Reserve Pricing in Repeated Second-Price Auctions with Strategic Bidders: SUPPLEMENTARY MATERIALS

Alexey Drutsa*

12 July 2020

Contents

A Missed proofs

A.1 Missed proofs from Section 3

A.1.1 Proof of Lemma 1

Proof. Let $\mathcal{I}^m = \{t_i^m\}_{i=1}^{I^m}$ be the set of rounds in which the bidder m is not eliminated by a barrage reserve pricing. Therefore, we have decomposition of the sequence of all rounds into the union of these sets: $\{1, \ldots, T\} = \cup_{m\in\mathbb{M}}\mathcal{I}^m$. Note that we also have a splitting in periods $\{1, \ldots, T\} = \cup_{i=1}^{I}\mathcal{T}_i$ and the intersection $\mathcal{I}^m \cap \mathcal{T}_i = \{t_i^m\}$ for $m \in \mathbb{M}$, $i = 1, \ldots, I^m$.

So, formally, we have

$$\mathrm{SReg}(T, \mathcal{A}, \mathbf{v}, \boldsymbol{\gamma}, \boldsymbol{\beta}) = \mathrm{Reg}\big(T, \mathcal{A}, \mathbf{v}, \mathring{\mathbf{b}}_{1:T}(T, \mathcal{A}, \mathbf{v}, \boldsymbol{\gamma}, \boldsymbol{\beta})\big) = \sum_{t=1}^{T}(\overline{v} - \overline{a}_t\overline{p}_t) = \sum_{m\in\mathbb{M}}\sum_{i=1}^{I^m}(\overline{v} - \overline{a}_{t_i^m}\overline{p}_{t_i^m}), \quad (A.1)$$

*Yandex, Moscow, Russia (www.yandex.com); Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russia; e-mail: adrutsa@yandex.ru.

where the two first identities follow from definitions, while the latter one is just a change of the order of summation (since $\{1,\dots,T\} = \cup_{m\in\mathbb{M}}\mathcal{I}^m = \cup_{m\in\mathbb{M}}\{t_i^m\}_{i=1}^{I^m}$). The terms in the sum could be decomposed in the following way: $\overline{v} - \overline{a}_{t_i^m}\overline{p}_{t_i^m} = \overline{v} - v^m + v^m - \overline{a}_{t_i^m}\overline{p}_{t_i^m}$. Also note, since, in each round t_i^m, the bidders $\mathbb{M}^{-m}$ are eliminated by a barrage reserve price, then the allocation indicator $\overline{a}_{t_i^m}$ and the transferred payment $\overline{p}_{t_i^m}$ depend only on the behavior of the bidder m in this round, i.e., $\overline{a}_{t_i^m} = a^m_{t_i^m}$, $\overline{p}_{t_i^m} = \overline{p}^m_{t_i^m}$, and, if $a^m_{t_i^m} = 1$, $\overline{p}_{t_i^m} = \overline{p}^m_{t_i^m} = p^m_{t_i^m}$. So, we can continue Eq. (A.1):

$$\begin{aligned}
\mathrm{SReg}(T,\mathcal{A},\mathbf{v},\boldsymbol{\gamma},\boldsymbol{\beta}) &= \sum_{m\in\mathbb{M}}\sum_{i=1}^{I^m}(\overline{v} - v^m + v^m - \overline{a}_{t_i^m}\overline{p}_{t_i^m}) \\
&= \sum_{m=1}^{M}\sum_{i=1}^{I^m}(\overline{v} - v^m) + \sum_{m=1}^{M}\sum_{i=1}^{I^m}(v^m - a^m_{t_i^m}p^m_{t_i^m}) \\
&= \sum_{m=1}^{M} I^m(\overline{v} - v^m) + \sum_{m=1}^{M}\mathrm{Reg}^m(\mathcal{I}^m,\mathcal{A}^m,v^m,\mathring{b}^m_{1:T}), \\
&= \mathrm{SReg}^{\mathrm{dev}}(T,\mathcal{A},\mathbf{v},\boldsymbol{\gamma},\boldsymbol{\beta}) + \mathrm{SReg}^{\mathrm{ind}}(T,\mathcal{A},\mathbf{v},\boldsymbol{\gamma},\boldsymbol{\beta}).
\end{aligned} \tag{A.2}$$

□

A.1.2 Proof of Proposition 1

Proof. The idea of the proof is close to the ones of propositions in [6, 2, 3, 4, 5, 8]. Let t be the round in which the bidder m reaches the node $\mathfrak{n}$ and rejects his reserve price p_t^m, which is equal to $p_t^m = \mathrm{p}(\mathfrak{n})$ by the construction of the algorithm $\mathrm{div}_M(\langle\mathcal{A}_1\rangle,\mathrm{sr})$. Note that, in the round t, all other bidders $\mathbb{M}^{-m}$ are eliminated by a barrage price and the reserve prices set by the div-algorithm $\mathrm{div}_M(\langle\mathcal{A}_1\rangle,\mathrm{sr})$ depend only on $\mathbf{a}_{1:T}$ (because $\mathcal{A}_1 \in \mathbf{A}^{\mathrm{RPPA}}$ and $\mathrm{sr}: \mathbb{M}\times\mathfrak{T}(\mathcal{A}_1)^M \to \mathtt{bool}$). Therefore, it is easy to see that, for any strategy σ, the expected future surplus $\mathrm{Sur}_{t:T}(\mathcal{A},\gamma_m,v^m,h_t^m,\beta^m,\sigma)$ of the bidder m as a function of the bid $b_t^m = \sigma(h_t^m)$ in the round t depends, in fact, only on the binary decision $a_t^m = \mathbb{I}_{\{b_t^m \ge p_t^m\}}$: more formally, the expected surplus is constant when the bid b_t^m is changed within $\{b_t^m \ge p_t^m\}$ and is constant when the bid b_t^m is changed within $\{b_t^m < p_t^m\}$. Moreover, since the buyers are divided (in the whole game) and $\mathcal{A}_1 \in \mathbf{A}^{\mathrm{RPPA}}$, if two strategies σ' and $\sigma'' \in \mathfrak{S}_T$ do not differ in their binary output, i.e., $\mathbb{I}_{\{\sigma'(h)\ge p_t^m\}} = \mathbb{I}_{\{\sigma''(h)\ge p_t^m\}}\ \forall h \in \mathbb{H}_{1:T}$, then they have the same future discounted surplus. Hence, any strategy can be treated as a map to binary decisions $\{0,1\}$ (instead of $\mathbb{R}_+$). Let $\hat{\sigma}_a$ denote an optimal strategy among all possible strategies in which the binary decision a_t^m in the round t is $a \in \{0,1\}$, i.e., $\mathbb{I}_{\{\hat{\sigma}_a(h_t^m)\ge p_t^m\}} = a$ and $\hat{\sigma}_a$ maximizes

