Weakly Consistent Optimal Pricing Algorithms in Repeated Posted-Price Auctions with Strategic Buyer: SUPPLEMENTARY MATERIALS*

Alexey Drutsa†

Contents

A Missed proofs

A.1 Missed proofs from Section 4

A.1.1 Proof of Proposition 1

Proof. For each node $\mathfrak{m} \in \mathfrak{T}(\mathcal{A})$, let $S(\mathfrak{m})$ be the surplus obtained by the buyer when playing an optimal strategy against $\mathcal{A}$ after reaching the node $\mathfrak{m}$. Since the price $p^{\mathfrak{n}}$ is rejected then the following inequality holds (see [5, Lemma 1])

$$\gamma^{t^{\mathfrak{n}}-1}(v - p^{\mathfrak{n}}) + S(\mathfrak{r}(\mathfrak{n})) < S(\mathfrak{l}^r(\mathfrak{n})). \tag{A.1}$$

The left subtree's surplus $S(\mathfrak{l}^r(\mathfrak{n}))$ can be upper bounded as follows (using $p^{\mathfrak{n}} \leq p^{\mathfrak{m}} \forall \mathfrak{m} \in \mathfrak{L}(\mathfrak{l}^{r-1}(\mathfrak{n}))$):

$$S(\mathfrak{l}^r(\mathfrak{n})) \leq \sum_{t=t^{\mathfrak{n}}+r}^{T} \gamma^{t-1}(v - p^{\mathfrak{n}}) < \frac{\gamma^{t^{\mathfrak{n}}+r-1}}{1-\gamma}(v - p^{\mathfrak{n}});$$

*This is the supplementary materials for the work published in Proceedings of the 35 th International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018

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

while, in contrast to the proof of [2, Prop.2], we lower bound the right subtree's surplus $S(\mathfrak{r}(\mathfrak{n}))$ by $\gamma^{t^{\mathfrak{n}}}(v - p^{\mathfrak{r}(\mathfrak{n})})$[1], because, after accepting $p^{\mathfrak{n}}$ at the round $t^{\mathfrak{n}}$, the buyer is able to earn at least this amount at the round $t^{\mathfrak{n}}+1$. We plug these bounds in Eq. (A.1), divide by $\gamma^{t^{\mathfrak{n}}-1}$, and obtain

$$(v - p^{\mathfrak{r}(\mathfrak{n})} + p^{\mathfrak{r}(\mathfrak{n})} - p^{\mathfrak{n}}) + \gamma(v - p^{\mathfrak{r}(\mathfrak{n})}) < \frac{\gamma^r}{1-\gamma}(v - p^{\mathfrak{r}(\mathfrak{n})} + p^{\mathfrak{r}(\mathfrak{n})} - p^{\mathfrak{n}}) \Leftrightarrow$$

$$\Leftrightarrow (v - p^{\mathfrak{r}(\mathfrak{n})})\left(1 + \gamma - \frac{\gamma^r}{1-\gamma}\right) < \left(\frac{\gamma^r}{1-\gamma} - 1\right)(p^{\mathfrak{r}(\mathfrak{n})} - p^{\mathfrak{n}}),$$

that implies Eq. (2), since $r > \log_\gamma(1-\gamma^2)$ implies $1 - \gamma^2 - \gamma^r > 0$. □

