025. These works focus on adversarial bids and provide regret rates with similar
35 dependency on the time horizon T of $\tilde{O}(\sqrt{T})$ under full-information feedback and $\tilde{O}(T^{2/3})$ under
36 bandit feedback. These rates have been proven to be tight, except for the usual uniform price auction
37 (referred to as the First Rejected Bid in Potfer et al., 2024).

We study the learning-to-bid problem in the case where opposing bids are stochastic. The stochasticity assumption ensures that our analysis characterizes the inherent difficulty of the auction format, rather than the difficulty arising from facing strategic or adversarial agents. We characterize when both problems can be learned with similar regret rates and when the rates differ. In the process, we provide the first worst-case tight lower bounds for uniform price auctions with bandit feedback, as well as the first instance-dependent regret bounds. Finally, we show that, against symmetric unit-demand adversaries, a clear separation of achievable regret appears, and we provide an efficient algorithm for the uniform auction that leverages the specific structure of this setting.

1.1 Related Work

Simultaneous auctions of multiple identical items, including uniform, discriminatory, and Vickrey-Clarke-Groves pricing, are standard in auction theory, as described in Krishna, 2009. Uniform and discriminatory auctions have been studied and compared from an empirical point of view (Nyborg & Sundaresan, 1996; Nyborg et al., 2002), and by theoretical approaches, focusing on the revenue they generate (Ausubel et al., 2014) as well as the social welfare they generate (De Keijzer et al., 2013; Syrgkanis & Tardos, 2013). Special cases involving symmetry and unit-demand participant are also of particular interest (Anderson & Holmberg, 2023).

The repeated setting of auctions has recently attracted some focus, as covered by Nedelec et al., 2022. This setting allows for leveraging online and statistical learning tools (Lattimore & Szepesvári, 2020) to explore dynamic strategies. These repeated settings were first studied with a focus on the auctioneer, enabling the learning of reserve prices as in the work of Mohri and Medina, 2014. Learning to bid, from the bidder’s perspective, was introduced later and studied in several settings, including sealed-bid first-price and second-price auctions (Achddou et al., 2021b; Balseiro et al., 2019; Weed et al., 2016).

Learning to bid in multi-unit auctions has only recently started to be studied, except for a partial result for uniform auctions by Feng et al., 2018 used as an example. The discriminatory pricing auction was studied by Galgana and Golrezaei, 2025, who proved regret rates of $\tilde{O}(K\sqrt{T})$ and $\tilde{O}(KT^{2/3})$ under full-information and bandit feedback, respectively. Regret bounds of $\tilde{O}(K\sqrt{T})$ were also obtained for the uniform-price format in the full-information case by Brânzei et al., 2023, as well as sub-optimal rates in the bandit case. Potter et al., 2024 provided algorithms with a $\tilde{O}(K^{4/3}T^{2/3})$ regret upper bound for uniform auction in bandit feedback and showed that it is tight for a less common variant of uniform pricing (Last Accepted Bid pricing, or LAB). They provide no matching lower bound for the usual uniform price auction (First Rejected Bid, or FRB). Golrezaei and Sahoo, 2024 has also studied uniform pricing in a repeated setting, but focuses on bidders with return-over-investment constraints; they obtain similar regret bounds of $\tilde{O}(K^{5/3}T^{2/3})$ in their setting.

1.2 Contribution

We propose a regret-based comparison of repeated uniform and discriminatory multi-unit auctions with stochastic opposing bids. For this comparison, we provide the first analysis of learning to bid in repeated multi-unit auctions when facing stochastic bids for both auction types.

We show that both auction formats admit tight worst-case regret rates of $\tilde{O}(K\sqrt{T})$ under the full-information setting (improving the best known rates for uniform price auction by a factor of \sqrt{K}). Our analysis also yields regrets rates $\tilde{O}(K^{5/3}T^{2/3})$ in the bandit feedback for uniform auction, and we show a matching lower bound. In the bandit setting, some differences arise as the uniform auction exhibits better regret rates when the bidder has low demand: 0 regret for unit-demand and $\tilde{O}(\sqrt{T})$ for two-unit-demand.

We provide an instance-dependent quantity characterizing instances for which regret bounds scaling as $\mathcal{O}(\sqrt{T})$ can be obtained for the uniform auction under bandit feedback. The discriminatory auction’s lower bound of $\Omega(T^{2/3})$ remains valid in these instances.

Finally, in the bandit feedback when the adversary are unit-demand and symmetric we show a similar regret separation providing an algorithm which guarantees regret of $\tilde{O}(\sqrt{T})$ in uniform auctions while lower bound of $\Omega(T^{2/3})$ for the discriminatory price auctions remains valid. In the analysis, we derive a concentration inequality for partially observed ordered statistics based on the DKW inequality, which may be of independent interest.

The following table summarizes the achievable regret rates for the different auction formats under bandit feedback considered in this work, we use the notation $\tilde{\Theta}(\cdot)$ when we have matching lower bounds $\Omega(\cdot)$ and upper bounds $\mathcal{O}(\cdot)$, up to log factors. Results that are entirely novel are highlighted in green, while blue denotes marginal improvements or generalizations over existing work.

	Worst Case			Instance Dependent	
	Unit Demand	Two-unit Demand	Any Demand	Δ - sep 4	I.I.D.
Discriminatory	$\tilde{\Theta}(T^{2/3})$			$\tilde{\Theta}(T^{2/3})$	
Uniform	0	$\tilde{\Theta}(\sqrt{T})$	$\tilde{\Theta}(T^{2/3})$	$\tilde{\Theta}(\sqrt{T})$	

Table 1: Achievable regret under bandit feedback.

93

94 2 Models

95 The following encompasses both uniform and discriminatory price multi-unit auctions, and matches those used in Brânzei et al., 2023 and Galgana and Golrezaei, 2025. Indeed, because both auction formats are very similar, we only need to specify different payment rules (equations (2) and (3)) to describe the mechanisms.

96 **Repeated auction** Let $(K, T) \in \mathbb{N}^2$ be, respectively, the number of items sold at each auction and the number of repeated auctions. We describe the repeated K -item uniform and discriminatory price auctions from the point of view of one buyer, that we call the *bidder*, facing other stochastic buyers, whose bids are aggregated and who, as a group, are called the *adversary*. The bidder has a fixed set of valuations that quantifies how much utility it derives from being allocated any number of items. We denote by $(v_1, \dots, v_K) \in [0, 1]^K$ the bidder's marginal valuations. These are marginal in the sense that when the bidder obtains n items, it derives $\sum_{l \leq n} v_l$ from it.

106 At each time-step $t \in [T]$, an instance of the auction proceeds as follows:

- 107 1. The bidder transmits its bids $\mathbf{b}^t := \{b_1^t, b_2^t, \dots, b_K^t\}$ to the auctioneer. For convenience, we consider \mathbf{b} ordered in non-increasing order and the bids in $[0, 1]^K$. We call B the corresponding subset of $[0, 1]^K$.
- 108 2. The bids of the adversary $\beta^t := \{\beta_1^t, \beta_2^t, \dots, \beta_K^t\} \in B$ are drawn according to a distribution over B , denoted \mathcal{D} , and then transmitted to the auctioneer.
- 109 3. The auctioneer orders all the bids in non-increasing order and determines the *allocation* (the number of items won) of the bidder that we denote $x(\mathbf{b}^t, \beta^t) \in [K]$ as well as the prices $p(\mathbf{b}^t, \beta^t) \in [0, 1]^K$.
- 110 4. The bidder receives their allocated items and pays the price for each item. This gives rise to the following utility: $u(\mathbf{b}^t, \beta^t) := \sum_{l=1}^{x(\mathbf{b}^t, \beta^t)} [v_l - p(\mathbf{b}^t, \beta^t)(l)]$
- 111 5. The bidder may receive additional information, referred to as the feedback.

112 **Allocation** Both auctions we study are standard; therefore, the items are allocated to the buyers who submitted the K highest bids. Thus, the allocation is formulated as follows¹:

$$x(\mathbf{b}, \beta) := \max\{k \in [K] \mid \forall j \leq k, b_j \geq \beta_{K+1-j}\} \quad (1)$$

120 **Pricing** As we noted previously, the pricing rules of the uniform and discriminatory auctions are different. We describe below how the price is set in both auctions:

- 121 • The uniform price auction sets the price of all allocated units identically: the value of the first rejected bid. Formally, it is the $(K + 1)^{\text{th}}$ largest element of (\mathbf{b}, β) , written:

$$p(\mathbf{b}, \beta) := (\max(b_{x(\mathbf{b}, \beta)+1}, \beta_{K-x(\mathbf{b}, \beta)+1}))_{k \in [K]} \quad (2)$$

- 122 • The discriminatory price auction sets the price of each allocated unit as the value of the corresponding bid (the bid that won its emitter this item). Formally, the discriminatory price is written as follows:

$$p(\mathbf{b}, \beta) := (b_k)_{k \in [K]} \quad (3)$$

¹This allocation breaks ties in favor of the bidder. The results can be extended to any tie-breaking rule by ensuring that ties occur with probability zero.

127 **Regret** The performance of a learning algorithm is quantified by the regret. This metric quantifies
 128 the expected cumulative difference between the maximum utility the bidder could have obtained and
 129 the utility they actually received. We define it as follows:

$$R_T = T \sup_{\mathbf{b} \in B} \left(\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] \right) - \sum_{t=1}^T \mathbb{E}_{\beta^t \sim \mathcal{D}} [u(\mathbf{b}^t, \beta^t)] \quad (4)$$

130 This definition aligns with what is usually defined as pseudo-regret. Note that, with the previous
 131 definition, minimizing the regret is equivalent to maximizing the cumulative expected utility. Using
 132 this metric allows us to take into account both how well and quickly a bidder can approximate the
 133 best bid, as well as how *costly* it is to acquire information about \mathcal{D} .

134 **Feedback** The bidder can sequentially improve its bids since it receives some information after each
 135 auction, which intuitively provides it with some knowledge about the distribution \mathcal{D} . We describe
 136 below the *feedback*, defined as the information received after one instance of the auction. We focus on
 137 two feedback types: the full-information and bandit feedback setting, which are the most common in
 138 online learning literature (Lattimore & Szepesvári, 2020) and in online learning in auctions (Achddou
 139 et al., 2021a; Feng et al., 2018). At time $t \in [T]$, at the end of the auction:

- 140 • in the *full-information* feedback setting, the bidder observes β^t ;
- 141 • in the *bandit* feedback setting, the bidder observes only $x(\mathbf{b}^t, \beta^t)$ and $p(\mathbf{b}^t, \beta^t)$.

142 Note that the full-information feedback is strictly more informative than the bandit feedback, because
 143 $x(\mathbf{b}^t, \beta^t)$ and $p(\mathbf{b}^t, \beta^t)$ can be computed by the bidder who knows both its bids \mathbf{b}^t and the auction's
 144 rules.

145 3 Learning with full-information feedback

146 We first focus on the full information feedback setting where the learner observes the full vector β^t
 147 at each time step. The bidder aims to minimize regret by choosing bids as close as possible to the
 148 optimal bid vector, namely $\mathbf{b}^* := \operatorname{argmax}_{\mathbf{b} \in B} \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)]$. Because the auction mechanisms
 149 are not truthful, \mathbf{b}^* depends on \mathcal{D} . To describe this dependence, let us define, for all $k \in [K]$, and
 150 $x \in [0, 1]$, $F_k(x) := \mathbb{P}_{\beta \sim \mathcal{D}}(\beta_k \leq x)$ the k^{th} marginal cumulative distribution function.

151 The following lemma provides the main intuition for our learning algorithms, which mainly focus on
 152 estimating the marginal CDFs rather than the optimal bid vector \mathbf{b}^* .

153 **Lemma 1.** *In both auction formats, the expected utility can be expressed as a function of the bidder's*
 154 *bid vector \mathbf{b} and the marginal cumulative distribution functions $(F_k)_{k \in [K]}$.*

155 The proof follows from writing the utility and taking the expectation, it is provided in Appendix A. We
 156 denote U_d the functions such that, in the discriminatory auction, for all $\mathbf{b} \in B$, $U_d((F_k)_{k \in [K]}, \mathbf{b}) =$
 157 $\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)]$, and U_u the equivalent in the uniform auction. Explicit formulas are provided in the
 158 appendix (see (14) and (15)) to save space.

159 3.1 Algorithm

160 At time $t \in [T]$, the bidder has observed $\beta^1, \dots, \beta^{t-1}$, which are i.i.d. samples from \mathcal{D} . These are
 161 sufficient to compute empirical marginal CDFs, which are known to have strong convergence proper-
 162 ties Massart, 1990. They are defined, for $k \in [K]$, and $x \in [0, 1]$ as $\hat{F}_k^t(x) := \frac{1}{t-1} \sum_{j=1}^{t-1} \mathbb{1}\{\beta_k^j \leq x\}$.
 163 Depending on the auction, we define the expected utility estimate as either $\hat{u}^t(\mathbf{b}) := U_u((\hat{F}_k^t), \mathbf{b})$ or
 164 $\hat{u}^t(\mathbf{b}) := U_d((\hat{F}_k^t), \mathbf{b})$. Equipped with these, we can now provide an algorithm that guarantees tight
 165 regret bounds for both auction formats.

Algorithm 1: Full-information feedback

- 1 **Input:** time horizon T
 - 2 **Output:** bids for each time step $(\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^{T-1}, \mathbf{b}^T) \in (B)^T$.
 - 3 **for** $t = 1, 2, \dots, T$ **do**
 - 4 Play $\mathbf{b}^t := \operatorname{argmax}_{\mathbf{b} \in B} \hat{u}^t(\mathbf{b})$.
 - 5 Receive the utility $u(\mathbf{b}^t, \beta^t)$ and observe β^t .
-

166 **Theorem 1.** *Under full-information feedback, Algorithm 1 achieves a regret $R_T = \tilde{O}(K\sqrt{T})$ for*
 167 *both the discriminatory and the uniform price auction.*

168 *Proof Sketch.* The proof for the discriminatory auction uses the Dvoretzky–Kiefer–Wolfowitz inequality
 169 from Massart, 1990 to build high-probability confidence bands for \hat{F}_k , which yields confidence
 170 bands for the utility estimate. Combining that with the optimality of \mathbf{b}^t with respect to \hat{u}^t allows
 171 us to show that the per-round regret is upper bounded by $2K\sqrt{\ln(2/\alpha)/2t}$, therefore providing the
 172 correct regret rates when summing.

173 The proof for the uniform auction’s is very similar. The main difference appears when ensuring error
 174 bands scales as K and not K^2 , this requires our concentration bands to leverage the strong negative
 175 correlation between the different terms estimated in expected utility. This is achieved by using the
 176 Natarajan dimension, a multi-class classifier generalization of VC dimension Shalev-Shwartz and
 177 Ben-David, 2014. \square

178 The complete proof is in Appendix B. These regret bounds are tight, as it is known that any algorithm
 179 must incur a regret growing as $\Omega(K\sqrt{T})$ in both auction format (Brânzei et al., 2023; Galgana &
 180 Golrezaei, 2025).

181 **Remark 1.** *Note that, because \hat{u}^t is always a piece-wise linear function and is separable into*
 182 *functions that only depend on two components of \mathbf{b} , a maximizer can be found by using dynamic*
 183 *programming in $\mathcal{O}(K^3t^2)$.*

184 4 Learning with bandit feedback

185 We now focus on learning in the bandit feedback setting. Let us recall that in this setting, at time
 186 t , the bidder only observes his allocation $x(\mathbf{b}^t, \beta^t)$ and the price paid per unit $p(\mathbf{b}^t, \beta^t)$. Since the
 187 allocation function is the same across both auction’s type (1), the observed allocation provides the
 188 same information :

$$(\mathbb{1}\{b_i^t \geq \beta_{K-i+1}^t\})_{i \in [K]} \quad (5)$$

189 However, the information conveyed by the price differs:

- 190 • The discriminatory price depends only on the bidder’s own bid and therefore provides no
 191 additional information.
- 192 • The uniform price can be set by an opposing bid β_{K-i+1} , potentially revealing information
 193 about the distribution \mathcal{D} .

194 The following lemma formalizes the feedback the bidder receives in the uniform price auction.

195 **Lemma 2.** *At time $t \in [T]$, let $\mathbf{b}^t, \beta^t \in B$ be the bids of the learner and the adversary, in the*
 196 *uniform auction with bandit feedback the bidder observes $(\mathbb{1}\{\beta_{K-i+1}^t \in (b_{i+1}^t, b_i^t]\})_{i \in [K]}$.*

197 This feedback reveals components of β^t if they fall into the right intervals, providing richer feedback
 198 than in the discriminatory case. We now examine achievable regret rates in both auction types and the
 199 effects of this feedback discrepancy.

200 4.1 Discriminatory price auction

201 The discriminatory price auction with fixed valuation has been studied by Galgana and Golrezaei,
 202 2025 who provided algorithms guaranteeing regret upper bounds of $\mathcal{O}(KT^{2/3})$. They also provided
 203 a regret lower bound of $\Omega(K^{2/3}T^{2/3})$ for discretized bid strategies. We provide a strengthening of
 204 this result with a lower bound that holds for *any* algorithm, including those operating over continuous
 205 bid spaces.

206 **Lemma 3.** *Any algorithms for bidding in repeated first price auctions with known valuation must*
 207 *incur a regret of $\Omega(T^{2/3})$*

208 Given this result and the almost matching upper bound, we believe the achievable regret rates are well
 209 characterized. In the following we characterize what can be achieved in the uniform price auction,
 210 and when the richer feedback allows for better rates than in discriminatory auctions.

4.2 Uniform price auction

The additional information provided by the richer bandit feedback in the uniform price auction can be used to design learning algorithms. We illustrate how the feedback allows to estimate the marginal CDFs F_k with Algorithm 2. During its estimation phase, it leverages the feedback detailed in Lemma 2 to ensure observation of each β_k in a round robin fashion. Once the estimated CDFs are precise enough, an estimated best bid $\mathbf{b}^{T_{expl}}$ is computed and then played repeatedly during the exploitation phase.

We define the empirical CDFs for the uniform price auction with bandit feedback as follows for $k \in [K]$:

$$\forall x \in [0, 1], \tilde{F}_k(x) := \frac{\sum_{j=1}^t \mathbb{1}\{x \in (b_{K+2-k}^t, b_{K+1-k}^t]\} \mathbb{1}\{\beta_k^t \leq x\}}{\sum_{j=1}^t \mathbb{1}\{x \in (b_{K+2-k}^t, b_{K+1-k}^t]\}} \quad (6)$$

We also define the empirical expected utility \tilde{u} , for $\mathbf{b} \in B$ as $U_u((\tilde{F})_{k \in [K]}, \mathbf{b})$. With these we can provide our algorithms and the associated regret guarantees.