$$\mathbb{E}[\sum_{s=t}^{T}\gamma_m^{s-1}\overline{a}_s^m(v^m - \overline{p}_s^m) \mid h_t^m, a_t^m = a, \sigma, \beta^m].$$

Given a strategy $\sigma \in \mathfrak{S}_T$, let us denote the future expected surplus when following this strategy by $S_t^m(\sigma) := \mathrm{Sur}_{t:T}(\mathcal{A},\gamma_m,v^m,h_t^m,\beta^m,\sigma)$. When the optimal strategy $\mathring{\sigma}^m$ (used by the buyer) is **pure**, we directly have $S_t^m(\hat{\sigma}_1) \le S_t^m(\mathring{\sigma}^m) = S_t^m(\hat{\sigma}_0)$, since the price p_t^m is rejected ($a_t^m = 0$) by our strategic buyer. In the general case, when the buyer's optimal strategy $\mathring{\sigma}^m$ is **mixed**, let α_0 be the probability of a reject ($a_t^m = 0$) and, thus, $1-\alpha_0$ be the probability of an acceptance ($a_t^m = 1$) in this strategy. Since the strategy is optimal, its surplus $S_t^m(\mathring{\sigma}^m) = \alpha_0 S_t^m(\hat{\sigma}_0) + (1-\alpha_0)S_t^m(\hat{\sigma}_1)$ must be no lower than the surplus $S_t^m(\hat{\sigma}_1)$ of the strategy $\hat{\sigma}_1$:

$$\alpha_0 S_t^m(\hat{\sigma}_0) + (1-\alpha_0)S_t^m(\hat{\sigma}_1) \ge S_t^m(\hat{\sigma}_1).$$

Since the price p_t^m is rejected, the probability $\alpha_0 > 0$ and, thus, $\alpha_0 S_t^m(\hat{\sigma}_0) \geq \alpha_0 S_t^m(\hat{\sigma}_1)$. In any way, we obtain:

$$S_t^m(\hat{\sigma}_1) \leq S_t^m(\hat{\sigma}_0). \tag{A.3}$$

Let us bound each side of this inequality:

$$\begin{aligned} S_t^m(\hat{\sigma}_1) &= \mathbb{E}[\sum_{s=t}^{T} \gamma_m^{s-1} \overline{a}_s^m (v^m - \overline{p}_s^m) \mid h_t^m, a_t^m = 1, \hat{\sigma}_1, \beta^m] = \\ &= \gamma_m^{t-1}(v^m - \mathrm{p}(\mathfrak{n})) + \mathbb{E}[\sum_{s=t+1}^{T} \gamma_m^{s-1} \overline{a}_s^m (v^m - \overline{p}_s^m) \mid h_t^m, a_t^m = 1, \hat{\sigma}_1, \beta^m] \geq \\ &\geq \gamma_m^{t-1}(v^m - \mathrm{p}(\mathfrak{n})), \end{aligned} \tag{A.4}$$

where, in the second identity, we used the fact that if the bidder accepts the price $\mathrm{p}(\mathfrak{n})$, then he necessarily gets the good since all other bidders $\mathbb{M}^{-m}$ are eliminated by a barrage price in this round t (**it is the key point of the proof!**). In the last inequality, we used that the expected surplus in rounds $s \geq t+1$ is at least non-negative, because the subalgorithm $\mathcal{A}_1 \in \mathbf{C_R}$ is right consistent and accepting of the offered price $\mathrm{p}(\mathfrak{m})$ in some reached node $\mathfrak{m} \in \mathfrak{T}(\mathcal{A}_1)$ s.t. $\mathrm{p}(\mathfrak{m}) > v^m$ will thus result in reserve prices for him higher than his valuation v^m in all subsequent rounds as well (so, the buyer has no incentive to get a local negative surplus in a round, because it will result in non-positive surplus in all subsequent rounds).

$$\begin{aligned} S_t^m(\hat{\sigma}_0) &= \mathbb{E}[\sum_{s=t}^{T} \gamma_m^{s-1} \overline{a}_s^m (v^m - \overline{p}_s^m) \mid h_t^m, a_t^m = 0, \hat{\sigma}_0, \beta^m] = \\ &= \mathbb{E}[\sum_{s=t_{i+r}^m}^{T} \gamma_m^{s-1} \overline{a}_s^m (v^m - \overline{p}_s^m) \mid h_t^m, a_t^m = 0, \hat{\sigma}_0, \beta^m] \leq \\ &\leq \sum_{s=t+r}^{T} \gamma_m^{s-1} (v^m - \mathrm{p}(\mathfrak{n}) + \delta_{\mathfrak{n}}^l) < \frac{\gamma_m^{t+r-1}}{1-\gamma_m}(v^m - \mathrm{p}(\mathfrak{n}) + \delta_{\mathfrak{n}}^l), \end{aligned} \tag{A.5}$$