A.1.2 Proof of Proposition 2

Proof. As in the proof of Prop. 1, let $S(\mathfrak{m})$ be the surplus obtained by the buyer when playing an optimal strategy against $\mathcal{A}$ after reaching the node $\mathfrak{m}$, for each node $\mathfrak{m} \in \mathfrak{T}(\mathcal{A})$. The condition $v < p \;\forall p \in \wp(\mathfrak{R}(\mathfrak{n}))$ implies that $S(\mathfrak{r}(\mathfrak{n})) = 0$ and the strategic buyer will thus gain exactly $\gamma^{t^{\mathfrak{n}}-1}(v - p^n)$ if he accepts the price p^n at the round $t^{\mathfrak{n}}$. Let us show that there exists a strategy in $\mathfrak{L}(\mathfrak{n})$ with a larger surplus. Indeed, if the buyer rejects r times the price $p^{\mathfrak{n}}$ and accepts this price G times after that, then he gets the following surplus:

$$\sum_{s=t^{\mathfrak{n}}+r}^{t^{\mathfrak{n}}+r+G-1} \gamma^{s-1}(v - p^n) = \frac{\gamma^{t^n+r-1} - \gamma^{t^n+r-1+G}}{1-\gamma}(v - p^n) = \gamma^{t^n-1}\gamma^r\frac{1-\gamma^G}{1-\gamma}(v - p^n) > \gamma^{t^{\mathfrak{n}}-1}(v - p^n),$$

where the last inequality holds due to the condition on G and

$$\gamma^r(1-\gamma^G)/(1-\gamma) > 1 \Leftrightarrow (1-\gamma^G) > (1-\gamma)\gamma^{-r} \Leftrightarrow \gamma^G < 1 - (1-\gamma)\gamma^{-r}.$$

□

A.1.3 Proof of Lemma 2

Proof. In order to get the claims of this lemma, one just needs to straightforwardly verify few inequalities.

<u>The case $\gamma \in ((\sqrt{5}-1)/2, 1)$.</u>

For Prop. 1, we have:

$$r = \lceil r_{\gamma,\varkappa} \rceil \geq r_{\gamma,\varkappa} = \log_\gamma\left((1-\gamma)\Big(1 + \frac{\varkappa}{1+\varkappa}\gamma\Big)\right) > \log_\gamma\left((1-\gamma)(1+\gamma)\right) = \log_\gamma(1-\gamma^2)$$

since $\varkappa/(1+\varkappa) < 1 \;\forall \varkappa > 0$; and

$$\eta_{r,\gamma} = \frac{\gamma^r + \gamma - 1}{1 - \gamma^2 - \gamma^r} \leq \frac{(1-\gamma)\Big(1 + \frac{\varkappa}{1+\varkappa}\gamma\Big) + \gamma - 1}{1 - \gamma^2 - (1-\gamma)\Big(1 + \frac{\varkappa}{1+\varkappa}\gamma\Big)} = \frac{1 + \frac{\varkappa}{1+\varkappa}\gamma - 1}{1 + \gamma - 1 - \frac{\varkappa}{1+\varkappa}\gamma} = \varkappa.$$

For Prop. 2, we have:

$$r = \lceil r_{\gamma,\varkappa} \rceil < r_{\gamma,\varkappa} + 1 = \log_\gamma(1-\gamma) + \log_\gamma\left(\Big(1 + \frac{\varkappa}{1+\varkappa}\gamma\Big)\gamma\right) < \log_\gamma(1-\gamma)$$

[1] This term may be negative (when $v < p^{\mathfrak{r}(\mathfrak{n})}$), but the lower bound on optimal surplus $S(\mathfrak{r}(\mathfrak{n}))$ holds a fortiori in this case.

since $\left(1+\frac{\varkappa}{1+\varkappa}\gamma\right)\gamma > 1$ due to $\varkappa > (1-\gamma)/(\gamma^2+\gamma-1)$; and, finally,

$$G \geq G_{\gamma,\varkappa} = \log_\gamma\left(1 - \left(1+\frac{\varkappa}{1+\varkappa}\gamma\right)^{-1}\gamma^{-1}\right) > \log_\gamma\left(1-(1-\gamma)\gamma^{-r}\right),$$

where we used $\left(1+\frac{\varkappa}{1+\varkappa}\gamma\right)\gamma < \gamma^r/(1-\gamma)$ since $r < r_{\gamma,\varkappa}+1$.