Algorithm 2: Estimate then commit for uniform price auction

- 1 **Input:** time horizon T , T_{expl} duration of the exploration.
 - 2 **Output:** bids for each time step $(\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^{T-1}, \mathbf{b}^T) \in B^T$.
 - 3 **Estimate:** for $t = 1, 2, \dots, T_{expl}$ **do**
 - 4 $k \leftarrow T - K \lfloor T/K \rfloor + 1$.
 - 5 Play \mathbf{b}^t such that $\forall i \leq k, b_i^t = 1$ and $\forall i > k, b_i^t = 0$.
 - 6 Receive the utility $u^t = u(\mathbf{b}^t, \beta^t)$ and the feedback.
 - 7 **Commit:** for $t = T_{expl} + 1, \dots, T$ **do**
 - 8 Play $\mathbf{b}^t = \arg\max_{\mathbf{b} \in B} \hat{u}^{T_{expl}}(\mathbf{b})$
-

Theorem 2. *The estimate then commit algorithms, with the exploration time $T_{expl} = K^{2/3}T^{2/3}$ achieves a regret bound of $\tilde{O}(K^{5/3}T^{2/3})$*

Proof Sketch. The proof separates the analysis of the regret incurred during the exploration phase and the commitment. During the exploration phase it is no bigger than $K^{5/3}T^{2/3}$. The regret of the exploitation phase can be upper bounded by $K^{5/3}T^{2/3}$; obtaining this bound requires the use of variants of the DKW concentration inequality, which exhibit local dependencies as in Bartl and Mendelson, 2023; Blanchard and Voracek, 2024. A simpler analysis, using DKW inequalities (Massart, 1990) yields a looser bound of $K^2T^{2/3}$ when setting $T_{expl} = KT^{2/3}$. \square

These regret guarantees match the rates in T of the ones for the discriminatory price auction. Since the regret rates in the discriminatory auction are tight in T , it is natural to ask whether that is also true for the uniform price auction. Theorem 3 provides a matching lower bound, proving that a regret that grows as $T^{2/3}$ is optimal.

Theorem 3. *In the uniform price auction, any algorithms must incur a regret of $\Omega(T^{2/3})$ for learning to bid with known valuations, when valuations are such that $v_3 > 0$.*

Proof idea. The proof of this theorems relies on standard approach (Cesa-Bianchi et al., 2024; Kleinberg & Leighton, 2003) to create regret lower bounds in bandit setting with continuous action space by embedding into our auction model a hard instance which is akin to a multiarmed bandit problem with $\mathcal{O}(T^{1/3})$ arms. The main challenge is that the bandit feedback in the uniform price auction is relatively rich, as pointed out in Lemma 2. We overcome this difficulty with the two following techniques, for a specific $i \in [K]$:

- the optimal bids are such that $b_i^* = b_{i+1}^*$ while the bidder needs to observe β_{K-i+1} ,
- when submitting a bid with $b_i \neq b_{i+1}$, the bidder suffers an additional regret term which scales linearly with $b_i - b_{i+1}$.

\square

246 We have shown that the achievable regret in the two auction formats behaves similarly regardless of
 247 the feedback types. The remainder of the paper is dedicated to showing that, beyond worst cases, the
 248 uniform auction can be easier to learn (regret scaling with \sqrt{T}) across different families of instances
 249 for which the discriminatory auctions remain equally hard (regret scaling as $T^{2/3}$). We provide
 250 general algorithms that exhibit better regret scaling when possible without information about the
 251 instance and then focus on an I.I.D setting, whose structure is especially interesting.

252 5 Bandit feedback, beyond worst-case

253 5.1 Instance dependant regret

254 To exhibit that the regret scales as \sqrt{T} on family of instances that are well-suited for the uniform price
 255 auction, we need an algorithm which adapts to such instances. This adaptability is key in guaranteing
 256 worst-case regret, as in Theorem 2, while also leveraging easy instances when presented with one.
 257 Algorithm 2 lacks such flexibility, because of both its fixed-length estimation and commitment phases,
 258 as well as its fixed marginal CDF estimation intervals (namely $[0, 1]$).

259 We propose two natural improvements: (i) using sequentially shrinking estimation intervals and
 260 (ii) using distinct intervals for each marginal CDF. Intuitively, the shrinking rule should allow us
 261 to estimate CDFs only However, the choice of the shrinking rule must be made with care. We
 262 propose a successive elimination approach (Audibert & Bubeck, 2010) which iteratively reduces a
 263 set of candidate bids $B^t \subset B$ based on the current utility estimates and their uncertainty. As the
 264 set of relevant bids B^t gets smaller, the intervals where estimates of the marginal CDFs are useful
 265 to estimate the utility do as well; this yields our shrinking rules for the intervals, that we denote
 266 $(\mathcal{I}_k^t)_{k \in [K]}$.

Algorithm 3: Bandit Feedback

Input: Time horizon T
Initialize: For each $i \in [K]$, initialize $\mathcal{I}_i^0 \leftarrow [0, 1]$
Output: Bids $(\mathbf{b}^1, \dots, \mathbf{b}^T) \in B^T$, with $\Delta^0 \leftarrow -1$

```

1 for  $t = 1, \dots, T$  do
2   if  $t \leq T_{\text{expl}}$  and  $\Delta^{t-1} \leq 0$  then
3      $k \leftarrow T - K \lfloor T/K \rfloor + 1$ ;
4      $b_k^t \leftarrow \max\{b_k \in \mathcal{I}_k^{t-1}\}$ ,  $b_{k+1}^t \leftarrow \min\{b_{k+1} \in \mathcal{I}_{k+1}^{t-1}\}$ ;
5      $\mathbf{b}^t \leftarrow \arg \max_{b_k=b_k^t, b_{k+1}=b_{k+1}^t} \tilde{u}^t(\mathbf{b})$ ;
267   else
6      $\tilde{k} \leftarrow T - 2(K-1) \lfloor T/2(K-1) \rfloor + 2$ ,  $k \leftarrow \lfloor \tilde{k}/2 \rfloor$ ;
7      $b_k^t \leftarrow \begin{cases} \max\{b_k \in \mathcal{I}_k^{t-1}\} & \text{if } k' \text{ is even} \\ \min\{b_k \in \mathcal{I}_k^{t-1}\} & \text{if } k' \text{ is odd} \end{cases}$ ;
8      $\mathbf{b}^t \leftarrow \arg \max_{b_k=b_k^t} \tilde{u}^t(\mathbf{b})$ ;
9   Play  $\mathbf{b}^t$  and receive feedback;
10  Update  $(\mathcal{I}_k^t)_{k \in [K]}$ ,  $\Delta^t$  using Algorithm 4;
11
```

Algorithm 4: Interval Refinement

Input: Time t , horizon T , intervals $(\mathcal{I}_k^{t-1})_{k \in [K]}$, utility estimate \tilde{u}^t
Output: Updated intervals $(\mathcal{I}_k^t)_{k \in [K]}$, gap Δ^t

```

268 1  $u_{\max}^t \leftarrow \max_{\mathbf{b} \in \prod_{k \in [K]} \mathcal{I}_k^{t-1}} \tilde{u}^t(\mathbf{b})$ ; // Max observed utility
2  $\prod_{k \in [K]} \mathcal{I}_k^t \leftarrow \text{conv} \left( U^{-1} \left( \left[ u_{\max}^t - \sqrt{\frac{\ln(2T^2)}{2 \lfloor t/K \rfloor}}, u_{\max}^t \right] \right) \right)$  //  $U^{-1}$  denote the preimage of  $\tilde{u}^t$ 
3  $\Delta^t \leftarrow \min_{k \in \{2, \dots, K\}} (\min \mathcal{I}_k^t - \max \mathcal{I}_{k+1}^t)$ ;

```

269 Let us denote $B^* \subset B$ the set of maximizers of the expected utility and the optimal intervals
 270 $\mathcal{I}_k^* \subset \mathbf{R}$ the smallest closed intervals such that $B^* \subset \prod_{k=1}^K \mathcal{I}_k^*$. We introduce the interval gap

271 $\Delta := \min_{k \in \{2, \dots, K\}} (\min \mathcal{I}_k^* - \max \mathcal{I}_{k+1}^*)$, the following theorem provides instance dependant
 272 bounds, valid when $\Delta > 0$ ($\mathcal{O}(\cdot)$ hides terms constant in T).

273 **Theorem 4.** *Algorithm 3 guarantee a regret of $\tilde{\mathcal{O}}(K^{5/3}T^{2/3})$ in general and, when $\Delta > 0$, an*
 274 *instance-dependent regret bounds of $\tilde{\mathcal{O}}(K\sqrt{T})$.*

275 *Proof.* Worst-case regret bounds are essentially proved as in Theorem 2. We only need to check
 276 that our shrinking rule is well-behaved. Noticing that with high probability, the intervals \mathcal{I}_i^t always
 277 contain the best-response set B^* suffices.

278 Instance-dependent bounds proofs rely on the fact that as soon as intervals become disjoint (ie
 279 $\Delta^t > 0$), the observation of Lemma 2 covers the whole intervals, which is why we recover full-
 280 information type bounds. \square

281 5.2 Regret separation

282 We illustrate the implications of Theorem 4 by describing instance families for which the achievable
 283 regret rates in uniform and discriminatory price auctions diverge significantly. This highlights how
 284 the choice of auction impacts learning complexity.

285 A straightforward example is when the bidder has *unit demand*, meaning it only values a single unit
 286 ($v_1 > 0$ and $v_i = 0$ for $i > 1$). In this case, the uniform auction is truthful (Lemma 7), so the bidder
 287 can bid truthfully and incur zero regret. In contrast, the discriminatory auction still requires strategic
 288 bidding, and the lower bound from Lemma 3 applies, resulting in a regret of $\Omega(T^{2/3})$.

289 **Two unit demand** A similar but weaker regret rate discrepancy also appears in the two-unit demand
 290 setting (i.e. $v_1 > 0, v_2 \geq 0$ and $v_i = 0$ for $i > 2$). In that setting, both auctions require strategic
 291 bidding; in discriminatory auction, Lemma 3 still applies resulting in regret $\Omega(T^{2/3})$ while Theorem 4
 292 yields a regret of $\mathcal{O}(\sqrt{T})$ on the regret in the uniform auction.

293 **Δ separated distributions** We finally consider a more general condition on the distribution of
 294 adversary bids. Let \mathcal{D} be a distribution over B such that when $\beta \sim \mathcal{D}$ each coordinate β_k almost
 295 surely lies in disjoint intervals \mathcal{I}_k , and adjacent intervals are at least separated by a gap Δ . We refer
 296 to these as Δ separated. To rule out degenerate cases, we focus on $\Delta < \frac{1}{2K}$:

297 **Lemma 4.** *Learning in a discriminatory or uniform auction when the adversary's distribution is Δ*
 298 *separated, with $\Delta < \frac{1}{2K}$ can be achieved with respectively regret of $\Omega(T^{2/3})$ and $\mathcal{O}(\sqrt{T})$.*

299 *Proof sketch.* The lower bound from Lemma 3 for the discriminatory auction can be easily extended
 300 by leveraging the fact that $\Delta > \frac{1-v_1}{K}$. Proving the upper bound is straightforward by noticing the
 301 following : (i) there is always an optimal bids against Δ -separated distribution whose component are
 302 only 0s and 1s, (ii) the \bar{F} of Δ -separated distribution are Δ -separated. \square

303 5.3 I.I.D. adversaries

304 This section characterizes achievable regret when the bidder faces $N \in \mathbb{N}$ symmetric, unit-demand
 305 participants. Each participant submits a single bid, so the opposing bid vector β consists of the K
 306 highest bids among them. By symmetry, all bids are i.i.d. from a distribution \mathcal{P} , and we assume
 307 independence in their bidding. We denote by \mathcal{P}_K^N the induced distribution on β .

308 **Opposing bids** The first K order statistics of N i.i.d. samples from \mathcal{P} constitute the opposing bids.
 309 This structure provides additional information on the relationships between marginal CDFs (Casella
 310 & Berger, 2024). Letting $F : [0, 1] \rightarrow [0, 1]$ denote the CDF of \mathcal{P} , we have for all $k \in [K]$:

$$F_k(x) := \mathbb{P}_{\beta \sim \mathcal{P}_K^N}(\beta_k \leq x) = F^{N-K}(x) \sum_{j=1}^k \binom{N}{N-j} F^{K-j+1}(x) (1 - F(x))^{j-1} \quad (7)$$

311 Note that this is a polynomial in F ; in the following, we denote it P_k .

312 **Concentration Inequalities for order statistics** Exploiting (7), we build cdf bands for all order
 313 statistics based on observations of a single order statistic. For simplicity, it is stated in the full
 314 information feedback. For $k, k' \in [K]^2$ define the empirical estimate of the CDF of the k'^{th} order
 315 statistics based on the k^{th} : $\hat{F}_{k \rightarrow k'}^t := P_{k'}(P_k^{-1}(\hat{F}_k^t))$

316 **Lemma 5.** Let $t, k \in \mathbb{N}$, and let X_j^i for $i \in [t], j \in [k]$ be i.i.d. samples from a distribution \mathcal{P} with
 317 cdf F . With probability $1 - \alpha$:

$$\hat{F}_{k \rightarrow k'}^t(x) - \alpha_{k,k'}\epsilon \leq F_{k'}(x) \leq \hat{F}_{k \rightarrow k'}^t(x) + \alpha_{k,k'}\epsilon \quad (8)$$

318 Where $\epsilon = \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}}$ and $\alpha_{k,k'}$ is a constant which only depends on k, k' and N .

319 *Proof.* The lemma is proved using the DKW inequality (Massart, 1990), the order statistics cdf
 320 formula (7) and by noticing that all P_k are strictly increasing on $(0, 1)$. \square

321 **Algorithm** These concentration inequalities effectively enable recovery of full-information feed-
 322 back in the uniform price auction by selecting the best estimate. For $k \in [K]$ and $x \in [0, 1]$, let $t_k(x)$
 323 be the number of times $\mathbb{1}\{\beta_k \leq x\}$ was observed (Lemma 2). Define $k^*(x) = \arg\max_{k \in [K]} t_k(x)$
 324 as the index with the most observations, and estimate marginal CDFs as $\bar{F}_k^t(x) := \tilde{F}_{k^*(x) \rightarrow k}^t(x)$. The
 325 resulting utility estimate is $\bar{u}^t(\cdot) := U_u((\bar{F}_k^t)_{k \in [K]}, \cdot)$

Algorithm 5: UBIID algorithm

- 1 **Input:** time horizon T , number of adversary N .
 - 2 **Output:** bids for each time step $(\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^{T-1}, \mathbf{b}^T) \in B^T$. **for** $t = 1, 2, \dots, T$ **do**
 - 3 \lfloor Play $\mathbf{b}^t := \arg\max_{\mathbf{b} \in B} \bar{u}^t(\mathbf{b})$
-

326 **Theorem 5.** When facing i.i.d. adversaries in the uniform auction with bandit feedback, Algorithm 5
 327 guarantees the regret is upper bounded by $\tilde{O}\left(K^{5/2}\sqrt{T}\right)$.

328 *Proof sketch.* The previous result is essentially proved as the bound for the full-information feedback,
 329 by noticing that for all $x \in [0, 1]$, $t_{k^*(x)} \geq t/K$. This ensures that the estimates \bar{F} concentrate as well
 330 as in the full-information feedback up to a factor \sqrt{K} , techniques similar to the ones for Theorem 1
 331 then yield the result. \square

332 **Lower bounds** Because the opposing bids are order statistics, previous lower bounds may not
 333 apply. Intuitively, this is because constraining the opposing bids to be ordered i.i.d. can make learning
 334 to bid easier (as Theorem 5 shows). Let us consider the discriminatory auction first. Notice that
 335 for 1-unit auction, the general setting and the i.i.d. setting coincide. Therefore the lower bound from
 336 Lemma 3 applies and regret is at least $\Omega(T^{2/3})$. The uniform auction requires a new lower bound,
 337 that we provide below:

338 **Lemma 6.** When facing i.i.d. adversaries in the uniform auction with bandit feedback, any learning
 339 algorithm must incur at least a regret of $\Omega(\sqrt{T})$.

340 *Proof idea.* Focusing on two-item auctions, we build upon the proof of the lower bound of Brânzei
 341 et al., 2023 for the uniform auction in the full information feedback. Since they use opposing bids
 342 which can only be worth 0 or $2/3$, it is possible to build distributions that mimic the behavior of the
 343 ones used in their lower bound, even when using i.i.d. bids. \square

344 **Remark 2.** We assume knowledge of the number of adversaries N in this section. However, since
 345 the components of β are order statistics, N can also be estimated from a finite number of samples.

346 6 Conclusion

347 We characterized the difficulty of learning to bid in multi-unit auctions through achievable regret rates.
 348 We show that, under full-information feedback, both auction formats can be learned with matching
 349 regret rates. In the bandit feedback setting, the picture is more nuanced: while worst-case regret
 350 guarantees are similar for both formats, the richer feedback structure of the uniform auction can make
 351 it easier to learn in instances that do not correspond to the worst-case distribution.

352 A key simplification in our analysis is the stochasticity assumption for opposing bids. It remains an
 353 open question whether similar results hold when adversaries are strategic, which would also enable
 354 the study of social welfare when multiple bidders learn simultaneously. Another interesting direction
 355 is to consider settings where bidders' valuations must also be learned, and to investigate how auction
 356 transparency (Cesa-Bianchi et al., 2024) impacts the learning process.