where i is the current period of the div-algorithm $\mathrm{div}_M(\langle \mathcal{A}_1 \rangle, \mathrm{sr})$, i.e., the round $t = t_i^m \in \mathcal{T}_i$ is such that the buyer m is the non-eliminated participant in this round (see Sec.3). In the second identity, we used the fact that if the bidder rejects the price p_t^m, then the future rounds $\{t_{i+j}^m\}_{j=1}^{r-1}$ (in which the bidder will be non-eliminated) will be reinforced penalization rounds (and the strategic bidder will reject prices in all of them as well). In the first inequality, we just upper bounded surplus by assuming that only this bidder left among the suspected bidders $\mathbb{S}_j, j > i$, and he receives the lowest possible reserve price from the left subtree $\mathfrak{L}(\mathfrak{n})$ of the node $\mathfrak{n}$. The latter inequality is just a simple arithmetic upper bound for the sum of discounts $\sum_{s=t+r}^{T} \gamma_m^{s-1}$.

We unite these bounds on $S_t^m(\hat{\sigma}_0)$ and $S_t^m(\hat{\sigma}_1)$ (i.e., Eq. (A.3), (A.4), and (A.5)), divide by γ_m^{t-1}, and get

$$(v^m - \mathrm{p}(\mathfrak{n})) \left(1 - \frac{\gamma_m^r}{1-\gamma_m}\right) < \frac{\gamma_m^r}{1-\gamma_m} \delta_{\mathfrak{n}}^l, \tag{A.6}$$

that implies the inequality claimed by the proposition, since $r > \log_{\gamma_m}(1-\gamma_m)$. □

A.2 Missed proofs from Section 4

A.2.1 Proof of Lemma 2

Proof. The idea of the proof is close to the ones of lemmas in [2, 3, 4]. The game has been played and $\mathring{\mathbf{b}}_{1:T} = \mathring{\mathbf{b}}_{1:T}(T, \mathrm{div}_M(\langle \mathcal{A}_1 \rangle, \mathrm{sr}), \mathbf{v}, \boldsymbol{\gamma}, \boldsymbol{\beta})$ are the resulted optimal bids of the strategic buyers $\mathbb{M}$. So, let $L^m := l^m_{I^m}$ be the number of phases conducted by the algorithm during the rounds $\mathcal{I}^m = \{t^m_i\}_{i=1}^{I^m}$ against the strategic buyer m. Then we decompose the total individual regret over these rounds into the sum of the phases' regrets: $\mathrm{Reg}^m(\mathcal{I}^m, \langle \mathcal{A}_1 \rangle, v^m, \mathring{b}^m_{1:T}) = \sum_{l=0}^{L^m} R^m_l$. For the regret R_l at each phase except the last one, the following identity holds:

$$R^m_l = \sum_{k=1}^{K^m_l} (v^m - p^m_{l,k}) + rv^m + g(l)(v^m - p^m_{l,K^m_l}), \quad l = 0, \ldots, L^m - 1, \tag{A.7}$$

where the first, second, and third terms correspond to the exploration rounds with acceptance, the reject-penalization rounds, and the exploitation rounds[1], respectively. Since the basis of the subalgorithm PRRFES $\mathcal{A}_1 \in \mathbf{C_R}$ is right-consistent [2], as discussed in the proof of Proposition 1 (see Appendix A.1.2), the optimal strategy of the bidder m is non-losing [2]: the buyer has no incentive to get a local negative surplus in a round, because it will result in non-positive surplus in all subsequent rounds.

Hence, since the price p^m_{l,K^m_l} is 0 or has been accepted, we have $p^m_{l,K^m_l} \leq v^m$. Second, since the price p^m_{l,K^m_l+1} is rejected, we have $v^m - p^m_{l,K^m_l+1} < (p^m_{l,K^m_l+1} - p^m_{l,K^m_l}) = \epsilon_l$ (by Proposition 1 since $\zeta_{r,\gamma_m} \leq 1$ for $r \geq r_{\gamma_0}$ and $\gamma_m \leq \gamma_0$). Hence, the valuation $v^m \in \big[p^m_{l,K^m_l}, p^m_{l,K^m_l} + 2\epsilon_l\big)$ and all accepted prices $p^m_{l+1,k}$, $\forall k \leq K^m_{l+1}$, from the next phase $l+1$ satisfy:

$$p^m_{l+1,k} \in [q^m_{l+1}, v^m) \subseteq \big[p^m_{l,K^m_l}, p^m_{l,K^m_l} + 2\epsilon_l\big) \quad \forall k \leq K^m_{l+1},$$

because any accepted price has to be lower than the valuation v^m for the strategic buyer (whose optimal strategy is locally non-losing one, as we stated above). This infers $K^m_{l+1} < 2\epsilon_l/\epsilon_{l+1} = 2N_{l+1}$, where $N_l := \epsilon_{l-1}/\epsilon_l = \epsilon^{-1}_{l-1} = 2^{2^{l-1}}$. Therefore, for the phases $l = 1, \ldots, L^m$, we have:

$$v^m - p^m_{l,K^m_l} < 2\epsilon_l; \qquad v^m - p^m_{l,k} < \epsilon_l\big(2N_l - k\big)\ \forall k \in \mathbb{Z}_{2N_l};$$

and

$$\sum_{k=1}^{K^m_l} (v^m - p^m_{l,k}) < \epsilon_l \sum_{k=1}^{2N_l - 1} \big(2N_l - k\big) = \epsilon_l \frac{2N_l - 1}{2}\big(2 \cdot 2N_l - 2N_l\big) \leq 2N_l \cdot N_l \epsilon_l = 2N_l \cdot \epsilon_{l-1} = 2,$$

where we used the definitions of N_l and ϵ_l. For the zeroth phase $l = 0$, one has trivial bound $\sum_{k=1}^{K^m_0} (v - p^m_{0,k}) \leq 1/2$. Hence, by definition of the exploitation rate $g(l)$, we have $g(l) = \epsilon^{-1}_l$ and, thus,

$$R^m_l \leq 2 + rv^m + g(l) \cdot 2\epsilon_l \leq rv^m + 4, \quad l = 0, \ldots, L - 1. \tag{A.8}$$