The case $\gamma \in (1/2, (\sqrt{5}-1)/2)$

For Prop. 1, we have:

$$r = \lceil r_{\gamma,\varkappa} \rceil = 1 > \log_\gamma(1-\gamma^2)$$

since $1-\gamma^2-\gamma > 0$ for this case of γ; and

$$\eta_{r,\gamma} = \frac{\gamma^r+\gamma-1}{1-\gamma^2-\gamma^r} = \frac{\gamma+\gamma-1}{1-\gamma^2-\gamma} = \frac{2\gamma-1}{1-\gamma^2-\gamma} \leq \varkappa.$$

For Prop. 2, we have:

$$r = \lceil r_{\gamma,\varkappa} \rceil = 1 < \log_\gamma(1-\gamma)$$

since $\gamma > 1/2$; and, finally,

$$G \geq G_{\gamma,\varkappa} = \log_\gamma(2\gamma-1) > \log_\gamma(2\gamma-1) - 1 = \log_\gamma\left(1-\frac{1-\gamma}{\gamma}\right) = \log_\gamma\left(1-(1-\gamma)\gamma^{-r}\right).$$

□

A.1.4 Proof of Theorem 1

First of all, let us remind the following notion. A buyer strategy $\mathbf{a}$ is said to be *locally non-losing* w.r.t. v and $\mathcal{A}$ if prices higher than v are never accepted[2] (i.e., $a_t = 1$ implies $p_t \leq v$).

We also introduce some notations: for $l \in \mathbb{N}$,

$$\epsilon_l := \epsilon_{l-1}^2 = 2^{-2^l}, \quad N_l := \epsilon_{l-1}/\epsilon_l = \epsilon_{l-1}^{-1} = 2^{2^{l-1}}. \tag{A.2}$$

Proof. Note that the conditions of this theorem allow us to apply Lemma 2, Prop. 1, and Prop. 2. So, let L be the number of phases conducted by the algorithm during T rounds, then we decompose the total regret over T rounds into the sum of the phases' regrets: $\mathrm{SReg}(T,\mathcal{A},v,\gamma) = \sum_{l=0}^{L} R_l$. For the regret R_l at each phase except the last one, the following identity holds:

$$R_l = \sum_{k=0}^{K_l-1}(v-p_{l,k}) + rv + g(l)(v-p_{l,K_l}), \quad l = 0,\ldots,L-1, \tag{A.3}$$

where the first, second, and third terms correspond to the exploration rounds with acceptance, the reject-penalization rounds, and the exploitation rounds, respectively. First, note that here, in the exploitation rounds, we directly use Proposition 2 (via Lemma 2 since $g(l) \geq G_{\gamma,\varkappa}$) to conclude that $p_{l,K_l} < v$ and the price p_{l,K_l} is thus accepted by the strategic buyer at the exploitation rounds (since the buyer's decisions at these rounds do not affect further pricing of the algorithm prePRRFES and $p_{l,K_l} < v$).

Second, since the price p_{l,K_l} is rejected, we have $v - p_{l,K_l+1} < \varkappa(p_{l,K_l+1} - p_{l,K_l}) = \varkappa\epsilon_l$ (by Proposition 1 via Lemma 2 since $\eta_{r,\gamma} \leq \varkappa$ for $r \geq \lceil r_{\gamma,\varkappa} \rceil$ and any $t \in \mathbb{N}$). Hence, the valuation

[2] Note that the optimal strategy of a strategic buyer may not satisfy this property: it is easy to imagine an algorithm that offers the price 1 at the first round and, if it is accepted, offers the price 0 all remaining rounds.