References

- Achddou, J., Cappé, O., & Garivier, A. (2021a). Efficient algorithms for stochastic repeated second-price auctions. *Algorithmic Learning Theory*, 99–150.
- Achddou, J., Cappé, O., & Garivier, A. (2021b). Fast rate learning in stochastic first price bidding. *Asian Conference on Machine Learning*, 1754–1769.
- Anderson, E., & Holmberg, P. (2023). *Multi-unit auctions with uncertain supply and single-unit demand* (tech. rep.). Faculty of Economics, University of Cambridge.
- Audibert, J.-Y., & Bubeck, S. (2010). Best arm identification in multi-armed bandits. *COLT-23th Conference on learning theory-2010*, 13–p.
- Ausubel, L. M., Cramton, P., Pycia, M., Rostek, M., & Weretka, M. (2014). Demand reduction and inefficiency in multi-unit auctions. *The Review of Economic Studies*, 81(4), 1366–1400.
- Balseiro, S., Golrezaei, N., Mahdian, M., Mirrokni, V., & Schneider, J. (2019). Contextual bandits with cross-learning. *Advances in Neural Information Processing Systems*, 32.
- Bartl, D., & Mendelson, S. (2023). On a variance dependent dvoretzky-kiefer-wolfowitz inequality. *arXiv preprint arXiv:2308.04757*.
- Blanchard, M., & Voracek, V. (2024). Tight bounds for local glivenko-cantelli. *International Conference on Algorithmic Learning Theory*, 179–220.
- Boucheron, S., Lugosi, G., & Massart, P. (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford University Press.
- Brânzei, S., Derakhshan, M., Golrezaei, N., & Han, Y. (2023). Learning and collusion in multi-unit auctions. *Advances in Neural Information Processing Systems*, 36, 22191–22225.
- Casella, G., & Berger, R. (2024). *Statistical inference*. CRC press.
- Cesa-Bianchi, N., Cesari, T., Colomboni, R., Fusco, F., & Leonardi, S. (2024). The role of transparency in repeated first-price auctions with unknown valuations. *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, 225–236.
- Combes, R., Proutière, A., & Fauquette, A. (2020). Unimodal bandits with continuous arms: Order-optimal regret without smoothness. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 4(1), 1–28.
- De Keijzer, B., Markakis, E., Schäfer, G., & Telelis, O. (2013). Inefficiency of standard multi-unit auctions. *European Symposium on Algorithms*, 385–396.
- Elmaghraby, W., & Oren, S. S. (1999). The efficiency of multi-unit electricity auctions. *The Energy Journal*, 20(4), 89–116.
- Feng, Z., Podimata, C., & Syrgkanis, V. (2018). Learning to bid without knowing your value. *Proceedings of the 2018 ACM Conference on Economics and Computation*, 505–522.
- Galgana, R., & Golrezaei, N. (2025). Learning in repeated multiunit pay-as-bid auctions. *Manufacturing & Service Operations Management*, 27(1), 200–229.
- Golrezaei, N., & Sahoo, S. (2024). Bidding in uniform price auctions for value maximizing buyers. *arXiv preprint arXiv:2406.03674*.
- Hortaçsu, A., & Puller, S. L. (2008). Understanding strategic bidding in multi-unit auctions: A case study of the texas electricity spot market. *The RAND Journal of Economics*, 39(1), 86–114.

- 397 Kleinberg, R., & Leighton, T. (2003). The value of knowing a demand curve: Bounds on regret for
398 online posted-price auctions. *44th Annual IEEE Symposium on Foundations of Computer
399 Science, 2003. Proceedings.*, 594–605.
- 400 Krishna, V. (2009). *Auction theory*. Academic press.
- 401 Lattimore, T., & Szepesvári, C. (2020). *Bandit algorithms*. Cambridge University Press.
- 402 Massart, P. (1990). The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The annals of
403 Probability*, 1269–1283.
- 404 Mohri, M., & Medina, A. M. (2014). Learning theory and algorithms for revenue optimization in
405 second price auctions with reserve. *International conference on machine learning*, 262–270.
- 406 Naaman, M. (2021). On the tight constant in the multivariate dvoretzky–kiefer–wolfowitz inequality.
407 *Statistics & Probability Letters*, 173, 109088.
- 408 Nedelec, T., Calauzènes, C., El Karoui, N., Perchet, V., et al. (2022). Learning in repeated auctions.
409 *Foundations and Trends® in Machine Learning*, 15(3), 176–334.
- 410 Nyborg, K. G., Rydqvist, K., & Sundaresan, S. M. (2002). Bidder behavior in multiunit auctions:
411 Evidence from swedish treasury auctions. *Journal of Political Economy*, 110(2), 394–424.
- 412 Nyborg, K. G., & Sundaresan, S. (1996). Discriminatory versus uniform treasury auctions: Evidence
413 from when-issued transactions. *Journal of Financial Economics*, 42(1), 63–104. [https://doi.org/https://doi.org/10.1016/0304-405X\(95\)00871-B](https://doi.org/https://doi.org/10.1016/0304-405X(95)00871-B)
414
- 415 Potfer, M., Baudry, D., Richard, H., Perchet, V., & Wan, C. (2024). Improved learning rates in
416 multi-unit uniform price auctions. *Advances in Neural Information Processing Systems*, 37,
417 130237–130264.
- 418 Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to
419 algorithms*. Cambridge university press.
- 420 Syrgkanis, V., & Tardos, E. (2013). Composable and efficient mechanisms. *Proceedings of the
421 forty-fifth annual ACM symposium on Theory of computing*, 211–220.
- 422 Weed, J., Perchet, V., & Rigollet, P. (2016). Online learning in repeated auctions. *Conference on
423 Learning Theory*, 1562–1583.

NeurIPS Paper Checklist

1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: The claims in the abstract are provided in more details in the contribution subsection. These claims are then provided in the main paper. Proofs are available in the appendix.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[Yes\]](#)

Justification: Limitation are discussed in the conclusion. The result stated are clearly only valid in the limit of our modelling.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best

judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: The proofs are mostly in the supplementary material. high level overview of the ideas in the proofs are provided in the main paper.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
 - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.

- 518 (b) If the contribution is primarily a new model architecture, the paper should describe
519 the architecture clearly and fully.
- 520 (c) If the contribution is a new model (e.g., a large language model), then there should
521 either be a way to access this model for reproducing the results or a way to reproduce
522 the model (e.g., with an open-source dataset or instructions for how to construct
523 the dataset).
- 524 (d) We recognize that reproducibility may be tricky in some cases, in which case
525 authors are welcome to describe the particular way they provide for reproducibility.
526 In the case of closed-source models, it may be that access to the model is limited in
527 some way (e.g., to registered users), but it should be possible for other researchers
528 to have some path to reproducing or verifying the results.

529 5. Open access to data and code

530 Question: Does the paper provide open access to the data and code, with sufficient instruc-
531 tions to faithfully reproduce the main experimental results, as described in supplemental
532 material?

533 Answer: [NA]

534 Justification: The paper does not include experiments.

535 Guidelines:

- 536 • The answer NA means that paper does not include experiments requiring code.
- 537 • Please see the NeurIPS code and data submission guidelines ([https://nips.cc/public/](https://nips.cc/public/guides/CodeSubmissionPolicy)
538 [guides/CodeSubmissionPolicy](https://nips.cc/public/guides/CodeSubmissionPolicy)) for more details.
- 539 • While we encourage the release of code and data, we understand that this might not be
540 possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not
541 including code, unless this is central to the contribution (e.g., for a new open-source
542 benchmark).
- 543 • The instructions should contain the exact command and environment needed to run
544 to reproduce the results. See the NeurIPS code and data submission guidelines ([https://](https://nips.cc/public/guides/CodeSubmissionPolicy)
545 nips.cc/public/guides/CodeSubmissionPolicy) for more details.
- 546 • The authors should provide instructions on data access and preparation, including how
547 to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- 548 • The authors should provide scripts to reproduce all experimental results for the new
549 proposed method and baselines. If only a subset of experiments are reproducible, they
550 should state which ones are omitted from the script and why.
- 551 • At submission time, to preserve anonymity, the authors should release anonymized
552 versions (if applicable).
- 553 • Providing as much information as possible in supplemental material (appended to the
554 paper) is recommended, but including URLs to data and code is permitted.

555 6. Experimental setting/details

556 Question: Does the paper specify all the training and test details (e.g., data splits, hyper-
557 parameters, how they were chosen, type of optimizer, etc.) necessary to understand the
558 results?

559 Answer: [NA]

560 Justification: The paper does not include experiments.

561 Guidelines:

- 562 • The answer NA means that the paper does not include experiments.
- 563 • The experimental setting should be presented in the core of the paper to a level of detail
564 that is necessary to appreciate the results and make sense of them.

565 • The full details can be provided either with the code, in appendix, or as supplemental
566 material.

567 7. Experiment statistical significance

568 Question: Does the paper report error bars suitably and correctly defined or other appropriate
569 information about the statistical significance of the experiments?

570 Answer: [NA]

571 Justification: The paper does not include experiments.

572 Guidelines:

- 573 • The answer NA means that the paper does not include experiments.
- 574 • The authors should answer "Yes" if the results are accompanied by error bars, confi-
575 dence intervals, or statistical significance tests, at least for the experiments that support
576 the main claims of the paper.
- 577 • The factors of variability that the error bars are capturing should be clearly stated (for
578 example, train/test split, initialization, random drawing of some parameter, or overall
579 run with given experimental conditions).
- 580 • The method for calculating the error bars should be explained (closed form formula,
581 call to a library function, bootstrap, etc.)
- 582 • The assumptions made should be given (e.g., Normally distributed errors).
- 583 • It should be clear whether the error bar is the standard deviation or the standard error
584 of the mean.
- 585 • It is OK to report 1-sigma error bars, but one should state it. The authors should
586 preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis
587 of Normality of errors is not verified.
- 588 • For asymmetric distributions, the authors should be careful not to show in tables or
589 figures symmetric error bars that would yield results that are out of range (e.g. negative
590 error rates).
- 591 • If error bars are reported in tables or plots, The authors should explain in the text how
592 they were calculated and reference the corresponding figures or tables in the text.

593 8. Experiments compute resources

594 Question: For each experiment, does the paper provide sufficient information on the com-
595 puter resources (type of compute workers, memory, time of execution) needed to reproduce
596 the experiments?

597 Answer: [NA]

598 Justification: The paper does not include experiments.

599 Guidelines:

- 600 • The answer NA means that the paper does not include experiments.
- 601 • The paper should indicate the type of compute workers CPU or GPU, internal cluster,
602 or cloud provider, including relevant memory and storage.
- 603 • The paper should provide the amount of compute required for each of the individual
604 experimental runs as well as estimate the total compute.
- 605 • The paper should disclose whether the full research project required more compute
606 than the experiments reported in the paper (e.g., preliminary or failed experiments that
607 didn't make it into the paper).

608 9. Code of ethics

609 Question: Does the research conducted in the paper conform, in every respect, with the
610 NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines?>

Answer: [Yes]

Justification: The research is conform to neurips code of ethics.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: This is a theoretical paper, and as such it has very limited societal implication. The main possible implication are discussed in the conclusion section.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The paper poses no such risk.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring

658 that users adhere to usage guidelines or restrictions to access the model or implementing
659 safety filters.

- 660 • Datasets that have been scraped from the Internet could pose safety risks. The authors
661 should describe how they avoided releasing unsafe images.
- 662 • We recognize that providing effective safeguards is challenging, and many papers do
663 not require this, but we encourage authors to take this into account and make a best
664 faith effort.

665 12. Licenses for existing assets

666 Question: Are the creators or original owners of assets (e.g., code, data, models), used in
667 the paper, properly credited and are the license and terms of use explicitly mentioned and
668 properly respected?

669 Answer: [NA]

670 Justification: The paper does not use any existing assets.

671 Guidelines:

- 672 • The answer NA means that the paper does not use existing assets.
- 673 • The authors should cite the original paper that produced the code package or dataset.
- 674 • The authors should state which version of the asset is used and, if possible, include a
675 URL.
- 676 • The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- 677 • For scraped data from a particular source (e.g., website), the copyright and terms of
678 service of that source should be provided.
- 679 • If assets are released, the license, copyright information, and terms of use in the package
680 should be provided. For popular datasets, paperswithcode.com/datasets has curated
681 licenses for some datasets. Their licensing guide can help determine the license of a
682 dataset.
- 683 • For existing datasets that are re-packaged, both the original license and the license of
684 the derived asset (if it has changed) should be provided.
- 685 • If this information is not available online, the authors are encouraged to reach out to
686 the asset's creators.

687 13. New assets

688 Question: Are new assets introduced in the paper well documented and is the documentation
689 provided alongside the assets?

690 Answer: [NA]

691 Justification: The paper does not release new assets

692 Guidelines:

- 693 • The answer NA means that the paper does not release new assets.
- 694 • Researchers should communicate the details of the dataset/code/model as part of their
695 submissions via structured templates. This includes details about training, license,
696 limitations, etc.
- 697 • The paper should discuss whether and how consent was obtained from people whose
698 asset is used.
- 699 • At submission time, remember to anonymize your assets (if applicable). You can either
700 create an anonymized URL or include an anonymized zip file.

701 14. Crowdsourcing and research with human subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: The paper does not involve crowdsourcing nor research with human subject

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. Institutional review board (IRB) approvals or equivalent for research with human subjects

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: The paper does not involve crowdsourcing nor research with human subject

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. Declaration of LLM usage

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: The paper does not involve any LLM in the models and theoretical framework studied.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.

Appendix

This appendix provides additional technical details, proofs, and supporting results for the main paper. It is organized as follows:

- **Section A** presents general simplifications and key properties of the auction models that are useful for our analysis.
- **Section B** contains detailed proofs of the regret upper bounds for both the full-information and bandit feedback settings, including concentration inequalities and technical lemmas.
- **Section C** provides the proofs of the main regret lower bounds for the discriminatory (first price) and uniform price auctions.
- **Section D** covers proofs and results specific to certain instances, such as Δ -separated distributions and the i.i.d. unit-demand adversary setting.

Throughout, all notation, definitions, and assumptions are consistent with those in the main text, unless explicitly stated otherwise. Where appropriate, we restate key lemmas and theorems for clarity. Cross-references to the main text and within the appendix are provided for ease of navigation.

A Appendix

A.1 General simplifications

This section presents general properties and simplifications of the auction models considered in the main paper. We highlight structural results that are useful for the design and analysis of learning algorithms, such as truthfulness in uniform price auctions and the implications for optimal bidding strategies. These results serve as foundational tools for the regret analyses in subsequent sections.

Uniform price auction with unit demand We focus our attention on the unit demand uniform price auction, in which the bidder only wishes to acquire one item. Formally, this means that $v_1 > 0$ and for all $k > 1$, $v_k = 0$. One can notice that in this specific case, the optimal bid is straightforwardly $\mathbf{b}^* = (v_1, 0, \dots, 0)$. Therefore, in this case, there is no need for an online learning algorithm and a bidder playing \mathbf{b}^* for every time $t \in [T]$ incurs regret 0.

The following lemma shows a more general property of which the previous claim is a direct consequence.

Lemma 7. *In the uniform price auction, for any distribution of opposing bid \mathcal{D} , let $\mathbf{b}^* = (b_1^*, b_2^*, \dots, b_K^*) := \operatorname{argmax}_{\mathbf{b} \in B} \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)]$ then $\tilde{\mathbf{b}}^* := (v_1, \min(b_2^*, v_2), \dots, \min(b_K^*, v_K))$ is also a maximizer.*

Proof. We will provide a stronger result, which is that, given any opposing bids $\beta \in B$, when we denote $u(\mathbf{b}^*, \beta)$ the utility of the bidder against these bids when bidding \mathbf{b}^* , then $u(\mathbf{b}^*, \beta) \leq u(\tilde{\mathbf{b}}^*, \beta)$. Let k^* be the biggest index such that $b_{k^*}^* > v_{k^*}$. Denote $\tilde{\mathbf{b}}_{k^*}^*$ the bid with $b_{k^*}^*$ replaced by v_{k^*} . Let us consider $x(\mathbf{b}^*, \beta)$ and $x(\tilde{\mathbf{b}}_{k^*}^*, \beta)$, either they are equals or $x(\mathbf{b}^*, \beta) > x(\tilde{\mathbf{b}}_{k^*}^*, \beta)$.

- In the first case, noticing that the price function $p(\cdot, \beta)$ is increasing in all the components of the bids, yields $u(\mathbf{b}^*, \beta) \leq u(\tilde{\mathbf{b}}_{k^*}^*, \beta)$.

- In the second case, we have $x(\mathbf{b}^*, \beta) > x(\tilde{\mathbf{b}}_{k^*}^*, \beta)$, which necessarily implies that $b_{k^*}^* \geq p(\mathbf{b}^*, \beta) \geq v_{k^*}$. Therefore, $v_{k^*} - p(\mathbf{b}^*, \beta) \leq 0$, noticing that this is the last term of the utility when playing \mathbf{b}^* and that the price is increasing in all the components of the bids, we can conclude that $u(\mathbf{b}^*, \beta) \leq u(\tilde{\mathbf{b}}_{k^*}^*, \beta)$.

We can repeat this process for all the components of the bids. Combining this with the fact that the utility is only increasing in the first component of the bid vector (usually denoted b_1) allows us to get that $u(\mathbf{b}^*, \beta) \leq u(\tilde{\mathbf{b}}^*, \beta)$. \square

Because of the previous property, when analyzing learning algorithms for the uniform price auction, we will always consider the algorithms' pick bids such that $b_1 = v_1$ when maximizing estimates of expected utility.

794 We now move on to providing the proofs and formulas necessary to justify our approach. We recall
 795 the Lemma 1.

796 **Lemma 1.** *In both auction formats, the expected utility can be expressed as a function of the bidder's*
 797 *bid vector \mathbf{b} and the marginal cumulative distribution functions $(F_k)_{k \in [K]}$.*

798 In fact, for the discriminatory auction, the utility can be written as follows:

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] = \sum_{i=1}^K F_{K-i+1}(b_i) (v_i - b_i) \quad (9)$$

799 In the uniform price auction, the expected utility can be written as follows :

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] = \sum_{i=1}^K (F_{K-i}(b_{i+1}) - F_{K-i+1}(b_{i+1})) \left(\sum_{l=1}^i v_l - b_{i+1} \right) \quad (10)$$

$$+ \sum_{i=1}^K \mathbb{E}_{\beta_{K-i+1} \sim \mathcal{D}_{K-i+1}} \left[\mathbb{1} \{b_i \geq \beta_{K-i+1} \geq b_{i+1}\} \left(\sum_{l=1}^i v_l - \beta_{K-i+1} \right) \right] \quad (11)$$

800 *Proof.* We begin by proving the lemma for the discriminatory auction. The utility can be rewritten
 801 as $u(\mathbf{b}, \beta) = \sum_{i=1}^K \mathbb{1} \{b_i \geq \beta_{K-i+1}\} (v_i - b_i)$. Taking the expectation and rewriting by using the
 802 definition F_k leads to the following equation, which concludes the proof.

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] = \sum_{i=1}^K F_{K-i+1}(b_i) (v_i - b_i) \quad (12)$$

803 We now move to the uniform price auction. To have a convenient formula, we write the utility of the
 804 bidder in the uniform price auction by decomposing it depending on the $K + 1^{\text{th}}$ bid.

$$u(\mathbf{b}, \beta) = \sum_{i=1}^K \mathbb{1} \{\beta_{K-i} > b_{i+1} \geq \beta_{K-i+1}\} \left(\sum_{l=1}^i v_l - b_{i+1} \right) \\ + \sum_{i=1}^K \mathbb{1} \{b_i \geq \beta_{K-i+1} > b_{i+1}\} \left(\sum_{l=1}^i v_l - \beta_{K-i+1} \right)$$

805 When taking the expectation, this leads to

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] = \sum_{i=1}^K \mathbb{P}(\beta_{K-i} > b_{i+1} \geq \beta_{K-i+1}) \left(\sum_{l=1}^i v_l - b_{i+1} \right) \\ + \sum_{i=1}^K \mathbb{P}(b_i \geq \beta_{K-i+1} > b_{i+1}) \left(\sum_{l=1}^i v_l \right) - i \mathbb{E} [\mathbb{1} \{b_i \geq \beta_{K-i+1} > b_{i+1}\} \beta_{K-i+1}]$$

806 Which we can simplify as

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] = \sum_{i=1}^K (F_{K-i+1}(b_{i+1}) - F_{K-i}(b_{i+1})) \left(\sum_{l=1}^i v_l - b_{i+1} \right) \\ + \sum_{i=1}^K (F_{K-i+1}(b_i) - F_{K-i+1}(b_{i+1})) \left(\sum_{l=1}^i v_l \right) \quad (13) \\ - \sum_{i=1}^K i \left(b_i F_{K-i+1}(b_i) - b_{i+1} F_{K-i+1}(b_{i+1}) - \int_{b_{i+1}}^{b_i} F_{K-i+1}(t) dt \right)$$

807 This concludes the proof. \square

With the previous proof in mind, we can define U_u and U_d , the corresponding functions that, given the CDFs, allow us to compute the expected utility.