Moreover, this inequality holds for the L^m-th phase, since it differs from the other ones only in possible absence of some rounds (reject-penalization or exploitation ones). Namely, for the L^m-th phase, we have:

$$R^m_L = \sum_{k=1}^{K^m_L} (v^m - p^m_{L^m,k}) + r_{L^m} v^m + g_{L^m}(L^m)(v^m - p^m_{L^m,K^m_{L^m}}), \tag{A.9}$$

[1] Note that the prices at the exploitation rounds p^m_{l,K^m_l} are equal to either 0 or an earlier accepted price, and are thus accepted by the strategic buyer (since the buyer's decisions at these rounds do not affect further pricing of the algorithm divPRRFES).

where r_{L^m} is the actual number of reject-penalization rounds and $g_{L^m}(L^m)$ is the actual number of exploitation ones in the last phase. Since $r_{L^m} \leq r$ and $g_{L^m}(L^m) \leq g(L^m)$, the right-hand side of Eq. (A.9) is upper-bounded by the right-hand side of Eq. (A.7) with $l = L^m$, which is in turn upper-bounded by the right-hand side of Eq. (A.8). Finally, one has

$$\mathrm{Reg}^m(\mathcal{I}^m, \mathrm{div}_M(\langle\mathcal{A}_1\rangle, \mathrm{sr}), v^m, \mathring{b}^m_{1:T}) = \sum_{l=0}^{L^m} R^m_l \leq (rv^m + 4)\,(L^m + 1).$$

Thus, one needs only to estimate the number of phases L^m by the subhorizon I^m. So, for $2 \leq I^m \leq 2 + r + g(0)$, we have $L^m = 0$ or 1 and thus $L^m + 1 \leq 2 \leq \log_2 \log_2 I^m + 2$. For $I^m \geq 2 + r + g(0)$, we have $I^m = \sum_{l=0}^{L^m-1}(K^m_l + r + g(l)) + K^m_{L^m} + r_{L^m} + g_{L^m}(L^m) \geq g(L^m - 1)$ with $L^m > 0$. Hence, $g(L^m - 1) = 2^{2^{L^m-1}} \leq I^m$, which is equivalent to $L^m \leq \log_2 \log_2 I^m + 1$. Summarizing, we get the claimed upper bound of the lemma. □

A.2.2 Proof of Lemma 3

Proof. Let $\overline{m} \in \overline{\mathbb{M}}$ be one of the bidders $\overline{\mathbb{M}} = \{m \in \mathbb{M} \mid v^m = \overline{v}\}$ that have the maximal valuation $\overline{v}$. Then, the stopping rule $\mathrm{sr}_{\mathcal{A}_1}$ (which is based on the rule $\rho(m, \mathbf{l}, \mathbf{q}) := \exists \hat{m} \in \mathbb{M}^{-m} : q^m + 2\epsilon_{l^m-1} < q^{\hat{m}}\ \forall \mathbf{l} \in \mathbb{Z}^M_+,\ \forall \mathbf{q} \in \mathbb{R}^M_+$) is executed no later than the period i' where the upper bound $q^m_{l^m_{i'}} + 2\epsilon_{l^m_{i'}-1}$ of the bidder m's valuation becomes lower than the lower bound $q^{\overline{m}}_{l^{\overline{m}}_{i'}}$ of the bidder $\overline{m}$'s valuation[2].

Moreover, since $v^m \in [q^m_{l^m_j}, q^m_{l^m_j} + 2\epsilon_{l^m_j-1}]$ and $v^{\overline{m}} \in [q^{\overline{m}}_{l^{\overline{m}}_j}, q^{\overline{m}}_{l^{\overline{m}}_j} + 2\epsilon_{l^{\overline{m}}_j-1}]$ for any period j, the stopping rule is executed no later than the period i where both the phase iteration parameter $\epsilon_{l^m_i}$ of the bidder m and the phase iteration parameter $\epsilon_{l^{\overline{m}}_i}$ of the bidder $\overline{m}$ become smaller than one quarter of the difference between the valuations of these bidders, i.e., $\epsilon_{l^m_i}$ and $\epsilon_{l^{\overline{m}}_i} < \frac{\overline{v}-v^m}{4}$ (because, in this case, the segments $[q^m_{l^m_i}, q^m_{l^m_i} + 2\epsilon_{l^m_i-1}]$ and $[q^{\overline{m}}_{l^{\overline{m}}_i}, q^{\overline{m}}_{l^{\overline{m}}_i} + 2\epsilon_{l^{\overline{m}}_i-1}]$ do not intersect at all, what implies $q^m_{l^m_i} + 2\epsilon_{l^m_i-1} < q^{\overline{m}}_{l^{\overline{m}}_i}$).