$v \in \big(p_{l,K_l}, p_{l,K_l} + (1+\varkappa)\epsilon_l\big)$ and all accepted prices $p_{l+1,k}$, $\forall k \le K_{l+1}$, from the next phase $l+1$ satisfy:

$$p_{l+1,k} \in (q_{l+1}, v) \subseteq \big(p_{l,K_l}, p_{l,K_l} + (1+\varkappa)\epsilon_l\big) \quad \forall k \le K_{l+1},$$

because any accepted price has to be lower than the valuation v for the strategic buyer (whose optimal strategy is locally non-losing one for $\mathcal{A} \in \mathbf{C_R}$, see the discussion after [2, Lemma 1]). This infers $K_{l+1} < (1+\varkappa)N_{l+1} \le \lceil (1+\varkappa)N_{l+1} \rceil =: N_{l+1,\varkappa}$ since $N_{l+1} = \epsilon_l/\epsilon_{l+1}$ by Eq. (A.2). Therefore, for the phases $l = 1, \ldots, L$, we have:

$$v - p_{l,K_l} < (1+\varkappa)\epsilon_l; \qquad v - p_{l,k} < \epsilon_l\big((1+\varkappa)N_l - k\big)\ \forall k \in \mathbb{Z}_{N_{l,\varkappa}};$$

and

$$\sum_{k=0}^{K_l-1}(v - p_{l,k}) < \epsilon_l \sum_{k=0}^{N_{l,\varkappa}-2}\big((1+\varkappa)N_l - k\big) = \epsilon_l \frac{N_{l,\varkappa}-1}{2}\big(2(1+\varkappa)N_l - N_{l,\varkappa} + 2\big) \le$$

$$\le \epsilon_l \frac{(1+\varkappa)N_l}{2}\big((1+\varkappa)N_l + 2\big) = \frac{(1+\varkappa)^2}{2} N_l \cdot N_l\epsilon_l + (1+\varkappa)N_l\epsilon_l = \frac{(1+\varkappa)^2}{2} + (1+\varkappa)\epsilon_{l-1},$$

where we used the definitions of N_l and ϵ_l (i.e., $N_l\epsilon_l = \epsilon_{l-1}$ and $N_l = \epsilon_l^{-1}$), given in Eq. (A.2). For the zeroth phase $l = 0$, one has trivial bound $\sum_{k=0}^{K_0-1}(v - p_{0,k}) \le 1$. Hence, by definition of the exploitation rate $g(l)$, we have

$$g(l) \cdot \epsilon_l = \max\{\epsilon_l^{-1} \cdot \epsilon_l, \lceil G_{\gamma,\varkappa} \rceil \cdot \epsilon_l\} \le \max\{1, \lceil G_{\gamma,\varkappa} \rceil / 2\},$$

and, thus,

$$R_l \le \frac{(1+\varkappa)(2+\varkappa)}{2} + rv + g(l)\cdot(1+\varkappa)\epsilon_l \le rv + \frac{(1+\varkappa)}{2}(2 + \max\{2, \lceil G_{\gamma,\varkappa} \rceil\} + \varkappa), \quad l = 0, \ldots, L-1. \tag{A.4}$$

The L-th phase differs from the other ones only in possible absence of some rounds: (reject-penalization or exploitation ones). In this phase, we consider two cases on the actual number of exploitation rounds $g_L(L)$: (a) $g_L(L) \ge \lceil G_{\gamma,\varkappa} \rceil$ and (b) $g_L(L) < \lceil G_{\gamma,\varkappa} \rceil$. In the case (a), we again apply Proposition 2 (via Lemma 2 since $g_L(L) \ge \lceil G_{\gamma,\varkappa} \rceil$) to get that $p_{L,K_L} < v$ and the price p_{L,K_L} is thus accepted by the strategic buyer at the exploitation rounds. In this case, we have thus:

$$R_L = \sum_{k=0}^{K_L-1}(v - p_{L,k}) + rv + g_L(L)(v - p_{L,K_L}). \tag{A.5}$$