Let $(G_k)_{k \in [K]}$ be a sequence of CDF functions. The following equations define the functions mentioned in Lemma 1 for: **the discriminatory auction** for $\mathbf{b} \in B$:

$$U_d((G_k)_{k \in [K]}, \mathbf{b}) = \sum_{i=1}^K G_{K-i+1}^t(b_i) (v_i - b_i), \quad (14)$$

the uniform auction for $\mathbf{b} \in B$:

$$U_u((G_k)_{k \in [K]}, \mathbf{b}) := \sum_{i=1}^K (G_{K-i+1}(b_{i+1}) - G_{K-i}(b_{i+1})) \left(\sum_{l=1}^i v_l - b_{i+1} \right) + \sum_{i=1}^K (G_{K-i+1}(b_i) - G_{K-i+1}(b_{i+1})) \sum_{l=1}^i v_l - i \int_{(b_{i+1}, b_i]} x dG_{K-i+1} \quad (15)$$

Lemma 2. At time $t \in [T]$, let $\mathbf{b}^t, \beta^t \in B$ be the bids of the learner and the adversary, in the uniform auction with bandit feedback the bidder observes $(\mathbb{1} \{ \beta_{K-i+1}^t \in (b_{i+1}^t, b_i^t] \} \beta_{K-i+1}^t)_{i \in [K]}$.

Proof. At time t , let $i \in [K]$, then $\beta_{K-i+1}^t \in (b_{i+1}^t, b_i^t]$, is equivalent to $b_{i+1}^t < \beta_{K-i}^t \leq b_i^t$ which is in turn equivalent to $\{x(\mathbf{b}^t, \beta^t) = i\} \cap \{p(\mathbf{b}^t, \beta^t) \neq b_{i+1}^t\}$, which is an observable event. Note that when the event is realized, the value of β_{K-i+1}^t is observed since it is the price.

Therefore, since it is the case for all $i \in [K]$, the bidder has observed $(\mathbb{1} \{ \beta_{K-i+1}^t \in [b_{i+1}^t, b_i^t] \} \beta_{K-i+1}^t)_{i \in [K]}$. \square

Remark 3. Notice that the previous Lemma 2 implies that at all times t and for all $i \in [K]$, the following function is also observed $\forall x \in (b_{i+1}^t, b_i^t]$, $\mathbb{1} \{ \beta_{K-i}^t \in [0, x] \}$.

B Regret Upper Bounds

In this section, we provide detailed proofs of the regret upper bounds stated in the main text for both the full-information and bandit feedback settings. We include all necessary technical lemmas, concentration inequalities, and step-by-step arguments for the algorithms analyzed. Our goal is to make the derivations transparent and self-contained, enabling readers to verify each step of the analysis.

B.1 Full-information feedback

Recall that in the full information feedback, the bidder observes β^t at the end of each time-step. therefore, at time t , it has observed $\beta^1, \dots, \beta^{t-1}$.

The following lemma, which links the quality of the estimates of the marginal CDF to the quality of the estimates of the utility, can be used to derive regret guarantees for both auction formats.

9

Lemma 8. Given $(\hat{F}_k^t)_{k \in [K]}$, the empirical estimates of the cumulative marginal distributions of \mathcal{D} from t samples, let $\alpha \in [0, 1]$, then in both the uniform price auction (with $\hat{u}(\cdot) := U_u((\hat{F}_k)_{k \in [K]}, \cdot)$) and in the discriminatory auction (with $\hat{u}(\cdot) := U_d((\hat{F}_k)_{k \in [K]}, \cdot)$). The following inequality holds :

$$\forall k \in [K], \mathbb{P} \left(\sup_{\mathbf{b} \in B} |\hat{u}(\mathbf{b}) - \mathbb{E}[u(\mathbf{b}, \beta)]| \leq K \epsilon^t(\alpha) \right) \geq 1 - K\alpha \quad (16)$$

Where for the discriminatory auction $\epsilon^t(\alpha) := \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}}$ and for the uniform auction $\epsilon^t(\alpha) := 3\sqrt{C \frac{2K \log(2K) + \log(\frac{1}{\alpha})}{t}}$

839 *Proof.* We begin with the proof of the guarantees for the discriminatory auction, which is more
 840 concise: Let $T \in \mathbb{N}$, at time $t \in [T]$, we apply the result from Massart, 1990, which in our notations
 841 provides that, for $\alpha \in [0, 1]$:

$$\forall k \in [K], \mathbb{P} \left(\sup_{x \in [0,1]} |\hat{F}_k^t(x) - F_k(x)| \leq \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}} \right) \geq 1 - \alpha \quad (17)$$

842 Doing a union over the opposite events, one easily gets :

$$\mathbb{P} \left(\forall k \in [K], \sup_{x \in [0,1]} |\hat{F}_k^t(x) - F_k(x)| \leq \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}} \right) \geq 1 - K\alpha \quad (18)$$

843 We can write the expected utility in the discriminatory price auction, for $\mathbf{b} = (b_1, \dots, b_K) \in B$ as
 844 $\mathbb{E}[u(\mathbf{b}, \beta)] = \sum_{k \in [K]} F_k(b_k) (v_k - b_k)$. Using the expression of \hat{u} , we get:

$$\begin{aligned} \sup_{\mathbf{b} \in B} |\hat{u}(\mathbf{b}) - \mathbb{E}[u(\mathbf{b}, \beta)]| &\leq \sup_{\mathbf{b} \in B} \sum_{k \in [K]} |\hat{F}_k(b_k) - F_k(b_k)| \\ &\leq \sum_{k \in [K]} \sup_{b_k \in [0,1]} |\hat{F}_k(b_k) - F_k(b_k)| \end{aligned}$$

845 Using (18) then yields the desired result.

846 We now proceed with the proof for the uniform price auction. We start by rewriting the utility from
 847 (13) as follows :

$$\begin{aligned} \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] &= \mathbb{E}_{\beta \sim \mathcal{D}} \left[\sum_{i=1}^K \mathbb{1} \{ \beta_{K-i} > b_{i+1} \geq \beta_{K-i+1} \} \left(\sum_{j=1}^i v_j - b_{i+1} \right) \right. \\ &\quad \left. + \sum_{i=1}^K \mathbb{1} \{ b_i \geq \beta_{K-i+1} > b_{i+1} \} \left(\sum_{j=1}^i v_j \right) \right] \\ &\quad - \sum_{i=1}^K i \left(b_i F_{K-i+1}(b_i) - b_{i+1} F_{K-i+1}(b_{i+1}) - \int_{b_{i+1}}^{b_i} F_{K-i+1}(t) dt \right) \\ &= \mathbb{E}_{\beta \sim \mathcal{D}} \left[\sum_{i=1}^K \mathbb{1} \{ \beta_{K-i} > b_{i+1} \geq \beta_{K-i+1} \} \left(\sum_{j=1}^i v_j - b_{i+1} \right) \right. \\ &\quad \left. + \sum_{i=1}^K \mathbb{1} \{ b_i \geq \beta_{K-i+1} > b_{i+1} \} \left(\sum_{j=1}^i v_j - b_{i+1} \right) \right] \\ &\quad - \sum_{i=1}^K i (b_i - b_{i+1}) F_{K-i+1}(b_i) \\ &\quad - \sum_{i=1}^K i \int_{b_{i+1}}^{b_i} F_{K-i+1}(t) dt \end{aligned}$$

848 The estimated utility can also be written similarly, yielding:

$$\begin{aligned} \hat{u}^t(\mathbf{b}) = & \hat{\mathbb{E}}^t \left[\sum_{i=1}^K \mathbb{1} \{ \beta_{K-i} > b_{i+1} \geq \beta_{K-i+1} \} \left(\sum_{j=1}^i v_j - b_{i+1} \right) \right. \\ & + \sum_{i=1}^K \mathbb{1} \{ b_i \geq \beta_{K-i+1} > b_{i+1} \} \left(\sum_{j=1}^i v_j - b_{i+1} \right) \Big] \\ & - \sum_{i=1}^K i (b_i - b_{i+1}) \hat{F}_{K-i+1}(b_i) \\ & - \sum_{i=1}^K i \int_{b_{i+1}}^{b_i} \hat{F}_{K-i+1}(t) dt \end{aligned}$$

849 Where $\hat{\mathbb{E}}^t$ denotes the empirical expectation.

850 We will show that the estimated utility is close to the true expectation by showing that it is the case
851 for each of the terms of the sum. We begin with the terms that only require DKW bound, which we
852 restate here: Let $T \in \mathbb{N}$, at time $t \in [T]$, the result from Massart, 1990, provides that, for $\alpha \in [0, 1]$:

$$\forall k \in [K], \mathbb{P} \left(\sup_{x \in [0, 1]} |\hat{F}_k^t(x) - F_k(x)| \leq \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}} \right) \geq 1 - \alpha \quad (19)$$

853 We can therefore get that with probability $1 - K\alpha$:

$$\left| \sum_{i=1}^K i \int_{b_{i+1}}^{b_i} F_{K-i+1}(t) dt - \sum_{i=1}^K i \int_{b_{i+1}}^{b_i} \hat{F}_{K-i+1}(t) dt \right| \leq \sum_{i=1}^K i (b_i - b_{i+1}) \epsilon \leq K\epsilon \quad (20)$$

854 Similarly, with probability $1 - K\alpha$:

$$\left| \sum_{i=1}^K i (b_i - b_{i+1}) F_{K-i+1}(b_i) - \sum_{i=1}^K i (b_i - b_{i+1}) \hat{F}_{K-i+1}(b_i) \right| \leq \sum_{i=1}^K i (b_i - b_{i+1}) \epsilon \leq K\epsilon \quad (21)$$

855 Note that (20) and (21) are also valid when taking the supremum over $b \in B$.

856 We now move to bounding the last difference needed. We first provide the required tools.

857 Usual concentration inequality for the CDF can be obtained by considering the family of indicator
858 functions $: x \mapsto \mathbb{1} \{ X \leq x \}$. In order to obtain tighter concentration bounds, we will use results on
859 uniform convergence results on multiclass learnability from Shalev-Shwartz and Ben-David, 2014 to
860 bound all the terms simultaneously.

861 First notice that for any $\beta \in B$, the family of events $(\{\beta_{K-i} > b_{i+1} \geq \beta_{K-i+1}\})_{i \in [K]} \cup$
862 $(\{b_i \geq \beta_{K-i+1} > b_{i+1}\})_{i \in [K]}$ are disjoint for any $\mathbf{b} \in B$. We will consider the multiclass clas-
863 sification functions associated with these events for $\mathbf{b} \in B$, and denote \mathcal{H}_B this family of classifi-
864 cation functions. We also denote f the function which maps an element of B to the corresponding
865 multiclassifier.

866 We begin by bounding the Natarajan dimension of \mathcal{H}_B , a generalisation of VC dimension. For a
867 reminder of the definition, see Shalev-Shwartz and Ben-David, 2014 section 29.1.

868 Notice that amongst the events, there are interval indicator functions for each coordinate of the
869 vector β . Therefore, we can use the same argument as for obtaining the VC dimension of the family
870 of functions $\mathbb{1} \{ X \in [a, b] \}$. The main difference being that we require $2K + 2$ element of B in
871 order to have at least three different values amongst one coordinate, than observing that one cannot
872 obtain labels $[1, 0, 1]$ in the corresponding class yields the following upper bound on the Natarajan
873 dimension \mathcal{N} of the family of multiclassifier at hand:

$$\mathcal{N}(\mathcal{H}_B) \leq 2K + 2 \quad (22)$$

874 Then, using the Multiclass Fundamental Theorem Shalev-Shwartz and Ben-David, 2014, **Theorem**
 875 **29.3** yields the following concentration inequality, with probability $1 - \alpha$:

$$\sup_{\mathbf{b} \in B} \left| \frac{\sum_{t=1}^T f(\mathbf{b})(\beta^t)}{T} - \mathbb{E}_{\beta \sim \mathcal{D}} [f(\mathbf{b})(\beta)] \right| \leq \epsilon^t \quad (23)$$

876 Where there is a universal constant $C > 0$ such that $\epsilon^T := \sqrt{C \frac{2K \log(K) + \log(\frac{1}{\alpha})}{T}}$

877 We can therefore use this bound in order to obtain the following with probability at least $1 - \alpha$:

$$\begin{aligned} \sup_{\mathbf{b} \in B} & \left| \hat{\mathbb{E}}^t \left[\sum_{i=1}^K (\mathbb{1}\{\beta_{K-i} > b_{i+1} \geq \beta_{K-i+1}\} + \mathbb{1}\{b_i \geq \beta_{K-i+1} > b_{i+1}\}) \left(\sum_{j=1}^i v_j - b_{i+1} \right) \right] \right. \\ & \left. - \hat{\mathbb{E}}^t \left[\sum_{i=1}^K (\mathbb{1}\{\beta_{K-i} > b_{i+1} \geq \beta_{K-i+1}\} + \mathbb{1}\{b_i \geq \beta_{K-i+1} > b_{i+1}\}) \left(\sum_{j=1}^i v_j - b_{i+1} \right) \right] \right| \\ & \leq K \epsilon^t \end{aligned} \quad (24)$$

878 With this equation, we can now bound the error of the expected utility using (20), (21), and (24), with
 879 probability at least $1 - \alpha$:

$$\sup_{\mathbf{b} \in B} \left| \hat{u}^t(\mathbf{b}) - \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] \right| \leq 3K \epsilon^t \quad (25)$$

880 which concludes the proof. \square

881 We restate the theorem providing upper bounds on the regret in the full-information feedback :

882 **Theorem 1.** *Under full-information feedback, Algorithm 1 achieves a regret $R_T = \tilde{\mathcal{O}}(K\sqrt{T})$ for*
 883 *both the discriminatory and the uniform price auction.*

884 *Proof.* Lemma 8 yields the following concentration inequality :

$$\mathbb{P} \left(|\hat{u}^t(\mathbf{b}) - \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)]| \leq K \epsilon^t \right) \geq 1 - K\alpha \quad (26)$$

885 Let's denote the optimal bid $\mathbf{b}^* := \operatorname{argmax}_{\mathbf{b} \in B} \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)]$. Since \mathbf{b}^t is the maximum of \hat{u}^t , we
 886 known that $\hat{u}^t(\mathbf{b}^*) \leq \hat{u}^t(\mathbf{b}^t)$. Applying the previous concentration inequality to both \mathbf{b}^* and \mathbf{b}^t and
 887 leveraging the optimality of \mathbf{b}^t gives us with probability $1 - K\alpha$:

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}^t, \beta)] \geq \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}^*, \beta)] + 6K \epsilon^t \quad (27)$$

888 Choosing $\alpha = \frac{1}{T^2}$ and summing over $t \in [T]$, then yields the desired regret bounds : $R_T =$
 889 $\tilde{\mathcal{O}}(K\sqrt{T})$. \square

891 B.2 Bandits Feedback

892 We provide in this section the details necessary in order to obtain the guarantees for the algorithms 2
 893 and 3. This subsection is therefore only focused on the uniform price auction, which has the most
 894 expressive bandit feedback of the two auction formats.

895 Let us recall that in this setting, the bidder only observes the outcome of the auctions at the end of
 896 each time step. Formally at time t , with β^t the bid of the adversary and \mathbf{b}^t its own bid, the bidder
 897 observes $x(\beta^t, \mathbf{b}^t)$ and $p(\beta^t, \mathbf{b}^t)$. Lemma 2 provides an alternative representation, while Remark 3
 898 points out how it relates to indicator functions used for CDFs empirical estimates.

899 We recall the formula of the estimates of the marginal CDFs for $k \in [K]$:

$$\forall x \in [0, 1], \tilde{F}_k(x) := \frac{\sum_{j=1}^t \mathbb{1}\{x \in (b_{K+2-k}^t, b_{K+1-k}^t]\} \mathbb{1}\{\beta_k^t \leq x\}}{\sum_{j=1}^t \mathbb{1}\{x \in (b_{K+2-k}^t, b_{K+1-k}^t]\}} \quad (28)$$

B.2.1 Explore then commit algorithm

We restate the theorem providing upper bounds on the regret in the bandit setting for the uniform auction.

Theorem 2. *The estimate then commit algorithms, with the exploration time $T_{expl} = K^{2/3}T^{2/3}$ achieves a regret bound of $\tilde{O}(K^{5/3}T^{2/3})$*

Proof. In order to prove the regret of the algorithm, we provide a two-part bound which matches the two-phase behavior of the algorithm. Indeed, we use the following regret decomposition between exploration regret and commitment regret :

$$R_T = R_T^{expl} + R_T^{com} \quad (29)$$

Where we define the exploration regret as

$$R_T^{expl} = T_{expl} \sup_{\mathbf{b} \in B} \left(\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] \right) - \sum_{t=1}^{T_{expl}} \mathbb{E}_{\beta^t \sim \mathcal{D}} [u(\mathbf{b}^t, \beta^t)] , \quad (30)$$

and the commitment regret as

$$R_T^{com} = (T - T_{expl}) \sup_{\mathbf{b} \in B} \left(\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] \right) - \sum_{t=T_{expl}}^T \mathbb{E}_{\beta^t \sim \mathcal{D}} [u(\mathbf{b}^t, \beta^t)] . \quad (31)$$

We begin by upper-bounding the exploration regret by noticing that the utility is upper-bounded by K . This leads to

$$R_T^{expl} \leq K T_{expl} \quad (32)$$

Notice that during the exploration phase, the bidder ensures that for every $k \in [K]$ it has at least observed $\lfloor T_{expl}/K \rfloor$ samples of β_k .

Notice that during the exploration phase, for every $k \in [K]$, the bidder has observed at least $\lfloor T_{expl}/K \rfloor$ samples of β_k . This follows directly from applying Lemma 2 to the bids in the exploration phase. This ensures our CDFS estimates \tilde{F} are valid on the whole domain and that we can create concentration bands.