Therefore, in the periods $i \leq I^m$, it is not possible to have simultaneously $\epsilon_{l^m_i} < \frac{\overline{v}-v^m}{4}$ and $\epsilon_{l^{\overline{m}}_i} < \frac{\overline{v}-v^m}{4}$. So, in the period $i = I^m$, either $\epsilon_{l^m_{I^m}} \geq \frac{\overline{v}-v^m}{4}$, or (not exclusively) $\epsilon_{l^{\overline{m}}_{I^m}} \geq \frac{\overline{v}-v^m}{4}$ holds. In particular, from the definition of the phase iteration parameter $\epsilon_l = 2^{-2^l}$, we have: if $\epsilon_l \geq \delta$ for some $l \in \mathbb{Z}_+$ and $\delta \in (0, 1/2)$, then

$$\epsilon_l = 2^{-2^l} \geq \delta \quad \Leftrightarrow \quad -2^l \geq \log_2 \delta \quad \Leftrightarrow \quad 2^l \leq \log_2 \frac{1}{\delta} \quad \Leftrightarrow \quad l \leq \log_2 \log_2 \frac{1}{\delta}.$$

Hence, in the period I^m, the following holds:

$$l^m_{I^m} \leq \log_2 \log_2 \frac{4}{\overline{v} - v^m} \quad \text{or (not exclusively)} \quad l^{\overline{m}}_{I^m} \leq \log_2 \log_2 \frac{4}{\overline{v} - v^m},$$

and, thus,

$$\min\{l^m_{I^m}, l^{\overline{m}}_{I^m}\} \leq \log_2 \log_2 \frac{4}{\overline{v} - v^m}. \tag{A.10}$$

Finally, we bound I^m. Let, $L^{m';m} := l^{m'}_{I^m}$ be the phase of a buyer $m' \in \{m, \overline{m}\}$ in the period I^m. As in the proof of Lemma 2 (see Appendix A.2.1) we decompose I^m into the numbers of

[2] Note that it is correct to consider l^m_i in any period i even though the buyer m is not suspected in this period, i.e., $m \notin \mathbb{S}_i$. This is because the algorithm stops change the tracking node $\mathfrak{n}^m_i$ in the subalgorithm tree $\mathfrak{T}(\langle\mathcal{A}_1\rangle)$ after the period I^m, but l^m_i just remains the same in all subsequent periods, i.e., we formally set $l^m_i = l^m_{I^m}$ for all $i > I^m$.

exploration, reject-penalization, and exploitation rounds in each phase $l = 0, \dots, L^{m';m}$ passed by the buyer m'. Namely,

$$I^m = \sum_{l=0}^{L^{m';m}-1} (K_l^{m'} + r + g(l)) + K_{L^{m';m}}^{m'} + r_{L^{m';m}}^{m'} + g_{L^{m';m}}^{m'}, \tag{A.11}$$

where $r_l^{m'}$ and $g_l^{m'}$ are the numbers of penalization rounds and exploitation rounds, resp., passed by the buyer m' in the last phase $l = L^{m';m}$ before reaching the period I^m. Let us trivially bound $r_{L^{m';m}}^{m'} \le r$ and $g_{L^{m';m}}^{m'} \le g(L^{m';m})$. We also know that, for any $l \in \mathbb{Z}_+$, $K_l^{m'} \le 2 \cdot 2^{2^{l-1}}$ (see the proof of Lemma 2 in Appendix A.2.1). Therefore, Eq. A.11 implies

$$I^m \le \sum_{l=0}^{L^{m';m}} (2 \cdot 2^{2^{l-1}} + r + 2^{2^l}) \le \sum_{l=0}^{L^{m';m}} (3 \cdot 2^{2^l} + r) \le (L^{m';m} + 1)r + 2 \cdot 3 \cdot 2^{2^{L^{m';m}}}, \tag{A.12}$$

Taking $m' = m$ and $m' = \overline{m}$, we get the following from Eq. (A.12):

$$I^m \le (\min\{l_{I^m}^m, l_{I^m}^{\overline{m}}\} + 1)r + 6 \cdot 2^{2^{\min\{l_{I^m}^m, l_{I^m}^{\overline{m}}\}}} \le r(\log_2 \log_2 \frac{4}{\overline{v} - v^m} + 1) + 6 \cdot \frac{4}{\overline{v} - v^m}, \tag{A.13}$$

where we used the definition of $L^{m';m} := l_{I^m}^{m'}$ and the upper bound for the phases $l_{I^m}^m$ and $l_{I^m}^{\overline{m}}$ in Eq. (A.10). So, Eq. (A.13) implies the claim of the lemma.

□

A.2.3 Proof of Theorem 1

Proof. From Lemma 1, we have:

$$\mathrm{SReg}(T, \mathcal{A}, \mathbf{v}, \boldsymbol{\gamma}, \boldsymbol{\beta}) = \sum_{m=1}^{M} \mathrm{Reg}^m(\mathcal{I}^m, \mathcal{A}^m, v^m, \mathring{b}_{1:T}^m) + \sum_{m=1}^{M} I^m(\overline{v} - v^m). \tag{A.14}$$

From Lemma 2, if $I^m \ge 2$, one can upper bound the first term in right-hand side of Eq. (A.14) since $\mathcal{A}^m = \langle \mathcal{A}_1 \rangle$:

$$\mathrm{Reg}^m(\mathcal{I}^m, \mathcal{A}^m, v^m, \mathring{b}_{1:T}^m) \le (rv^m + 4)(\log_2 \log_2 I^m + 2) \le (r\overline{v} + 4)(\log_2 \log_2 T + 2), \tag{A.15}$$

where we bounded the subhorizon I^m of each bidder $m \in \mathbb{M}$ by the time horizon T (i.e., $I^m \le T$) and the valuation v^m of each bidder $m \in \mathbb{M}$ by the maximal valuation (i.e., $v^m \le \overline{v}$). Note that the latter bound of Eq. (A.15) holds for $\mathrm{Reg}^m(\mathcal{I}^m, \mathcal{A}^m, v^m, \mathring{b}_{1:T}^m)$ in the case of $I^m = 1$ as well (this case has not been provided by Lemma 2).