The right-hand side of Eq. (A.5) is upper-bounded by the right-hand side of Eq. (A.3) with $l = L$, which is in turn upper-bounded by the right-hand side of Eq. (A.4). In the case (b), we have no guarantee that $p_{L,K_L} < v$ and, hence, p_{L,K_L} may be rejected by the strategic buyer at the exploitation rounds. Hence, we have to estimate the regret in the last phase in the following way:

$$R_L = \sum_{k=0}^{K_L-1}(v - p_{L,k}) + r_L v + g_L(L)v \le \frac{(1+\varkappa)(2+\varkappa)}{2} + (r + \lceil G_{\gamma,\varkappa} \rceil - 1)v, \tag{A.6}$$

where r_L the actual number of reject-penalization rounds, $r_L \le r$.

Finally, using $(\lceil G_{\gamma,\varkappa} \rceil - 1)v - \max\{1, \lceil G_{\gamma,\varkappa} \rceil/2\} \le \lceil G_{\gamma,\varkappa} \rceil/2 - 1$, one has

$$\mathrm{SReg}(T, \mathcal{A}, v, \boldsymbol{\gamma}) = \sum_{l=0}^{L} R_l \le \left(rv + \frac{(1+\varkappa)}{2}(2 + \max\{2, \lceil G_{\gamma,\varkappa} \rceil\} + \varkappa)\right)(L+1) + \frac{\lceil G_{\gamma,\varkappa} \rceil}{2} - 1.$$

Thus, one needs only to estimate the number of phases L by the number of rounds T. So, for $2 \le T \le 2+r+g(0)$, we have $L = 0$ or 1 and thus $L+1 \le 2 \le \log_2 \log_2 T+2$. For $T \ge 2+r+g(0)$, we have $T = \sum_{l=0}^{L-1}(K_l+r+g(l))+K_L+r_L+g_L(L) \ge g(L-1)$ with $L > 0$. Hence, $2^{2^{L-1}} \le g(L-1) \le T$, which implies $L \le \log_2 \log_2 T + 1$. Summarizing, we get Eq. (4). $\square$

A.2 Missed proofs from Section 5

A.2.1 Proof of Corollary 1

Before the proof, we remind the definition of a regular weakly consistent algorithm and of a dense algorithm.

Definition A.1 ([2]). *A weakly consistent algorithm $\mathcal{A}$ is said to be* regular *($\mathcal{A}$ in the class* **RWC***) if, for any node $\mathfrak{n} \in \mathfrak{T}(\mathcal{A})$:*

- *when* $p^{\mathfrak{l}(\mathfrak{n})} = p^{\mathfrak{n}} = p^{\mathfrak{r}(\mathfrak{n})}$, $[p^{\mathfrak{m}} = p^{\mathfrak{n}} \ \ \forall \mathfrak{m} \in \mathfrak{R}(\mathfrak{l}(\mathfrak{n})) \cup \mathfrak{L}(\mathfrak{r}(\mathfrak{n}))]$ *or* $[\mathfrak{L}(\mathfrak{n}) \cong \mathfrak{R}(\mathfrak{n})]$;
- *when* $p^{\mathfrak{l}(\mathfrak{n})} = p^{\mathfrak{n}} \ne p^{\mathfrak{r}(\mathfrak{n})}$, $[p^{\mathfrak{m}} = p^{\mathfrak{n}} \ \ \forall \mathfrak{m} \in \mathfrak{R}(\mathfrak{l}(\mathfrak{n}))]$ *or* $[\mathfrak{R}(\mathfrak{l}(\mathfrak{n})) \cong \mathfrak{R}(\mathfrak{n})]$;
- *when* $p^{\mathfrak{l}(\mathfrak{n})} \ne p^{\mathfrak{n}} = p^{\mathfrak{r}(\mathfrak{n})}$, $[p^{\mathfrak{m}} = p^{\mathfrak{n}} \ \ \forall \mathfrak{m} \in \mathfrak{L}(\mathfrak{r}(\mathfrak{n}))]$ *or* $[\mathfrak{L}(\mathfrak{r}(\mathfrak{n})) \cong \mathfrak{L}(\mathfrak{n})]$.