We will use concentration results on the empirical CDF to show that our estimate of the expected utility is close to the actual one. We begin by stating the concentration results that we will use :

We can use Massart, 1990 to show that for all $k \in [K]$: :

$$\forall k \in [K], \mathbb{P} \left(\sup_{x \in [0,1]} \left| \tilde{F}_k^{T_{expl}}(x) - F_k(x) \right| \leq \sqrt{\frac{\ln \left(\frac{2}{\alpha} \right)}{2 \lfloor T_{expl}/K \rfloor}} \right) \geq 1 - \alpha \quad (33)$$

We will need another concentration result. It is well known that the family of indicator functions of intervals has VC dimensions 2, therefore, it follows from section 12.5 of Boucheron et al., 2013 that for any $k \in [K]$:

$$\mathbb{P} \left(\sup_{(a,b)^2 \in [0,1], a < b} \left| \frac{\tilde{F}_k^{T_{expl}}(b) - \tilde{F}_k^{T_{expl}}(a) - F_k(b) + F_k(a)}{\sqrt{F_k(b) - F_k(a)}} \right| \leq \epsilon (\lfloor T_{expl}/K \rfloor \alpha) \right) \geq 1 - \alpha \quad (34)$$

Where $\epsilon(T, \alpha) = \sqrt{2 \frac{\ln \left(\frac{2}{\alpha} \right) + \log(eT)}{T}}$.

To upper bound $\sup_{\mathbf{b} \in B} \left| \mathbb{E}_{\beta \sim \mathcal{D}} (\mathbf{b}, \beta) - \tilde{u}^{T_{expl}}(\mathbf{b}) \right|$, we will use the following formulation of the expected utility into terms well suited for showing uniform convergence.

$$\begin{aligned}
\mathbb{E}[u(\mathbf{b}, \beta)] &= \sum_{i=1}^K F_{K-i+1}(b_i) v_i - \sum_{i=1}^K i b_{i+1} (F_{K-i+1}(b_{i+1}) - F_{K-i}(b_{i+1})) \\
&\quad - \sum_{i=1}^K i b_{i+1} (F_{K-i+1}(b_i) - F_{K-i+1}(b_{i+1})) \\
&\quad - \sum_{i=1}^K i \int_{b_{i+1}}^{b_i} F_{K-i+1}(t) dt - \sum_{i=1}^K i (b_i - b_{i+1}) F_{K-i+1}(b_i)
\end{aligned} \tag{35}$$

927 Notice first that we can obtain the two following equations as in the proof of Lemma 8, we can use
928 (33) to obtain similar bounds as in (20) and (21).

$$\left| \sum_{i=1}^K i \int_{b_{i+1}}^{b_i} F_{K-i+1}(t) dt - \sum_{i=1}^K i \int_{b_{i+1}}^{b_i} \tilde{F}_{K-i+1}^{T_{expl}}(t) dt \right| \leq \sum_{i=1}^K i (b_i - b_{i+1}) \epsilon \leq K \epsilon \tag{36}$$

$$\left| \sum_{i=1}^K i (b_i - b_{i+1}) F_{K-i+1}(b_i) - \sum_{i=1}^K i (b_i - b_{i+1}) \tilde{F}_{K-i+1}^{T_{expl}}(b_i) \right| \leq \sum_{i=1}^K i (b_i - b_{i+1}) \epsilon \leq K \epsilon \tag{37}$$

929 We can also get the following bound:

$$\left| \sum_{i=1}^K \left(F_{K-i+1}(b_i) - \tilde{F}_{K-i+1}^{T_{expl}}(b_i) \right) v_i \right| \leq K \epsilon \tag{38}$$

930 It only remains to get concentration bounds for the following term:

$$\begin{aligned}
&\left| \sum_{i=1}^K i b_{i+1} (2F_{K-i+1}(b_{i+1}) - F_{K-i}(b_{i+1}) - F_{K-i+1}(b_i)) \right. \\
&\quad \left. - i b_{i+1} (2\tilde{F}_{K-i+1}(b_{i+1}) - \tilde{F}_{K-i}(b_{i+1}) - \tilde{F}_{K-i+1}(b_i)) \right|
\end{aligned} \tag{39}$$

931 We start by the following computations :

$$\begin{aligned}
\sum_{i=1}^K i b_{i+1} (F_{K-i+1}(b_{i+1}) - F_{K-i}(b_{i+1})) &= \sum_{i=1}^K i b_{i+1} (F_{K-i+1}(b_{i+1})) \\
&\quad - \sum_{i=2}^{K+1} (i-1) b_i (F_{K-i+1}(b_i)) \\
&= \sum_{i=1}^K b_{i+1} (F_{K-i+1}(b_{i+1})) + \sum_{i=2}^K b_{i+1} (i-1) (F_{K-i+1}(b_{i+1}) - F_{K-i+1}(b_i)) \\
&\quad - \sum_{i=2}^{K+1} (i-1) (b_i - b_{i+1}) (F_{K-i+1}(b_i)) - K b_{K+1}
\end{aligned}$$

932 Adding the other terms in F of (39) yields :

$$\begin{aligned}
\sum_{i=1}^K i b_{i+1} (2F_{K-i+1}(b_{i+1}) - F_{K-i}(b_{i+1}) - F_{K-i+1}(b_i)) &= \sum_{i=1}^K b_{i+1} (F_{K-i+1}(b_{i+1})) \\
&\quad + \sum_{i=2}^K b_{i+1} (2i-1) (F_{K-i+1}(b_{i+1}) - F_{K-i+1}(b_i)) \\
&\quad - \sum_{i=2}^{K+1} (i-1) (b_i - b_{i+1}) (F_{K-i+1}(b_i)) - K b_{K+1}
\end{aligned} \tag{40}$$

933 Naturally, \tilde{u} can be written the same way, we will therefore bound (39) by bounding the error on each
 934 sum independently. By upper bounding using both (33) and (34) we can upper bound the last terms
 935 as follows, with probability $1 - 2K\alpha$:

$$\sup_{\mathbf{b} \in B} \left| \sum_{i=1}^K i b_{i+1} (2F_{k-i+1}(b_{i+1}) - F_{k-i}(b_{i+1}) - F_{K-i+1}(b_i)) \right. \\ \left. - i b_{i+1} (2\tilde{F}_{k-i+1}(b_{i+1}) - \tilde{F}_{k-i}(b_{i+1}) - \tilde{F}_{K-i+1}(b_i)) \right| \leq K\epsilon + K^{3/2}\epsilon + K\epsilon \quad (41)$$

936 Combining equations therefore yields with probability $1 - 2K\alpha$

$$\sup_{b \in B} \left| \mathbb{E}_{\beta \sim \mathcal{D}} (\mathbf{b}, \beta) - \tilde{u}^{T_{expl}}(\mathbf{b}) \right| \leq 6K^{3/2}\epsilon \quad (42)$$

937 Where $\epsilon = \sqrt{2 \frac{\ln(\frac{2}{\alpha}) + \log(e \lfloor T_{expl}/K \rfloor)}{\lfloor T_{expl}/K \rfloor}}$.

938 Then notice from Algorithm 2 that for all $t > T_{expl}$, all the bids \mathbf{b}^t are the same (because they are
 939 the solution of the same maximization problem). By optimality we get $\tilde{u}^{T_{expl}}(\mathbf{b}^{T_{expl}}) \geq \tilde{u}^{T_{expl}}(\mathbf{b}^*)$.
 940 Combining with (42) then yields for all $t \geq T_{expl}$ with probability $1 - 2K\alpha$:

$$\mathbb{E}[u(\mathbf{b}^*, \beta)] - \mathbb{E}[u(\mathbf{b}^t, \beta)] \leq 6K^{3/2}\epsilon \quad (43)$$

941 By summing over $T_{expl} \leq t \leq T$ and recognizing the formula of R_T^{com} we get, with probability
 942 $1 - 2K\alpha$:

$$R_T^{com} \leq 12K^{5/3}T^{2/3} \sqrt{\ln\left(\frac{2}{\alpha}\right) + \ln(e \lfloor T_{expl}/K \rfloor)}. \quad (44)$$

943 Since $T_{expl} = K^{2/3}T^{2/3}$.

944 And from (32) we get:

$$R_T^{expl} \leq K^{5/3}T^{2/3} \quad (45)$$

945 Choosing $\alpha = \frac{1}{T^2}$ and summing R_T^{com} and R_T^{expl} yields the desired regret bounds.

946 □

947 B.2.2 Improved bandit algorithm

948 We now move to the improved bandit algorithm, which can allow for a better regret bound when the
 949 instance it faces makes it possible.

950 We restate the theorem providing upper bounds on the regret in the bandit setting, including the
 951 instance-dependent bounds. Before providing the proof, we introduce several instance-dependent
 952 quantities which are relevant for the instance-dependent bounds and the proofs.

953 **Theorem 4.** *Algorithm 3 guarantee a regret of $\tilde{O}(K^{5/3}T^{2/3})$ in general and, when $\Delta > 0$, an*
 954 *instance-dependent regret bounds of $\tilde{O}(K\sqrt{T})$.*

955 *Proof.* For the analysis, we will consider two cases, one that we call the worst-case, and one that is
 956 the good instance. We talk about good instances if during the first phase of the algorithm for some t^* ,
 957 Δ^{t^*} is positive, and if not, it is a worst-case instance.

958 **Worst-case** Let us first analyse what happens in worst-case instances : $\forall t \in [T_{expl}], \Delta^t \leq 0$. As
 959 for the proof of Theorem 2, we decompose the regret into exploration and rest regret.

$$R_T = R_T^{expl} + R_T^{com} \quad (46)$$

960 We can, by noticing the utility is positive and smaller than K , upper bound the exploration regret by
 961 KT_{expl} , which yields

$$R_T^{expl} \leq K^{5/3}T^{2/3} \quad (47)$$

In order to provide a similar analysis of the regret of the second phase as for Algorithm 2, we need to show that after the exploration phase, the bidder has effectively observed enough samples of each component of $\beta \sim \mathcal{D}$ on the right intervals (because observation is no longer guaranteed on $[0, 1]$).

We recall the formula of the marginal CDF estimators for $k \in [K]$:

$$\forall x \in [0, 1], \tilde{F}_k(x) := \frac{\sum_{j=1}^t \mathbb{1}\{x \in (b_{K+2-k}^t, b_{K+1-k}^t]\} \mathbb{1}\{\beta_k^t \leq x\}}{\sum_{j=1}^t \mathbb{1}\{x \in (b_{K+2-k}^t, b_{K+1-k}^t]\}} \quad (48)$$

It suffices to notice that during the exploration phase, at time t , when $k_t := t - K\lfloor t/K \rfloor + 1$, the interval corresponding to \tilde{F}_{K-k_t+1} , that reads $\mathbb{1}\{x \in (b_{k_t+1}^t, b_{k_t}^t]\}$ strictly contains $\mathcal{I}_{k_t}^t \cup \mathcal{I}_{k_t+1}^t$, and by monotonicity of the intervals, contains $\mathcal{I}_k^{t'} \cup \mathcal{I}_{k+1}^{t'}$ for any $t' \geq t$. We can therefore write : For any $t \in [T]$, for any $k \in [K]$, $\forall x \in \mathcal{I}_k^t \cup \mathcal{I}_{k+1}^t$, $\sum_{j=1}^t \mathbb{1}\{x \in (b_{k_t+1}^t, b_{k_t}^t]\} \geq \min(\lfloor t/K \rfloor, \lfloor T_{expl}/K \rfloor)$.

Notice from (15) that for $t' \geq t$, for $\mathbf{b} \in \prod_{k \in [K]} \mathcal{I}_k^{t'}$, \tilde{F}_{K-k_t+1} is only evaluated on $\mathcal{I}_k^{t'} \cup \mathcal{I}_{k+1}^{t'}$. The previous statement is therefore sufficient for the concentration results of Theorem 2 to be reproduced.

The only modification needed is to notice that we can only guarantee that the optimal bid \mathbf{b}^* remains in the $\prod_{k \in [K]} \mathcal{I}_k^{t'}$ with probability $1 - K\alpha$ at each time. The probability of the bad event "the best bids are out of our confidence intervals" is at most $K\alpha T$.

We can therefore obtain that with probability $1 - 2K\alpha T$:

$$R_T^{com} \leq 12K^{5/3}T^{2/3} \sqrt{\ln\left(\frac{2}{\alpha}\right) + \ln(e\lfloor T_{expl}/K \rfloor)}. \quad (49)$$

Choosing $\alpha = \frac{1}{T^3}$ and summing R_T^{com} and R_{expl}^T yields the desired worst-case regret bounds of $K^{5/3}T^{2/3}$.

Instance dependent In this section of the proofs, we examine the case where for some $t^* \in [T_{expl}]$, we have $\Delta^{t^*} > 0$. First note that because intervals are *shrinking*, this implies that for all $t \geq t^*$, $\Delta^t > 0$.

As for the previous case, we focus our attention on how samples that allow us to estimate the marginal CDFs are collected. The same proof as in the previous section holds to show that, for times t up to time t^* , at least $\lfloor t/K \rfloor$ samples of each β_{K-k+1} are observed, on the relevant intervals: $\mathcal{I}_k^t \cup \mathcal{I}_{k+1}^t$. We now focus on the behavior of the samples in the *else* statement, willing to show a similar statement. Let $t \in [T]$ such that $\Delta^t > 0$ and denote $k'_t := t - 2K\lfloor t/2K \rfloor + 2$ and $k_t := k'_t/2$. Then notice that if k'_t is even $\mathcal{I}_{k_t}^t \subseteq (b_{k'_t+1}^t, b_{k'_t}^t]$ and if k'_t is odd $\mathcal{I}_{k_t+1}^t \subseteq (b_{k'_t+1}^t, b_{k'_t}^t]$. Therefore, after one round (i.e. $2K$ steps) the bidder observes at least one sample of β_{K-k+1} on $\mathcal{I}_k^t \cup \mathcal{I}_{k+1}^t$.

We can therefore write the two following concentration inequalities as in the proof of Algorithm 2, here on specific intervals, and in any time variant. For any $t \in [T]$ and $k \in [K]$:

The first from Massart, 1990:

$$\forall k \in [K], \mathbb{P}\left(\sup_{x \in \mathcal{I}_k^t \cup \mathcal{I}_{k+1}^t} \left| \tilde{F}_{K-k+1}^t(x) - F_{K-k+1}(x) \right| \leq \sqrt{\frac{\ln(\frac{2}{\alpha})}{2\lfloor t/2K \rfloor}}\right) \geq 1 - \alpha \quad (50)$$

The second :

$$\mathbb{P}\left(\sup_{(a,b)^2 \in \mathcal{I}_k^t \cup \mathcal{I}_{k+1}^t} \left| \frac{\tilde{F}_{K-k+1}^t(b) - \tilde{F}_{K-k+1}^t(a) - F_{K-k+1}(b) + F_{K-k+1}(a)}{\sqrt{F_k(b) - F_k(a)}} \right| \leq \epsilon(\lfloor t/2K \rfloor \alpha) \right) \geq 1 - \alpha \quad (51)$$

Where $\epsilon(t, \alpha) = \sqrt{2 \frac{\ln(\frac{2}{\alpha}) + \log(et)}{t}}$.

This leads, using the same procedure as in Theorem 2 to with probability $1 - 2K\alpha$

$$\sup_{\mathbf{b} \in \prod_{i \in [K]} \mathcal{I}_i^t} \left| \mathbb{E}_{\beta \sim \mathcal{D}}(\mathbf{b}, \beta) - \tilde{u}^t(\mathbf{b}) \right| \leq 6K^{3/2}\epsilon \quad (52)$$

994 Where $\epsilon = \sqrt{2 \frac{\ln(\frac{2}{\alpha}) + \log(e \lfloor T_{expl}/K \rfloor)}{\lfloor T_{expl}/K \rfloor}}$.

995 Using the optimality of u_{max}^t , and the fact that as long as (52) is true, \mathbf{b}^* remains in $\prod_{i \in [K]} \mathcal{I}_k^t$
 996 ensures that, for $t \geq t^*$ with proba $1 - tK\alpha$:

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}^t, \beta)] \geq \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}^*, \beta)] + 12K^{3/2}\epsilon^t \quad (53)$$

997 The cumulative regret therefore follows by summing and choosing $\alpha = \frac{1}{T^3}$.

$$R_T \leq Kt^* + \sum_{t=t^*}^T 12K^{3/2}\epsilon^t \leq Kt^* + \sum_{t=t^*}^T 12K^2 \sqrt{\frac{12 \log t}{t}} \quad (54)$$

998 All that is left is to show that we can upper bound t^* based on some instance-dependent quantity.
 999 We introduce these quantities here: we denote $B^* \subset B$ the set of maximizers of the expected util-
 1000 ity and the optimal intervals $\mathcal{I}_k^* \subset \mathbf{R}$ the smallest closed intervals such that $B^* \subset \prod_{k=1}^K \mathcal{I}_k^*$.
 1001 We also introduce the interval gap $\Delta := \min_{k \in \{2, \dots, K\}} (\min \mathcal{I}_k^* - \max \mathcal{I}_{k+1}^*)$, and the util-
 1002 ity gaps: $\delta_k^+ := \mathbb{E}[u(\mathbf{b}^*, \beta)] - \max_{\mathbf{b} \text{ s.t. } b_k \leq \min b_k^* - \Delta/2} \mathbb{E}[u(\mathbf{b}, \beta)]$. and $\delta_k^- := \mathbb{E}[u(\mathbf{b}^*, \beta)] -$
 1003 $\max_{\mathbf{b} \text{ s.t. } b_k \geq \max b_k^* + \Delta/2} \mathbb{E}[u(\mathbf{b}, \beta)]$.

1004 Notice that from equation (52), that necessarily, as soon as $12K^{3/2}\epsilon^t \leq \min((\delta_k^+, \delta_k^-)_{k \in [K]})$, for
 1005 all $k \in [K]$, the intervals \mathcal{I}_k^t are at most $\Delta/2$ wider than \mathcal{I}_k^* , and therefore $\Delta^t > 0$. Therefore,
 1006 $t^* \leq \frac{12K}{\min((\delta_k^+, \delta_k^-)_{k \in [K]})}^2$. Furthermore, note that by assumption $t^* \leq T_{expl} = K^{2/3}T^{2/3}$

1007 With these quantities, we can now write the following regret bound, which completes the proof :

$$R_T \leq \tilde{O} \left(\min \left(K^{5/3}T^{2/3}, \frac{12K}{\min((\delta_k^+, \delta_k^-)_{k \in [K]}, \max(\Delta, 0))}^2 + K^2\sqrt{T} \right) \right) \quad (55)$$

1008

□

1009 C Regret lower bounds

1010 This section contains the proofs of the main regret lower bounds for the discriminatory (first price)
 1011 and uniform price auctions, namely Lemma 3 and Theorem 3. We construct hard instances and use
 1012 information-theoretic arguments to establish fundamental limits on the achievable regret in these
 1013 settings. The techniques used here are inspired by and extend prior work, as referenced in the main
 1014 text. The *difficult instance* that we consider is significantly more straightforward to describe for the
 1015 proof of Lemma 3. We therefore begin with this one, as it should make understanding both proofs
 1016 easier.