From Lemma 3, one can upper bound the second term in right-hand side of Eq. (A.14):

$$\sum_{m=1}^{M} I^m(\overline{v} - v^m) \le \sum_{\{m \in \mathbb{M} | v^m \ne \overline{v}\}} \frac{24 + 5r}{\overline{v} - v^m}(\overline{v} - v^m) \le (24 + 5r)(M - 1), \tag{A.16}$$

where we used that at least one bidder $\overline{m} \in \mathbb{M}$ has $v^{\overline{m}} = \overline{v}$ and, hence, $|\{m \in \mathbb{M} \mid v^m \ne \overline{v}\}| \le M - 1$.

Thus, plugging Eq. (A.15) and Eq. (A.16) into Eq. (A.14), we obtain the claimed bound for the strategic regret of divPRRFES.

□

B The pseudo-codes

B.1 The pseudo-code of div-transformation

Algorithm B.1 Pseudo-code of a div-transformation $\mathrm{div}_M(\mathcal{A}_1, \mathrm{sr})$ of a RPPA algorithm $\mathcal{A}_1 \in \mathbf{A}^{\mathrm{RPPA}}$.

Input: $M \in \mathbb{N}$, $\mathcal{A}_1 \in \mathbf{A}^{\mathrm{RPPA}}$, $\mathrm{sr} : \mathbb{M} \times \mathfrak{T}(\mathcal{A}_1)^M \to \texttt{bool}$
Initialize: $t := 1$, $\mathbb{S} := \mathbb{M}$, $\mathfrak{n}[\,] := \{\mathfrak{e}(\mathfrak{T}(\mathcal{A}_1))\}_{m=1}^{M}$
while $t \le T$ **do**
  for all $m \in \mathbb{S}$ **do**
    Set the price $\mathrm{p}(\mathfrak{n}[m])$ as reserve to the buyer m
    Set the price p^{bar} as reserve to the buyers from $\mathbb{M}^{-m}$
    $\mathbf{b}[\,] \leftarrow$ get bids from the buyers $\mathbb{M}$
    if $\mathbf{b}[m] \ge \mathrm{p}(\mathfrak{n}[m])$ **then**
      Allocate t-th good to the buyer m for the price $\mathrm{p}(\mathfrak{n}[m])$
      $\mathfrak{n}[m] := \mathfrak{r}(\mathfrak{n}[m])$
    else
      $\mathfrak{n}[m] := \mathfrak{l}(\mathfrak{n}[m])$
    end if
    $t := t + 1$
    if $t > T$ **then**
      break
    end if
  end for
  $\mathbb{S}^{\mathrm{old}} := \mathbb{S}$
  for all $m \in \mathbb{S}^{\mathrm{old}}$ **do**
    if $\mathrm{sr}(m, \mathfrak{n}[\,])$ **then**
      $\mathbb{S} := \mathbb{S} \setminus \{m\}$
    end if
  end for
end while

B.2 The pseudo-code of divPRRFES

Algorithm B.2 Pseudo-code of the algorithm divPRRFES.

```
Input: M ∈ ℕ, r ∈ ℕ, and g : ℤ+ → ℤ+
Initialize: t := 1, 𝕊 := 𝕄, q[ ] := {0}_{m=1}^M, l[ ] := {0}_{m=1}^M, x[ ] := {0}_{m=1}^M, state[ ] := {"explore"}_{m=1}^M
while t ≤ T do
  for all m ∈ 𝕊 do
    if state[m] = "penalize" then
      p := 1 // a reinforced penalization round for the buyer m
      x[m] := x[m] − 1
    end if
    if state[m] = "explore" then
      p := q[m] + 2^{−2^{l[m]}} // an exploration round for the buyer m
    else
      p := q[m] // an exploitation round for the buyer m
      x[m] := x[m] − 1
    end if
    Set the price p as reserve to the buyer m
    Set the price p^bar as reserve to the buyers from 𝕄^{−m}
    b[ ] ← get bids from the buyers 𝕄
    if b[m] ≥ p then
      Allocate t-th good to the buyer m for the price p
      q[m] := p
      if state[m] = "penalize" then
        x[m] := −1 // a reinforced penalization price is accepted; set 1 to the buyer m all his rounds
      end if
    else
      if state[m] = "explore" then
        state[m] := "penalize"
        x[m] := r // an exploration price is rejected; move the buyer m to penalization
      end if
    end if
    if state[m] = "penalize" and x[m] = 0 then
      state[m] := "exploit"
      x[m] := g(l[m]) // penalization rounds are ended; move the buyer m to exploitation
    end if
    if state[m] = "exploit" and x[m] = 0 then
      state[m] := "explore"
      l[m] := l[m] + 1 // exploitation rounds are ended; move the buyer m to the next phase
    end if
    t := t + 1
    if  t > T then
      break
    end if
  end for
  𝕊^old := 𝕊
  q_max := max_{m∈𝕄}(q[m])
  for all m ∈ 𝕊^old do
    if  q[m] + 2 ∗ 2^{−2^{l[m]−1}} < q_max then
      𝕊 := 𝕊 \ {m} // remove the buyer m from suspected ones if the stopping rule is satisfied
    end if
  end for
end while
```

C Summary on used notations

Note that we use several mnemonic notations:

- upper index for a value of a particular buyer (e.g., v^m, a_t^m, p_t^m, etc.);
- boldface for a vector of values for all bidders (e.g., $\mathbf{v}$, $\mathbf{a}_t$, $\mathbf{p}_t$, etc.);
- bar (overline) for terms associated with best value / winning (e.g., the winner $\overline{m}_t$, the highest valuation $\overline{v}$, etc.); etc.

The full list of used notations is summarized below in the following tables.

C.1 General notations

Table C.1: General notations: part I.

Notation	Expression	Description
$\mathbb{E}[\cdot]$		expectation
$\mathbb{I}_B$		the indicator: $\mathbb{I}_B = 1$, when B holds, and 0, otherwise.
T		the [time] horizon, the number of rounds in the repeated game
t		a round in the repeated game, $t \in \{1, \ldots, T\}$
v^m		the valuation of a buyer m
$\overline{v}$	$= \max_{m \in \mathbb{M}} v^m$	the highest valuation among the buyers
$\overline{\overline{v}}$	$= \max_{m \in \mathbb{M} \setminus \overline{\mathbb{M}}} v^m$	the maximal valuation among non-highest valuations ot the buyers (if exists)
$\overline{m}$		a buyer that has the highest valuation $\overline{v}$
$\overline{m}_t$	$= \mathrm{argmax}_{m \in \mathbb{M}_t} b_t^m$	the winning bidder in a round t for a given play of the game (if exists)
b_t^m		the bid of a buyer m in a round t for a given play of the game
p_t^m		the reserve price set to a buyer m in a round t for a given play of the game
a_t^m	$= \mathbb{I}_{b_t^m \geq p_t^m}$	indicator of bidding higher than the reserve price by a buyer m in a round t for a given play of the game
$\overline{a}_t^m$	$= \mathbb{I}_{\{\mathbb{M}_t \neq \varnothing \& m = \overline{m}_t\}}$	the allocation outcome of a round t for a bidder m for a given play of the game
$\overline{a}_t$	$= \mathbb{I}_{\{\mathbb{M}_t \neq \varnothing\}}$	the allocation outcome of a round t over all bidders for a given play of the game
$\overline{p}_t^m$	$= \overline{a}_t^m \overline{p}_t$	the payment outcome of a round t for a bidder m for a given play of the game
$\overline{p}_t$	$= \max\{p_t^{\overline{m}_t}, \max_{m \in \mathbb{M}_t^{-\overline{m}_t}} b_t^m\}$	the payment outcome of a round t over all bidders for a given play of the game
$\mathbf{x}$	$= \{x^m\}_{m=1}^M$	the vector of buyer values of some notion x (e.g., valuations $\mathbf{v}$, bids $\mathbf{b}_t$, reserve prices $\mathbf{p}_t$, payments $\overline{\mathbf{p}}_t$, allocations $\overline{\mathbf{a}}_t$ and $\mathbf{a}_t$ etc)
$x_{t_1:t_2}$	$= \{x_t\}_{t=t_1}^{t_2}$	the subseries of some time series $\{x_t\}_{t=1}^T$ (e.g., bids $\mathbf{b}_{1:T}$, reserve prices $\mathbf{p}_{1:T}$, payments $\overline{\mathbf{p}}_{1:T}$, allocations $\overline{\mathbf{a}}_{1:T}$ and $\mathbf{a}_{1:T}$ etc)
$\mathbf{A}_M$		the set of pricing algorithms of the seller against M buyers
$\mathbf{A}^{\mathrm{RPPA}}$	$\subset \mathbf{A}_1$	the subclass of 1-buyer pricing algorithms for repeating posted-price auctions
$\mathcal{A}$		a pricing algorithm (generally, from the set $\mathbf{A}_M$)
M		the number of buyers in the repeated game
$\mathbb{M}$	$= \{1, \ldots, M\}$	the set of buyers (bidders)
$\overline{\mathbb{M}}$	$= \{m \in \mathbb{M} \mid v^m = \overline{v}\}$	the set of buyers whose valuation is the highest one $\overline{v}$
$\mathbb{M}^{-m}$	$= \mathbb{M} \setminus \{m\}$	the set of buyers (bidders) without the buyer m
$\mathbb{M}_t$	$= \{m \in \mathbb{M} \mid b_t^m \geq p_t^m\}$	the set of actual buyers in a round t (they bid higher than reserve prices)

Table C.2: General notations: part II.

Notation	Expression	Description
$\mathrm{Reg}(\ldots)$		regret of a pricing algorithm
$\mathrm{SReg}(\ldots)$		strategic regret of a pricing algorithm
$\mathrm{Sur}(\ldots)$		expected surplus of a buyer (bidder)
γ_m		the discount rate of a buyer $m \in \mathbb{M}$
$\boldsymbol{\gamma}$	$= \{\gamma_m\}_{m=1}^M$	the vector of the discount rates of the buyers
h		a buyer history
h_t^m	$= (b_{1:t-1}^m, p_{1:t}^m, \overline{a}_{1:t-1}^m, \overline{p}_{1:t-1}^m)$	the history available to a buyer m in a round t for a given play of the game
σ	$\in \mathfrak{S}_T$	a buyer strategy
β^m	$\in \mathfrak{S}_T^{M-1}$	the beliefs of a buyer m on the strategies of the other bidders
$\boldsymbol{\beta}$	$= \{\beta^m\}_{m=1}^M$	the beliefs of all buyers
$\mathbb{H}_t$		the set of all possible histories in a round t
$\mathbb{H}_{t_1:t_2}$	$= \sqcup_{t=t_1}^{t_2} \mathbb{H}_t$	the disjoint union of the sets of histories in rounds $t_1, \ldots, t_2$
$\mathfrak{S}_T$		the set of all possible buyer strategies
$\mathring{\sigma}^m$		an optimal strategy of a buyer m in a round t
$\mathring{b}_t^m$		the optimal bid of a buyer m in a round t for a given play of the game
$\mathring{\mathbf{b}}_t$	$= \{\mathring{b}_t^m\}_{m=1}^M$	the optimal bids of all buyers in a round t for a given play of the game
$\mathring{\mathbf{b}}_{1:T}$		the optimal bids of all buyers in all rounds for a given play of the game

Table C.3: General notations: part III (related to RPPA algorithms).