1017 C.1 Lower bound for the first price auction

1018 Let us restate the lemma before providing the proof.

1019 **Lemma 3.** *Any algorithms for bidding in repeated first price auctions with known valuation must*
 1020 *incur a regret of $\Omega(T^{2/3})$*

1021 *Proof.* The first-price auction is a special case of the discriminatory auction when only one item is
 1022 available. The utility simply writes: $u(b, \beta) = \mathbb{1}\{b \geq \beta\} (v - b)$. Notably, when β has a random
 1023 distribution of Cumulative Distribution Function F , $\mathbb{E}[u(b, \beta)] = F(b) (v - b)$.

1024 Throughout this proof, we will use these notations, with the bid of the bidder $b \in [0, 1]$ and the
 1025 opposing bid $\beta \in [0, 1]$.

1026 To prove a lower bound on the regret of learning in the repeated first price auction, we use techniques
 1027 similar to those in Cesa-Bianchi et al., 2024 (ie, we show that in this problem we can embed $T^{1/3}$
 1028 "bandit arms" into our continuous actions space).

Outline of the proof We devise a base probability distribution of β , the adversary's bid, which is described by its CDF F . This distribution is such that the utility of the bidder is constant on a wide range of bids (ie, $[0, c]$ with c a constant). From this base probability, we derive a family of (\tilde{F}_i) , multiple slightly perturbed versions of F . The main characteristic of the perturbation is that it leaves the CDF of F unchanged outside of an interval of width $\sim T^{-1/3}$. We provide $\tilde{\mathcal{O}}(T^{1/3})$ of such perturbation. The instance we consider is a bidder such that the distribution of the opposing bids is drawn uniformly at random amongst the perturbed distribution. Quantifying the distance between the feedback distribution generated by each perturbed distribution allows us to prove the lower bound.

The base distribution \mathcal{D} is characterized by its CDF F that we define to be:

$$F(b) = \frac{1}{3(1-b)} \mathbb{1}\left\{b < \frac{1}{3}\right\} + \left(\frac{1}{4} + \frac{3b}{4}\right) \mathbb{1}\left\{b \geq \frac{1}{3}\right\} \quad (56)$$

It is easy to check that $\mathbb{E}_{\beta \sim \mathcal{D}}[u(\cdot, \beta)]$ is constant on $[0, \frac{1}{3}]$, and is decreasing on $[\frac{1}{3}, 1]$.

We define the intervals $(\mathcal{I}_i^T)_{(i \in [0, \lfloor 3T^{1/3} \rfloor])}$ on which the i^{th} perturbation takes place as follows : $\mathcal{I}_i^T := [\frac{i}{9}T^{-1/3}, \frac{i+1}{9}T^{-1/3}]$.

Let's then define the corresponding perturbations $\varepsilon_i^T : [0, 1] \rightarrow \mathbb{R}$:

$$\varepsilon_i^T(b) = \left(F\left((i+1)T^{-1/3}\right) - F(b)\right) \mathbb{1}\{b \in [i \in \mathcal{I}_i^T]\} \quad (57)$$

We can now define $F_i^T := F + \varepsilon_i^T$, the i^{th} perturbed CDF, and \mathcal{D}_i^T the corresponding distribution. It is straightforward to see that F_i^T and F coincide outside of \mathcal{I}_i^T . It is also clear, since F_i^T is constant on \mathcal{I}_i^T , that the utility is maximized in $\frac{i}{9}T^{-1/3}$.

We can now describe the instance faced by the bidder : Let $T \in \mathbb{N}$, the difficult instance is as follows:

- An index $i \in [\lfloor 3T^{1/3} \rfloor]$ is drawn uniformly at random.
- For t in $[T]$:
 - The bidder picks a bid b^t based on the past feedback Z_1, \dots, Z_{t-1}
 - The adversaries bid β^t is drawn according to \mathcal{D}_i^T , the distribution with CDF F_i^T .
 - The bidder receives the feedback $Z_t = \mathbb{1}\{b^t \geq \beta^t\}$, and the utility $u_t := (v - b^t) \mathbb{1}\{b^t \geq \beta^t\}$

From here onward, we focus on determining the minimum regret any algorithm must incur when facing the previously described instance of the first-price auction. We also restrict our analysis to *sensible* algorithms, that is, whose bids are always in $[0, \frac{1}{2}]$. These assumptions are relaxed at the end of the proof.

Some Notations: We denote \mathbb{P}^0 the probability measure over β when its drawn according to \mathcal{D} and \mathbb{P}^i the probability measure of β conditionally on the value of i (ie when its drawn according to \mathcal{D}_i^T). The notation extends naturally to expectations \mathbb{E}^0 and \mathbb{E}^i as well as regret R_T^0 and R_T^i .

For $t \in [T]$, we denote $Y_t : [0, 1] \rightarrow \{0, 1\}$ such that $Y_t(b) = \mathbb{1}\{b \geq \beta_t\}$. And Z_t is the feedback received at time t , it is a random variable and depends on the bid played by the algorithm \mathcal{A} at time t . For now, we restrict ourselves to deterministic algorithms, which allows us to write $b^t : \{0, 1\}^{t-1} \rightarrow [0, 1]$ the function which maps the feedback received up to time $t-1$ to the bid played by the algorithm \mathcal{A} at time t . With these notation, notice that $Z_t = Y_t(b^t(Z_1, \dots, Z_{t-1}))$.

With these in mind, the following lemma constitutes the core of the proof. We postpone its proofs just after the end of this proof, essentially, one uses the fact that distributions of feedbacks Z_t are too close to be distinguished.

Lemma 9. For all $i \leq \lfloor 3T^{1/3} \rfloor$ and any deterministic algorithm \mathcal{A} , denoting $N_T(i)$ the number of times algorithm \mathcal{A} produces a bid in \mathcal{I}_i^T , we have :

$$\mathbb{E}^i[N_T(i)] - \mathbb{E}^0[N_T(i)] \leq \frac{9}{40}T^{2/3}\sqrt{\mathbb{E}^0[N_T(i)]} \quad (58)$$

1070 We now use the fact that we know that in the hard instance, i is drawn uniformly at random.

1071 Averaging for $i \leq \lceil 3T^{1/3} \rceil$, we get:

$$\frac{\sum_{i=1}^{\lceil 3T^{1/3} \rceil} \mathbb{E}^i [N_T(i)]}{\lceil 3T^{1/3} \rceil} \leq \frac{\sum_{i=1}^{\lceil 3T^{1/3} \rceil} \mathbb{E}^0 [N_T(i)]}{\lceil 3T^{1/3} \rceil} + \frac{9}{40} T^{2/3} \sqrt{\frac{\sum_{i=1}^{\lceil 3T^{1/3} \rceil} \mathbb{E}^0 [N_T(i)]}{\lceil 3T^{1/3} \rceil}} \quad (59)$$

$$\leq \frac{T}{\lceil 3T^{1/3} \rceil} + \frac{9}{40} T^{2/3} \sqrt{\frac{T}{\lceil 3T^{1/3} \rceil}} \quad (60)$$

$$R_T = \frac{1}{\lceil 3T^{1/3} \rceil} \sum_{i=1}^{\lceil 3T^{1/3} \rceil} r_i \mathbb{E}^i [T - N_T(i)] \quad (61)$$

$$\geq \min_i r_i \left(T - \frac{T}{\lceil 3T^{1/3} \rceil} - \frac{9}{40} T^{2/3} \sqrt{\frac{T}{\lceil 3T^{1/3} \rceil}} \right) \quad (62)$$

$$\geq \min_i r_i \left(T - \frac{T}{3T^{1/3}} - \frac{9}{40} T^{2/3} \sqrt{\frac{T}{3T^{1/3}}} \right) \quad (63)$$

$$\geq \min_i r_i \left(T - \frac{1}{3} T^{2/3} - \frac{3\sqrt{3}}{40} T \right) \quad (64)$$

1072 Where r_i is the sub-optimality of not playing interval i when it is optimal.

$$r_i = \frac{1}{3 \left(1 - \frac{i+1}{9} T^{-1/3} \right)} \left(1 - \frac{i}{9} T^{-1/3} \right) - \frac{1}{3}$$

$$r_i = \frac{1}{9} T^{-1/3} \frac{1}{3 \left(1 - \frac{i+1}{9} T^{-1/3} \right)} \geq \frac{1}{27} T^{-1/3}$$

1073 Reinjecting this lower bound on r_i yields the desired lower bound of $\Omega(T^{2/3})$. This bound is valid
 1074 for any algorithm which issues bids in $[0, \frac{1}{2}]$, Lemma 11 generalises it to any deterministic bid. The
 1075 lower bound generalizes to random algorithms by Yao's minimax principle.

1076 □

1077 C.1.1 Technical lemma

1078 **Lemma 9.** For all $i \leq \lceil 3T^{1/3} \rceil$ and any deterministic algorithm \mathcal{A} , denoting $N_T(i)$ the number of
 1079 times algorithm \mathcal{A} produces a bid in \mathcal{I}_i^T , we have :

$$\mathbb{E}^i [N_T(i)] - \mathbb{E}^0 [N_T(i)] \leq \frac{9}{40} T^{2/3} \sqrt{\mathbb{E}^0 [N_T(i)]} \quad (58)$$

1080 *Proof.* We want to upper bound the following quantity :

$$\mathbb{E}^i [N_T(i)] - \mathbb{E}^0 [N_T(i)] = \sum_{t=1}^T \mathbb{P}^i [b^t(Z_1, \dots, Z_{t-1}) \in \mathcal{I}_i^T] - \mathbb{P}^0 [b^t(Z_1, \dots, Z_{t-1}) \in \mathcal{I}_i^T] \quad (65)$$

1081 We will upper each term of the sum independently. The algorithm can be anything deterministic;
 1082 therefore, $b^t(\cdot)$ can be any function, which is why we bound the previous quantity as follows :

$$\mathbb{P}^i [b^t (Z_1, \dots, Z_{t-1}) \in \mathcal{I}_i^T] - \mathbb{P}^0 [b^t (Z_1, \dots, Z_{t-1}) \in \mathcal{I}_i^T] \quad (66)$$

$$\leq \left\| \mathbb{P}_{(Z_1, \dots, Z_{t-1})}^i - \mathbb{P}_{(Z_1, \dots, Z_{t-1})}^0 \right\|_{TV} \quad (67)$$

$$\leq \sqrt{\frac{1}{2} KL \left(\mathbb{P}_{(Z_1, \dots, Z_{t-1})}^i ; \mathbb{P}_{(Z_1, \dots, Z_{t-1})}^0 \right)} \quad (68)$$

$$\leq \sqrt{\frac{1}{2} \left(KL \left(\mathbb{P}_{Z_1}^i ; \mathbb{P}_{Z_1}^0 \right) + \sum_{j=2}^t KL \left(\mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^i ; \mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^0 \right) \right)} \quad (69)$$

1083 Where $\mathbb{P}_{(Z_1, \dots, Z_{t-1})}^i$ is the push-forward probability of (Z_1, \dots, Z_{t-1}) .

1084 Since we focus on deterministic algorithms, $b^j(Z_1, \dots, Z_{j-1})$ conditionally on Z_1, \dots, Z_{t-1} is fixed.

$$KL \left(\mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^i ; \mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^0 \right) = \mathbb{E}^0 \left[\ln \left(\frac{\mathbb{P}^0 [Z_j = 0 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i [Z_j = 0 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0 [Z_j = 0 | Z_1, \dots, Z_{j-1}] \right. \quad (70)$$

$$\left. + \ln \left(\frac{\mathbb{P}^0 [Z_j = 1 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i [Z_j = 1 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0 [Z_j = 1 | Z_1, \dots, Z_{j-1}] \right] \quad (71)$$

$$= \mathbb{E}^0 \left[\left(\ln \left(\frac{\mathbb{P}^0 [Y_j(b^j) = 0]}{\mathbb{P}^i [Y_j(b^j) = 0]} \right) \mathbb{P}^0 [Y_j(b^j) = 0] \right. \quad (72)$$

$$\left. + \ln \left(\frac{\mathbb{P}^0 [Y_j(b^j) = 1]}{\mathbb{P}^i [Y_j(b^j) = 1]} \right) \mathbb{P}^0 [Y_j(b^j) = 1] \right) \mathbb{1} \{b^j \in \mathcal{I}_i^T\} \right] \quad (73)$$

$$\leq \frac{(F^0(b^j) - F^i(b^j))^2}{F^i(b^j) (1 - F^i(b^j))} \mathbb{P}^0 (b^j \in \mathcal{I}_i^T) \quad (74)$$

1085 Where $\mathbb{1} \{b^j \in \mathcal{I}_i^T\}$ appears because for other values of b^j , the logarithmic terms are 0.

1086 We obtain (74) from the fact that the KL between two Bernoulli \mathcal{P}, \mathcal{Q} of parameters p, q is

$$1087 \quad KL(\mathcal{P}, \mathcal{Q}) \leq \frac{(p-q)^2}{q(1-q)}.$$

1088 Since F^i is non-decreasing and $F^i(b^t) \in [\frac{1}{3}, \frac{1}{2}]$, by concavity of $x \mapsto x(1-x)$ on this interval.

$$KL \left(\mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^i ; \mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^0 \right) \leq \frac{(F^0(b^t) - F^i(b^t))^2}{\frac{2}{9}} \mathbb{P}^0 (b^t \in \mathcal{I}_i^T) \quad (75)$$

$$\leq \frac{9}{2} \left(F^0 \left(\frac{i}{9} T^{-1/3} \right) - F^0 \left(\frac{i+1}{9} T^{-1/3} \right) \right)^2 \mathbb{P}^0 (b^t \in \mathcal{I}_i^T) \quad (76)$$

1089 We can then use Lemma 10 in combination with (69) and (76) to obtain:

$$\mathbb{P}^i [b^t (Z_1, \dots, Z_{t-1}) \in \mathcal{I}_i^T] - \mathbb{P}^0 [b^t (Z_1, \dots, Z_{t-1}) \in \mathcal{I}_i^T] \leq \frac{9}{40} T^{-1/3} \sqrt{\mathbb{E}^0 [N_t(i)]} \quad (77)$$

1090 We can therefore re-inject in (65) to get :

$$\mathbb{E}^i [N_T(i)] - \mathbb{E}^0 [N_T(i)] \leq \sum_{t=1}^T \frac{9}{40} T^{-1/3} \sqrt{\mathbb{E}^0 [N_t(i)]} \leq \frac{9}{40} T^{2/3} \sqrt{\mathbb{E}^0 [N_T(i)]} \quad (78)$$

1091

□

1092 **Lemma 10.** For $i \leq \lceil 3T^{1/3} \rceil$, we have the following inequality :

$$\left(F^0\left(\frac{i}{9}T^{-1/3}\right) - F^0\left(\frac{i+1}{9}T^{-1/3}\right) \right)^2 \leq \frac{9T^{-2/3}}{400} \quad (79)$$

Proof.

$$\left(F^0\left(\frac{i}{9}T^{-1/3}\right) - F^0\left(\frac{i+1}{9}T^{-1/3}\right) \right)^2 = \left(\frac{1}{3(1 - \frac{i}{9}T^{-1/3})} - \frac{1}{3(1 - \frac{i+1}{9}T^{-1/3})} \right)^2 \quad (80)$$

$$= \left(\frac{1}{3(9 - iT^{-1/3})} - \frac{1}{3(9 - (i+1)T^{-1/3})} \right)^2 \quad (81)$$

$$= \left(\frac{3T^{-1/3}}{(9 - iT^{-1/3})(9 - (i+1)T^{-1/3})} \right)^2 \quad (82)$$

$$\leq \frac{9T^{-2/3}}{400} \quad (83)$$

1093 Where we get (83) from the fact that $i \leq \lceil 3T^{1/3} \rceil \implies (9 - iT^{-1/3})(9 - (i+1)T^{-1/3}) \geq 20$,
 1094 (given $T > 1$ but thats always the case) □

1095 **Lemma 11.** Let \mathcal{A} be any algorithm for the first price auction, their exists an algorithm \mathcal{A}^* achieving
 1096 a lower regret that only plays bids in $[0, \frac{1}{2}]$.

1097 *Proof.* For any algorithm \mathcal{A} , if at time t the algorithms picks a bid b^t outside of $[0, \frac{1}{2}]$, replace b^t
 1098 by $\tilde{b}^t = \frac{1}{4}$, ignore the feedback received from the auction and provide a simulated feedback drawn
 1099 according to $(F(b^t))$ to \mathcal{A} . Because the feedback we provide follows the same distribution, and both
 1100 are drawn independently from any other randomness (namely other β^t), on expectation, the algorithm
 1101 behaves identically for the following time steps, and utility-wise the cumulative reward receive by the
 1102 algorithms has strictly increased, since $u(\tilde{b}^t) \geq u(b^t)$.

1103

□

1104 C.2 Lower bound for the uniform price auction

1105 This section is dedicated to showing the lower bound on the rate of learning in uniform price auctions.
 1106 The lower bounds is valid when the learner is willing to acquire at least 3 items (ie $v_3 > 0$, and
 1107 $K \geq 3$).

1108 We restate the theorem :

1109 **Theorem 3.** In the uniform price auction, any algorithms must incur a regret of $\Omega(T^{2/3})$ for learning
 1110 to bid with known valuations, when valuations are such that $v_3 > 0$.

1111 *Proof.* For simplicity, we prove the lower bounds in the uniform 3-item auction. The lower bound
 1112 naturally extends to $K > 3$ by considering adversary bid distribution where for all $k \leq K - 3$,
 1113 $\mathbb{P}(\beta_k = 1) = 1$. The proof revolves around constructing a hard instance of the learning to bid
 1114 problem, and then showing one cannot avoid regret, which grows as $T^{2/3}$ in this instance.

1115 **The hard instance** In the instance of the problem, we focus on the bidder who has the following
 1116 valuations $v_1, v_2, v_3 = 1, \frac{1}{3}, \frac{1}{3}$. As before, $\beta = (\beta_1, \beta_2, \beta_3)$ is the adversary's bid and for $k \in$
 1117 $\{1, 2, 3\}$, F_k is the CDF of β_k , we also denote f_k the corresponding density. Since the quantity that
 1118 matters in our problems are the marginals CDFs (Lemma 1), we first define our hard instance by
 1119 the densities and marginals (checking then that there indeed exists a distribution with these is then
 1120 enough).