Notation	Expression	Description
$\mathfrak{T}(\mathcal{A}_1)$		the complete binary tree associated with a RPPA algorithm $\mathcal{A}_1$
$\mathfrak{n}$ or $\mathfrak{m}$		a node in the complete binary tree $\mathfrak{T}(\mathcal{A}_1)$ of a RPPA algorithm $\mathcal{A}_1$
$\mathfrak{r}(\mathfrak{n})$		the right child of a node $\mathfrak{n}$
$\mathfrak{l}(\mathfrak{n})$		the left child of a node $\mathfrak{n}$
$\mathfrak{R}(\mathfrak{n})$		the right subtree of a node $\mathfrak{n}$ (its root is $\mathfrak{r}(\mathfrak{n})$)
$\mathfrak{L}(\mathfrak{n})$		the left subtree of a node $\mathfrak{n}$ (its root is $\mathfrak{l}(\mathfrak{n})$)
$\mathfrak{e}(\mathfrak{T})$		the root of a tree $\mathfrak{T}$
$\mathrm{p}(\mathfrak{n})$		the price in a node $\mathfrak{n}$ (that is offered to a buyer when an algorithm reaches this node)
$\mathfrak{T}_1 \cong \mathfrak{T}_2$		the trees $\mathfrak{T}_1$ and $\mathfrak{T}_2$ are price-equivalent
$\delta_{\mathfrak{n}}^l$	$= \mathrm{p}(\mathfrak{n}) - \inf_{\mathfrak{m} \in \mathfrak{L}(\mathfrak{n})} \mathrm{p}(\mathfrak{m})$	the left increment of a node $\mathfrak{n}$

C.2 Notations related to dividing algorithms

Table C.4: Notations related to dividing algorithms.

Notation	Expression	Description
i		a period of a dividing algorithm (**do not confuse with** (1) a round of the game and (2) a phase of PRRFES algorithm!)
t_i^m		the round in a period i in which the bidder m is not eliminated by a barrage price (i.e., m is non-eliminated participant) of a dividing algorithm for a given play of the game
$p^{m,\text{bar}}$ or p^{bar}		a barrage reserve price
$\mathbb{S}_i$		the set of bidders suspected by a dividing algorithm in a period i for a given play of the game
$\mathcal{T}_i$		the rounds of a period i for a given play of the game
$\mathcal{I}^m$	$= \{t_i^m\}_{i=1}^{I^m}$	the rounds in which the bidder m is not eliminated by a barrage price (i.e., m is non-eliminated participant) of dividing algorithm for a given play of the game
I^m	$= \|\mathcal{I}^m\|$	the subhorizon of a buyer m (the number of periods in which he is suspected, i.e., $m \in \mathbb{S}_i$) for a given play of the game
$\mathcal{A}^m$		the subalgorithm of a dividing algorithm that acts against a buyer m
$\text{Reg}^m(\dots)$		Regret of the subalgorithm of a dividing algorithm that acts against a buyer m
$\text{div}_M(\dots)$		a div-transformation of 1-buyer pricing algorithm to the case of M buyers
$\text{SReg}^{\text{ind}}(\dots)$		individual strategic regret of a dividing algorithm
$\text{SReg}^{\text{dev}}(\dots)$		deviation strategic regret of a dividing algorithm
sr		a stopping rule used in a div_M-transformation of 1-buyer pricing algorithm
$\langle\mathcal{A}\rangle$		a transformation of a RPPA algorithm $\mathcal{A}$ s.t. all penalization sequences of nodes are replaced by reinforced penalization ones
$\mathfrak{n}_i^m$		the tracking node of a buyer m by div_M-transformed RPPA algorithm in a period i for a given play of the game

C.3 Notations related to divPRRFES

Table C.5: Notations related to divPRRFES.

Notation	Expression	Description
r		the number of penalization rounds (a parameter of PRRFES)
$g(l)$		the exploitation rate (a parameter of PRRFES)
l		a phase of PRRFES
ε_l	$= 2^{-2^l}$	the iteration parameter of a phase l
q_l^m		the last accepted price by a buyer m before a phase l for a given play of the game
$p_{l,k}^m$		the k-th exploration price of a buyer m in a phase l for a given play of the game
K_l^m		the last accepted exploration price of a buyer m in a phase l for a given play of the game
l_i^m		the current phase of a buyer m in a period i for a given play of the game
$l(\mathfrak{n})$		the phase of a node $\mathfrak{n}$ from the tree of the algorithm PRRFES
$q(\mathfrak{n})$		the last accepted price before the current phase of a node $\mathfrak{n}$ from the tree of the algorithm PRRFES

References

[1] K. Amin, A. Rostamizadeh, and U. Syed. Learning prices for repeated auctions with strategic buyers. In *NIPS'2013*, pages 1169–1177, 2013.

[2] A. Drutsa. Horizon-independent optimal pricing in repeated auctions with truthful and strategic buyers. In *WWW'2017*, pages 33–42, 2017.

[3] A. Drutsa. On consistency of optimal pricing algorithms in repeated posted-price auctions with strategic buyer. *CoRR*, abs/1707.05101, 2017.

[4] A. Drutsa. Weakly consistent optimal pricing algorithms in repeated posted-price auctions with strategic buyer. In *ICML'2018*, pages 1318–1327, 2018.

[5] A. Drutsa. Optimal non-parametric learning in repeated contextual auctions with strategic buyer. In *ICML'2020*, 2020.

[6] M. Mohri and A. Munoz. Optimal regret minimization in posted-price auctions with strategic buyers. In *NIPS'2014*, pages 1871–1879, 2014.

[7] A. Vanunts and A. Drutsa. Optimal pricing in repeated posted-price auctions with different patience of the seller and the buyer. In *Advances in Neural Information Processing Systems*, pages 939–951, 2019.

[8] A. Zhiyanov and A. Drutsa. Bisection-based pricing for repeated contextual auctions against strategic buyer. In *ICML'2020*, 2020.