1121 The overview of the hard instance is as follows: we characterize (via the marginals) a distribution \mathcal{D}
 1122 (called base instance) of adversary bids such that there exists an interval $I \subset [0, 1]$ such that the set
 1123 of maximizers of the expected utility can be written as v_1, y, y for $y \in I$. We then build a family of
 1124 $\tilde{\mathcal{O}}(T^{1/3})$ distributions based on perturbed versions of \mathcal{D} (via additive perturbation of the marginals)

1125 for which there exists a single expected utility maximizer. The hard instance then consists in selecting
 1126 uniformly at random one of the perturbed distributions when $t = 0$.

1127 The densities of the base instance are as follows:

$$f_3(y) = \begin{cases} 101 & \text{if } y \in [0, \frac{5}{900}] \\ \frac{1}{3(v_3-y)} & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ 0 & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \\ \alpha_3 & \text{if } y \in (\frac{1}{3}, 1] \end{cases}$$

$$f_2(y) = \begin{cases} 21 & \text{if } y \in [0, \frac{5}{900}] \\ \frac{1}{3(v_2-y)} & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ 0 & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \\ \alpha_2 & \text{if } y \in (\frac{1}{3}, 1] \end{cases}$$

$$f_1(y) = \begin{cases} 1 & \text{if } y \in [0, \frac{5}{900}] \\ \frac{1}{3(v_2-y)} & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ 0 & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \\ \alpha_1 & \text{if } y \in (\frac{1}{3}, 1] \end{cases}$$

1128 Where $\alpha_1, \alpha_2, \alpha_3$ are normalisation constant. The values of α are not important because they define
 1129 the density above v_2 and v_3 , ie in regions of the bid space B that we know are suboptimal from
 1130 Lemma 7.

1131 This results in the following CDFs for $y \in [0, v_2]$:

$$F_3(y) = \begin{cases} 101y & \text{if } y \in [0, \frac{5}{900}] \\ \frac{5}{9} + F_1(y) & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ \frac{5}{9} + F_1(\frac{1}{6}) & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \end{cases}$$

$$F_2(y) = \begin{cases} 21y & \text{if } y \in [0, \frac{5}{900}] \\ \frac{1}{9} + F_1(y) & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ \frac{1}{9} + F_1(\frac{1}{6}) & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \end{cases}$$

$$F_1(y) = \begin{cases} y & \text{if } y \in [0, \frac{5}{900}] \\ \frac{1}{3} [\ln(v_2 - \frac{5}{900}) - \ln(v_2 - y)] + \frac{5}{900} & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ F_1(\frac{1}{6}) & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \end{cases}$$

1132 As before, the opponents' bids follow a distribution, $\beta \sim \mathcal{D}$. Given a sample of this random variable
 1133 $\beta = (\beta_1, \beta_2, \beta_3)$ the utility of the learner is the following function :

$$u_\beta(b_1, b_2, b_3) = \mathbb{1}\{b_1 \geq \beta_3 \geq b_2\} (v_1 - \beta_3) \\ + \mathbb{1}\{\beta_2 \geq b_2 \geq \beta_3\} (v_1 - b_2) \\ + \mathbb{1}\{b_2 \geq \beta_2 \geq b_1\} (v_1 + v_2 - 2\beta_2) \\ + \mathbb{1}\{\beta_1 \geq b_3 \geq \beta_2\} (v_1 + v_2 - 2b_3) \\ + \mathbb{1}\{b_3 \geq \beta_1\} (v_1 + v_2 + v_3 - 3\beta_1)$$

1134 For $k \in [K]$, $F_k : [0, 1] \rightarrow [0, 1]$ is the CDF of the k^{th} element of the vector β . And we denote the
 1135 corresponding density f_k . The expected utility then writes :

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u_\beta(b_1, b_2, b_3)] = v_1 (F_3(b_1) - F_3(b_2)) - \int_{b_2}^{b_1} y f_3(y) dy \\ + [F_3(b_2) - F_2(b_2)] (v_1 - b_2) \\ + (v_1 + v_2) (F_2(b_2) - F_2(b_1)) - 2 \int_{b_3}^{b_2} y f_2(y) dy \\ + [F_2(b_3) - F_1(b_3)] (v_1 + v_2 - 2b_3) \\ + (v_1 + v_2 + v_3) F_1(b_3) - 3 \int_0^{b_3} y f_1(y) dy$$

1136 We can hence compute partial derivatives :

$$\frac{\partial}{\partial b_1} \mathbb{E}_{\beta \sim \mathcal{D}} [u_{\beta}(b_1, b_2, b_3)] = (v_1 - b_1) f_3(b_1) \quad (84)$$

1137

$$\frac{\partial}{\partial b_2} \mathbb{E}_{\beta \sim \mathcal{D}} [u_{\beta}(b_1, b_2, b_3)] = F_2(b_2) - F_3(b_2) + (v_2 - b_2) f_2(b_2) \quad (85)$$

1138

$$\frac{\partial}{\partial b_3} \mathbb{E}_{\beta \sim \mathcal{D}} [u_{\beta}(b_1, b_2, b_3)] = 2(F_1(b_3) - F_2(b_3)) + (v_3 - b_3) f_1(b_3) \quad (86)$$

1139 We can therefore write the partial derivative explicitly :

$$\frac{\partial}{\partial b_2} \mathbb{E}_{\beta \sim \mathcal{D}} [u_{\beta}(b_1, b_2, b_3)] = \begin{cases} \frac{21}{3} - 101y & \text{if } y \in [0, \frac{5}{900}] \\ -\frac{1}{9} & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ -\frac{4}{9} & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \end{cases} \quad (87)$$

1140

$$\frac{\partial}{\partial b_3} \mathbb{E}_{\beta \sim \mathcal{D}} [u_{\beta}(b_1, b_2, b_3)] = \begin{cases} \frac{1}{3} - 21y & \text{if } y \in [0, \frac{5}{900}] \\ +\frac{1}{9} & \text{if } y \in (\frac{5}{900}, \frac{1}{6}] \\ -\frac{2}{9} & \text{if } y \in (\frac{1}{6}, \frac{1}{3}] \end{cases} \quad (88)$$

1141 Recall, that $\mathbf{b} \in B \implies b_2 \geq b_3$, it should be clear from (88) and (87) that the maximum is
 1142 attained for all $b_2, b_3 \in [\frac{5}{900}, \frac{1}{6}]$, such that $b_2 = b_3$. Furthermore, there exist $c \in \mathbb{R}$ such that for all
 1143 $b_2, b_3 \in [\frac{5}{900}, \frac{1}{6}]$, if $b_2 \neq b_3$ the utility at least $c(b_2 - b_3)$ sub-optimal.

1144 Let $\epsilon > 0$, for $k \in [0, \lfloor \frac{1}{7\epsilon} \rfloor]$ we denote $I_{\epsilon}^k := (\frac{5}{900} + k\epsilon, \frac{5}{900} + (k+1)\epsilon)$. We will use these
 1145 intervals to define local perturbations of the CDFs. Note that with these notation, $I_{\frac{\epsilon}{2}}^{2k}$ and $I_{\frac{\epsilon}{2}}^{2k+1}$ are
 1146 respectively the first and second half of I_{ϵ}^k .

1147 For the purpose of the lower bound, we only need to introduce a perturbation of F_2 . We therefore
 1148 define $F_{2,\epsilon}^k$ as follows :

$$F_{2,\epsilon}^k(y) = \begin{cases} F_2(y) + (y - k\epsilon) & \text{if } y \in I_{\frac{\epsilon}{2}}^{2k} \\ F_2(y) + \epsilon - (y - k\epsilon) & \text{if } y \in I_{\frac{\epsilon}{2}}^{2k+1} \\ F_2(y) & \text{otherwise} \end{cases} \quad (89)$$

1149 With the following density

$$f_{2,\epsilon}^k(y) = \begin{cases} f_2(y) + 1 & \text{if } y \in I_{\frac{\epsilon}{2}}^{2k} \\ f_2(y) - 1 & \text{if } y \in I_{\frac{\epsilon}{2}}^{2k+1} \\ f_2(y) & \text{otherwise} \end{cases} \quad (90)$$

1150 We therefore define $\mathcal{D}_{\epsilon}^k \in \mathcal{P}([0, 1]^3)$, a distribution such that if $(\beta_1, \beta_2, \beta_3) \sim \mathcal{D}_{\epsilon}^k$ then the
 1151 corresponding CDFs of $\beta_1, \beta_2, \beta_3$ are $F_1, F_{2,\epsilon}^k, F_3$.

1152 We have now defined what we need to define the hard instance. It is defined by the following process:

- 1153 • Given T , set ϵ
- 1154 • Draw k uniformly at random in $[0, \frac{1}{7\epsilon}]$
- 1155 • For t in $[T]$:
 - 1156 – The bidder picks bids b_1^t, b_2^t, b_3^t based on the past feedbacks Z_1, \dots, Z_{t-1} .
 - 1157 – The adversaries bids are drawn $\beta^t \sim \mathcal{D}_{\epsilon}^k$.
 - 1158 – The bidder receives the utility $u_{\beta^t}(b_1^t, b_2^t, b_3^t)$ and the feedback $Z_t := Z(\beta^t, \mathbf{b}^t)$

Note that ϵ depends on T , intuitively, it should behave as $T^{-\frac{1}{3}}$. The exact dependency required is determined later in the proof.

Because in this instance, the only thing a bidder can learn is $F_{2,\epsilon}^k$ or equivalently which distribution \mathcal{D}_ϵ^k it faces, we will discard in the analysis the part of the feedback which doesn't provide information about the second marginal distribution. Reducing to $Z_t := Z(\beta^t, \mathbf{b}^t) = (\mathbb{1}\{\beta_2^t \leq b_3\}, \mathbb{1}\{\beta_2^t \in (b_3, b_2]\}\beta_2^t, \mathbb{1}\{\beta_2^t > b_2\})$ (this is directly implied by Lemma 2). To provide a simpler analysis, it is convenient to work with a one dimensional feedback, notice that the previously described feedback is equivalent (one can define a bijection) with $Z(\beta^t, \mathbf{b}^t) = \max(\mathbb{1}\{\beta_2^t \in (b_3, b_2]\}\beta_2^t, \mathbb{1}\{\beta_2^t > b_2\})$, this is the formulation we use in the following proofs. We also only consider algorithms that, at all times t , bid $b_1^t = v_1$. This is without loss of generality, as not doing so would only add additional regret and provide no additional useful feedback.

We also only consider deterministic algorithms, the lower bound then generalizes to randomized algorithms by Yao's minimax principle.

Our final reduction is the following: we consider in the following analysis only algorithms such that $\sum_{t=1}^T b_2^t - b_3^t = o(T^{2/3})$. This is without loss of generality: since the optimal bids b_2^* and b_3^* are always equal and belong to I_ϵ^k . The expression of the derivatives (87) and (88) ensures that for all bids with b_2 and b_3 in the interval $[5/900, \frac{1}{6}]$ the algorithms incurs a regret which scales with $b_2 - b_3$. Furthermore, if either b_2 or b_3 is not in $[5/900, \frac{1}{6}]$, the algorithms incur an additional regret term of constant size. Therefore any algorithms such that we do not have $\sum_{t=1}^T b_2^t - b_3^t = o(T^{2/3})$ incurs a regret which scales as $\Omega(T^{2/3})$ in this instance.

Some Notations: We denote \mathbb{P}^0 a probability measure over β when the CDFs are respectively F_1, F_2, F_3 and \mathbb{P}^k a probability measure over β conditionally on the value of k (ie when its distributed according to \mathcal{D}_ϵ^k). The notation extends naturally to expectations \mathbb{E}^0 and \mathbb{E}^k .

Lemma 12. For all $i \leq \lceil 3T^{-1/3} \rceil$ and any deterministic algorithm \mathcal{A} , denoting $N_T(i)$ the number of times algorithm \mathcal{A} produces a bid such that $(b_3^t, b_2^t] \cap \mathcal{I}_i^T \neq \emptyset$, we have :

$$\mathbb{E}^i [N_T(i)] - \mathbb{E}^0 [N_T(i)] \leq 10\epsilon T \sqrt{\mathbb{E}^0 [N_T(i)]} \quad (91)$$

Averaging for $i \leq \lfloor \frac{1}{7\epsilon} \rfloor$, we get:

$$\frac{\sum_{i=1}^{\lfloor \frac{1}{7\epsilon} \rfloor} \mathbb{E}^i [N_T(i)]}{\lfloor \frac{1}{7\epsilon} \rfloor} \leq \frac{\sum_{i=1}^{\lfloor \frac{1}{7\epsilon} \rfloor} \mathbb{E}^0 [N_T(i)]}{\lfloor \frac{1}{7\epsilon} \rfloor} + 10\epsilon T \sqrt{\frac{\sum_{i=1}^{\lfloor \frac{1}{7\epsilon} \rfloor} \mathbb{E}^0 [N_T(i)]}{\lfloor \frac{1}{7\epsilon} \rfloor}} \quad (92)$$

$$\leq \frac{T}{\lfloor \frac{1}{7\epsilon} \rfloor} + 10\epsilon T \sqrt{\frac{T}{\lfloor \frac{1}{7\epsilon} \rfloor}} \quad (93)$$

Let us denote $\tilde{N}_T(i)$ the number of times algorithm \mathcal{A} produces a bid such that $(b_3^t, b_2^t] \cap \mathcal{I}_i^T \neq \emptyset$, and i is the smallest such index. We can recover for this one that $\sum_{i=1}^{\lfloor \frac{1}{7\epsilon} \rfloor} \mathbb{E}^0 [\tilde{N}_T(i)] = T$.

Notice that because we focus on algorithms such that $\sum_{t=1}^T b_2^t - b_3^t = o(T^{2/3})$, the difference $\sum_{i=1}^{\lfloor \frac{1}{7\epsilon} \rfloor} \mathbb{E}^0 [\tilde{N}_T(i)] - \sum_{i=1}^{\lfloor \frac{1}{7\epsilon} \rfloor} \mathbb{E}^0 [N_T(i)] = o\left(\frac{T^{2/3}}{\epsilon}\right)$. We continue the computation without this additional term of size $o\left(\frac{T^{2/3}}{\epsilon}\right)$ as it is clear the bound remains unchanged from this addition (it would only add to a term of order $o\left(\frac{T^{2/3}}{\epsilon}\right)$ in the right hand side of the inequality).

$$R_T \geq \frac{1}{\lfloor \frac{1}{7\epsilon} \rfloor} \sum_{i=1}^{\lfloor \frac{1}{7\epsilon} \rfloor} r_i \mathbb{E}^i [T - N_T(i)] \quad (94)$$

$$\geq \min_i r_i \left(T - \frac{T}{\lfloor \frac{1}{7\epsilon} \rfloor} - 10\epsilon T \sqrt{\frac{T}{\lfloor \frac{1}{7\epsilon} \rfloor}} \right) \quad (95)$$

Where r_i is the sub-optimality of not playing bids in interval i when it is optimal. Using (87) and (88) we obtain

$$r_i \geq \frac{\epsilon}{12}$$

1191 Choosing $\epsilon = \frac{1}{700}T^{-1/3}$ yields the desired lower bound of $\Omega(T^{2/3})$

1192 □

1193 C.2.1 Technical lemma

1194 **Lemma 12.** *For all $i \leq \lceil 3T^{-1/3} \rceil$ and any deterministic algorithm \mathcal{A} , denoting $N_T(i)$ the number*
 1195 *of times algorithm \mathcal{A} produces a bid such that $(b_3^t, b_2^t] \cap \mathcal{I}_i^T \neq \emptyset$, we have :*

$$\mathbb{E}^i [N_T(i)] - \mathbb{E}^0 [N_T(i)] \leq 10\epsilon T \sqrt{\mathbb{E}^0 [N_T(i)]} \quad (91)$$

1196 *Proof.* For convenience and to have shorter expression, we denote $(b_3^t, b_2^t](\cdot) := (b_3^t(\cdot), b_2^t(\cdot)]$.

1197 We want to upper bound the following quantity :

$$\begin{aligned} \mathbb{E}^i [N_T(i)] - \mathbb{E}^0 [N_T(i)] &= \sum_{t=1}^T \mathbb{P}^i [(b_3^t, b_2^t](Z_1, \dots, Z_{t-1}) \cap \mathcal{I}_i^T \neq \emptyset] \\ &\quad - \mathbb{P}^0 [(b_3^t, b_2^t](Z_1, \dots, Z_{t-1}) \cap \mathcal{I}_i^T \neq \emptyset] \end{aligned} \quad (96)$$

1198 We will upper each term of the sum independently. The algorithm can be anything deterministic,
 1199 therefore $(b_3^t, b_2^t](\cdot)$ can be any function, which is why we bound the previous quantity as follows :

$$\mathbb{P}^i [(b_3^t, b_2^t](Z_1, \dots, Z_{t-1}) \cap \mathcal{I}_i^T \neq \emptyset] - \mathbb{P}^0 [(b_3^t, b_2^t](Z_1, \dots, Z_{t-1}) \cap \mathcal{I}_i^T \neq \emptyset] \quad (97)$$

$$\leq \left\| \mathbb{P}_{(Z_1, \dots, Z_{t-1})}^i - \mathbb{P}_{(Z_1, \dots, Z_{t-1})}^0 \right\|_{TV} \quad (98)$$

$$\leq \sqrt{\frac{1}{2} KL \left(\mathbb{P}_{(Z_1, \dots, Z_{t-1})}^i ; \mathbb{P}_{(Z_1, \dots, Z_{t-1})}^0 \right)} \quad (99)$$

$$\leq \sqrt{\frac{1}{2} \left(KL \left(\mathbb{P}_{Z_1}^i ; \mathbb{P}_{Z_1}^0 \right) + \sum_{j=2}^t KL \left(\mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^i ; \mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^0 \right) \right)} \quad (100)$$

1200 Where $\mathbb{P}_{(Z_1, \dots, Z_{t-1})}^i$ is the push-forward probability of (Z_1, \dots, Z_{t-1}) .

1201 Since we focus on deterministic algorithms, $b^j(Z_1, \dots, Z_{j-1})$ conditionally on Z_1, \dots, Z_{j-1} is fixed.

1202 We can therefore write the following equation to bound terms in the square root of (100).

$$KL \left(\mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^i ; \mathbb{P}_{Z_j | Z_1, \dots, Z_{j-1}}^0 \right) = \mathbb{E}^0 \left[\ln \left(\frac{\mathbb{P}^0 [Z_j = 0 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i [Z_j = 0 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0 [Z_j = 0 | Z_1, \dots, Z_{j-1}] \right. \quad (101)$$

$$\left. + \int_0^1 \ln \left(\frac{p^0 [Z_j = u | Z_1, \dots, Z_{j-1}]}{p^i [Z_j = u | Z_1, \dots, Z_{j-1}]} \right) p^0 [Z_j = u | Z_1, \dots, Z_{j-1}] du \right. \quad (102)$$

$$\left. + \ln \left(\frac{\mathbb{P}^0 [Z_j = 1 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i [Z_j = 1 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0 [Z_j = 1 | Z_1, \dots, Z_{j-1}] \right] \quad (103)$$

1203 Note that this can be written as such because Z_j takes values 1 and 0 with positive probability, it can
 1204 also take values in b_3^j, b_2^j and its distribution admits a density because the distribution of β_2 admits a

1205 density. We continue forward in our calculation, using the fact that the distribution of β_2 and hence
 1206 Z_j between \mathbb{P}^0 and \mathbb{P}^i only differs on \mathcal{I}_ϵ^i .

$$KL\left(\mathbb{P}_{Z_j|Z_1,\dots,Z_{j-1}}^i; \mathbb{P}_{Z_j|Z_1,\dots,Z_{j-1}}^0\right) = \mathbb{E}^0 \left[\ln \left(\frac{\mathbb{P}^0[Z_j = 0 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i[Z_j = 0 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0[Z_j = 0 | Z_1, \dots, Z_{j-1}] \right. \\ \left. (104) \right.$$

$$+ \sum_{k=0}^{\frac{1}{\epsilon}} \int_{I_\epsilon^k} \ln \left(\frac{p^0[Z_j = u | Z_1, \dots, Z_{j-1}]}{p^i[Z_j = u | Z_1, \dots, Z_{j-1}]} \right) p^0[Z_j = u | Z_1, \dots, Z_{j-1}] du \\ (105)$$

$$+ \ln \left(\frac{\mathbb{P}^0[Z_j = 1 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i[Z_j = 1 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0[Z_j = 1 | Z_1, \dots, Z_{j-1}] \Big] \\ (106)$$

$$= \mathbb{E}^0 \left[\ln \left(\frac{\mathbb{P}^0[Z_j = 0 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i[Z_j = 0 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0[Z_j = 0 | Z_1, \dots, Z_{j-1}] \right. \\ (107)$$

$$+ \int_{I_\epsilon^i} \ln \left(\frac{p^0[Z_j = u | Z_1, \dots, Z_{j-1}]}{p^i[Z_j = u | Z_1, \dots, Z_{j-1}]} \right) p^0[Z_j = u | Z_1, \dots, Z_{j-1}] du \\ (108)$$

$$+ \ln \left(\frac{\mathbb{P}^0[Z_j = 1 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i[Z_j = 1 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0[Z_j = 1 | Z_1, \dots, Z_{j-1}] \Big] \\ (109)$$

$$= \mathbb{E}^0 \left[\left(\ln \left(\frac{\mathbb{P}^0[Z_j = 0 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i[Z_j = 0 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0[Z_j = 0 | Z_1, \dots, Z_{j-1}] \right. \right. \\ (110)$$

$$+ \ln \left(\frac{\mathbb{P}^0[Z_j = 1 | Z_1, \dots, Z_{j-1}]}{\mathbb{P}^i[Z_j = 1 | Z_1, \dots, Z_{j-1}]} \right) \mathbb{P}^0[Z_j = 1 | Z_1, \dots, Z_{j-1}] \Big) \mathbf{1} \left\{ I_\epsilon^i \cap [b_3^j, b_2^j] \neq \emptyset \right\} \\ (111)$$

$$+ \int_{I_\epsilon^i} \ln \left(\frac{p^0[Z_j = u | Z_1, \dots, Z_{j-1}]}{p^i[Z_j = u | Z_1, \dots, Z_{j-1}]} \right) p^0[Z_j = u | Z_1, \dots, Z_{j-1}] du \mathbf{1} \left\{ I_\epsilon^i \cap [b_3^j, b_2^j] \neq \emptyset \right\} \Big] \\ (112)$$

1207 Let us denote $Y_j(\mathbf{b}^j)$ is the function which has values 1 when $\beta_2^j \leq b_3^j$ and value 0 when $\beta_2^j > b_2^j$.
 1208 And let us use the inequality between the KL divergence and χ^2 distance.

$$KL\left(\mathbb{P}_{Z_j|Z_1,\dots,Z_{j-1}}^i; \mathbb{P}_{Z_j|Z_1,\dots,Z_{j-1}}^0\right) \leq \mathbb{E}^0 \left[\left(\ln \left(\frac{\mathbb{P}^0[Y_j(b^j) = 0]}{\mathbb{P}^i[Y_j(b^j) = 0]} \right) \mathbb{P}^0[Y_j(b^j) = 0] \right. \right. \\ + \ln \left(\frac{\mathbb{P}^0[Y_j(b^j) = 1]}{\mathbb{P}^i[Y_j(b^j) = 1]} \right) \mathbb{P}^0[Y_j(b^j) = 1] \Big) \mathbf{1} \left\{ I_\epsilon^i \cap [b_3^j, b_2^j] \neq \emptyset \right\} \\ + \int_{I_\epsilon^i} \frac{(f_2(u) - f_{2,\epsilon}^i(u))^2}{f_{2,\epsilon}^i(u)} du \mathbf{1} \left\{ I_\epsilon^i \cap (b_3^j, b_2^j] \neq \emptyset \right\} \Big] \\ \leq \left(\max_{x \in I_\epsilon^i} \frac{(F_2^0(x) - F_2^i(x))^2}{F_2^i(x)(1 - F_2^i(x))} + 6\epsilon^2 \right) \mathbb{P}^0 \left(I_\epsilon^i \cap (b_3^j, b_2^j] \neq \emptyset \right) \\ \leq \epsilon^2(81 + 6) \mathbb{P}^0 \left(I_\epsilon^i \cap (b_3^j, b_2^j] \neq \emptyset \right) \\ (113)$$

1209 We can now use (100) and (113) to get :

$$\mathbb{P}^i \left[(b_3^t, b_2^t) (Z_1, \dots, Z_{t-1}) \cap \mathcal{I}_t^T \neq \emptyset \right] - \mathbb{P}^0 \left[(b_3^t, b_2^t) (Z_1, \dots, Z_{t-1}) \cap \mathcal{I}_t^T \neq \emptyset \right] \leq 10\epsilon \sqrt{\mathbb{E}^0 [N_t(i)]} \\ (114)$$

1210 Reinjecting in (96) allows us to recover :

$$\mathbb{E}^i [N_T(i)] - \mathbb{E}^0 [N_T(i)] \leq 10T\epsilon\sqrt{\mathbb{E}^0 [N_T(i)]} \quad (115)$$

1211

□

1212 D Instance specific proofs

1213 Here, we present proofs and results that are specific to certain instances of the auction problem,
1214 such as Δ -separated distributions and the i.i.d. unit-demand adversary setting. These specialized
1215 analyses provide further insight into the behavior of learning algorithms under particular structural
1216 assumptions.

1217 We begin by providing the proof of Lemma 4, which we first restate for convenience.

1218 **Lemma 4.** *Learning in a discriminatory or uniform auction when the adversary's distribution is Δ*
1219 *separated, with $\Delta < \frac{1}{2K}$ can be achieved with respectively regret of $\Omega(T^{2/3})$ and $\mathcal{O}(\sqrt{T})$.*

1220 *Proof.* We first provide the proof of the lower bound in the discriminatory auction. We begin by
1221 providing the necessary tools for the reader to understand the reduction of the case of Δ -separated
1222 distribution with $\Delta \leq \frac{1}{2K}$. Our first remark is the following: while the proof of Lemma 3 makes
1223 use of an adversary distribution whose support is $[0, 1]$, one can easily provide another difficult
1224 instance for adversary bids on $[0, \frac{1}{2}]$. Indeed, consider, with the notations used in the proof, $\tilde{F}(b) =$
1225 $\max(1, F(2b))$, and a valuation $v = \frac{1}{2}$. Then the bidder faces a problem identical to the original one,
1226 scaled down on $[0, \frac{1}{2}]$, and the lower bound holds.

1227 With this in mind, consider the multi-unit discriminatory auction, when the bids of the adversary are
1228 such that $\beta := (1, 1 - \frac{1}{2K}, \dots, 1 - \frac{K-2}{2K}, b)$ where b follows the distribution corresponding to the
1229 scaled down first price auction mentioned above. Note that this constitutes a Δ -separated distribution
1230 for $\Delta \geq \frac{1}{2K}$. Remember that for any bidder such that $v_1 \leq \frac{1}{2}$, it is strategy dominant to only emit
1231 bids smaller or equal to $\frac{1}{2}$. Therefore, this reduces to a first-price auction problem, and as such, the
1232 lower bound of $\Omega(T^{2/3})$ extends to this setting for the discriminatory auction.

1233 We now move to providing an upper bound for the uniform price auction. Our main argument is
1234 based on the following property of the utility: let $\mathbf{b} \in B$, then the events which identify which bid
1235 is the $K+1$, formalized as follow $(\{b_i \geq \beta_{K-i+1} > b_{i+1}\}, \{\beta_{K-i} > b_{i+1} \geq \beta_{K-i+1}\})_{i \in [K]}$ are
1236 disjoint. The utility writes :

$$\begin{aligned} u(\mathbf{b}, \beta) = & \sum_{i=1}^K \mathbb{1}_{\{b_i \geq \beta_{K-i+1} > b_{i+1}\}} \left(\sum_{j=1}^K v_j - \beta_{K-i+1} \right) \\ & + \sum_{i=1}^K \mathbb{1}_{\{\beta_{K-i} > b_{i+1} \geq \beta_{K-i+1}\}} \left(\sum_{j=1}^K v_j - b_{i+1} \right) \end{aligned} \quad (116)$$

1237 Notice when taking expectation of the utility that, since the \mathcal{I}_k which are the support of each k^{th}
1238 marginal distribution of \mathcal{D} are disjoint, only one of the pairs $\mathbb{P}(b_i \geq \beta_{K-i+1} > b_{i+1}), \mathbb{P}(\beta_{K-i} >$
1239 $b_{i+1} \geq \beta_{K-i+1})$ can be nonzero. Besides notice that the term involving $\mathbb{P}(b_i \geq \beta_{K-i+1} > b_{i+1})$ is
1240 always bigger. Therefore, define $i^* := \operatorname{argmax}_{i \in [K]} \mathbb{E} \left[\sum_{j=1}^K v_j - \beta_{K-i+1} \right]$ optimal bids are all
1241 the bids which ensure $\mathcal{I}_{K-i+1} \subseteq (b_{i+1}, b_i]$.

1242 Now notice that given the intervals, there exists \mathbf{b} such that $u(\mathbf{b}, \beta) =$
1243 $\max_{i \in [K]} \left(\mathbb{E} \left[\sum_{j=1}^K v_j - \beta_{K-i+1} \right] \right)$, and this is the maximum utility since for the other
1244 events, $b_{i+1} \geq \beta_{K-i+1}$.

1245 Let us now suppose that, in Algorithm 3, when there is a tie between several maximizers, the
1246 algorithm chooses, if possible, bids with values 0 or 1.

1247 With this in mind and the fact that the estimates built for the algorithm $(\tilde{F}_k)_{k \in [K]}$ necessarily
1248 correspond to a Δ separated distribution, it should be clear that, at any time t , the algorithms only

1249 plays bids composed of 0 and 1. Because there are only K such bids, the bidder essentially plays a
 1250 finite arms bandit problem.

1251 Furthermore, the concentration bounds developed in the proof of Theorem 4 ensure with probability
 1252 $1 - 2K\alpha$

$$\sup_{\mathbf{b} \in \prod_{i \in [K]} \mathcal{I}_k^t} \left| \mathbb{E}_{\beta \sim \mathcal{D}} (\mathbf{b}, \beta) - \tilde{u}^t(\mathbf{b}) \right| \leq 6K^{3/2}\epsilon \quad (117)$$

1253 And therefore, if we denote $\Delta_i^* := \min_{i \in [K], i \neq i^*} \mathbb{E} \left[\sum_{j=1}^K v_j - \beta_{K-i^*+1} \right] -$
 1254 $\mathbb{E} \left[\sum_{j=1}^K v_j - \beta_{K-i+1} \right]$, as soon as $6K^{3/2}\epsilon \leq \Delta_i^*$, the bidder only plays the optimal arm.

1255 The regret therefore scales as $\tilde{O}(\sqrt{T})$ in an instance dependent fashion (depending on Δ_i^*).

1256 \square

1257 We now move to providing the postponed proofs for the unit-demand symmetric case. We begin by
 1258 proving our concentration result Lemma 5.

1259 **Lemma 5.** *Let $t, k \in \mathbb{N}$, and let X_j^i for $i \in [t], j \in [k]$ be i.i.d. samples from a distribution \mathcal{P} with*
 1260 *cdf F . With probability $1 - \alpha$:*

$$\hat{F}_{k \rightarrow k'}^t(x) - \alpha_{k,k'}\epsilon \leq F_{k'}(x) \leq \hat{F}_{k \rightarrow k'}^t(x) + \alpha_{k,k'}\epsilon \quad (8)$$

1261 Where $\epsilon = \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}}$ and $\alpha_{k,k'}$ is a constant which only depends on k, k' and N .

1262 *Proof.* Let $k, k' \in [K]^2$, and $t \in [T]$ such that one has observed $X_{(k)}^1, \dots, X_{(k)}^t$, the k^{th} ordered
 1263 statistic t times. We can apply the DKW inequality from Massart, 1990 to bound the distance from
 1264 its empirical CDF to its actual CDFs denoted $\hat{F}_{(k)}$ and $F_{(k)}$. This yields the following

$$\mathbb{P} \left(\sup_{x \in [0,1]} |\hat{F}_{(k)}^t(x) - F_{(k)}(x)| \leq \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}} \right) \geq 1 - \alpha \quad (118)$$

1265 We denote $\epsilon = \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t}}$. Using (7), one has that for F the CDF corresponding to \mathcal{P} , $P_k(F) = F_{(k)}$,
 1266 where P_k is the polynomial defined by (7). This polynomial is strictly increasing on $(0, 1)$, we can
 1267 therefore get the following inequalities with probability $1 - \alpha$, $\forall x \in [0, 1]$:

$$P_k^{-1}(F_{(k)}(x) + \epsilon) \leq F(x) \leq P_k^{-1}(F_{(k)}(x) - \epsilon) \quad (119)$$

1268 Using (7) once again allows us to get with probability $1 - \alpha$, $\forall x \in [0, 1]$:

$$P_{k'}(P_k^{-1}(F_{(k)}(x) + \epsilon)) \leq F_{(k')}(x) \leq P_{k'}(P_k^{-1}(F_{(k)}(x) - \epsilon)) \quad (120)$$

1269 We can then obtain concentration inequality in the usual formulation, with probability $1 - \alpha$:

$$|P_{k'}(P_k^{-1}(F_{(k)}(x))) - F_{(k')}(x)| \leq \epsilon \alpha_{k,k'} \quad (121)$$

1270 Where $\alpha_{k,k'}$ only depends on the upper bound of the derivatives of both $P_{k'}$ and P_k^{-1} on $(0,1)$. \square

1271 With this, we now have the necessary tools to prove the Theorem 5. Let us first restate the theorem :
 1272 **Theorem 5.** *When facing i.i.d. adversaries in the uniform auction with bandit feedback, Algorithm 5*
 1273 *guarantees the regret is upper bounded by $\tilde{O}(K^{5/2}\sqrt{T})$.*

1274 *Proof.* As noted in Lemma 7, it is known in advance that bidding such that $b_1 = v_1$ is optimal.
 1275 Since this is compatible with our Algorithm 5, we assume that amongst the maximizers of \tilde{u}^t , the
 1276 algorithm always chooses one such that $b_1 = v_1$. Let $t \in [T]$, recall that $\tilde{u}^t(\cdot) =: U_u((\bar{F}_k)_{k \in [K]}, \cdot)$
 1277 and $\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] := U_u((F_k)_{k \in [K]}, \cdot)$. Notice that the observations formalized in Lemma 2 ensure

1278 that for any $x \in [0, v_1]$, at time t , there exists k' such that $t^{k'(x)} \geq t/K$. We can therefore leverage
 1279 Lemma 5 to obtain the following for $k \in [K]$ with probability $1 - \alpha$:

$$\sup_{x \in [0, v_1]} |\bar{F}_k^t(x) - F_k(x)| \leq \epsilon^t \gamma \quad (122)$$

1280 Where $\gamma = \max_{k, k' \in [K]^2} \alpha_{k, k'}$ and $\epsilon^t := \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t/K}}$.

1281 Note that we can ensure the bound is true across all $k \in [K]$ with probability $1 - K\alpha$

1282 Reinjecting directly and using the formula from (15), one gets the following : with probability $1 - K\alpha$
 1283 :

$$\sup_{\substack{\mathbf{b} \in B \\ b_1 = v_1}} \left| \bar{u}^t(\mathbf{b}) - \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}, \beta)] \right| \leq 3K^2 \epsilon^t \gamma \quad (123)$$

1284 Where $\gamma = \max_{k, k' \in [K]^2} \alpha_{k, k'}$ and $\epsilon^t := \sqrt{\frac{\ln(\frac{2}{\alpha})}{2t/K}}$.

1285 Using the optimality of \mathbf{b}^t with respect to $\bar{u}^t(\mathbf{b})$, we get that with probability $1 - K\alpha$:

$$\mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}^t, \beta)] \geq \mathbb{E}_{\beta \sim \mathcal{D}} [u(\mathbf{b}^*, \beta)] - 6K^2 \epsilon^t \gamma \quad (124)$$

1286 Summing over $t \in [T]$ and taking $\alpha = \frac{1}{T^3}$ yields the desired regret bound.

1287 □

1288 We finally provide the reduction necessary to obtain the lower bound of Lemma 6, which we restate
 1289 first :

1290 **Lemma 6.** *When facing i.i.d. adversaries in the uniform auction with bandit feedback, any learning*
 1291 *algorithm must incur at least a regret of $\Omega(\sqrt{T})$.*

1292 *Proof.* We focus on showing that we can build a reduction to one of the *hard instances* used in the
 1293 proof of the full-information feedback lower bound in Brânzei et al., 2023. We focus on the 2 item
 1294 auction.

1295 Let $p = \frac{2}{3}$, if we sample two binomial random variable X with probability p multiply them by $\frac{2}{3}$ and
 1296 denote that β then with probability $\frac{1}{9}$, $\beta = (0, 0)$, with probability $\frac{4}{9}$, $\beta = (\frac{2}{3}, 0)$ and with probability
 1297 $\frac{4}{9}$, $\beta = (\frac{2}{3}, \frac{2}{3})$. First notice that having with probability $\frac{1}{9}$, $\beta = (0, 0)$ does not change the best
 1298 strategy (since the utility $u(\cdot, (0, 0))$ is constant). Then notice that the probability of $\beta = (\frac{2}{3}, 0)$ and
 1299 $\beta = (\frac{2}{3}, \frac{2}{3})$ have derivatives in $p = \frac{2}{3}$ of opposing signs which ensures we can create perturbations of
 1300 size δ in opposing direction by small changes in p . These are sufficient to ensure one can reproduce
 1301 the same proofs as in Brânzei et al., 2023 Theorem 4 and lower bound the regret in this setting by
 1302 $\Omega(\sqrt{T})$. □