An algorithm $\mathcal{A}$ is said to be *dense* if the set of its prices $\wp(\mathcal{A})$ is dense in $[0, 1]$ (i.e., $\overline{\wp(\mathcal{A})} = [0, 1]$).

Proof. If the algorithm $\mathcal{A}$ is not dense, then the theorem holds since any non-dense horizon-independent regular weakly consistent algorithm has linear strategic regret (see [2, Cor.1]). First, let us consider the case when the first offered price $p_1 := p^{\mathfrak{e}(\mathfrak{T}(\mathcal{A}))} \in (0, 1)$ and show existence of a path $\tilde{\mathbf{a}}$ in the tree $\mathfrak{T}(\mathcal{A})$ that satisfies Eq. (7) from Definition 6.

Indeed, since $\mathcal{A}$ is dense there exists a node $\mathfrak{n} \in \mathfrak{T}(\mathcal{A})$ s.t. $p^{\mathfrak{n}} \in (0, p_1)$; let us take the one with the smallest depth $t^{\mathfrak{n}}$, denote $p' := p^{\mathfrak{n}}; t' := t^{\mathfrak{n}}$, and consider the path $\hat{\mathbf{a}}_{1:t'-1}$ from the root to this node $\mathfrak{n}$. For the corresponding price sequence $\{\hat{p}_t\}_{t=1}^{t'}$, the following holds:

- $\hat{p}_t \le p_1 \ \forall t \le t'$ due to the weak consistency of the algorithm $\mathcal{A}$;
- $\hat{p}_t \in \{0, p_1\} \ \forall t < t'$ due to the choice of the node $\mathfrak{n}$ with minimal $t^{\mathfrak{n}}$.

Since the algorithm $\mathcal{A}$ is regular weakly consistent, for any path $\mathbf{a}$ from the root s.t. its price sequence $\{p_t\}_{t=1}^{\infty}$ contains a price lower than p_1, the price sequence $\{p_t\}_{t=1}^{\infty}$ must be similar to $\{\hat{p}_t\}_{t=1}^{t'}$ at the beginning. Namely, there exists a node $\mathfrak{m} \in \mathfrak{T}(\mathcal{A})$ s.t. the path $\mathbf{a}$ passes through this node $\mathfrak{m}$, $p^{\mathfrak{m}} = p_{t^{\mathfrak{m}}} = p'$, and $p_t \in \{0, p_1\} \ \forall t < t^{\mathfrak{m}}$. Moreover, $\mathfrak{T}(\mathfrak{m}) \cong \mathfrak{T}(\mathfrak{n})$ since $\mathcal{A} \in$ **RWC** as well. Hence, if $\wp(\mathfrak{T}(\mathfrak{n})) \cap (0, p') = \varnothing$, then $\wp(\mathfrak{T}(\mathcal{A})) \cap (0, p') = \varnothing$, that contradicts to the density of the algorithm $\mathcal{A}$. Therefore, there exists a node $\hat{\mathfrak{n}} \in \mathfrak{T}(\mathfrak{n})$ s.t. $p^{\hat{\mathfrak{n}}} \in (0, p')$. Continuing the path $\hat{\mathbf{a}}_{1:t'-1}$ to this node $\hat{\mathfrak{n}}$, one gets the desired path $\tilde{\mathbf{a}}$ in the tree $\mathfrak{T}(\mathcal{A})$ that satisfies Eq. (7) from Definition 6. Hence, Theorem 2 implies a linear strategic regret for the algorithm $\mathcal{A}$ in the case $p_1 \in (0, 1)$.

Let us consider the case of $p_1 = 0$ or 1. Since $\mathcal{A}$ is dense, then, there exists a node $\mathfrak{n} \in \mathfrak{T}(\mathcal{A})$ such that $p^{\mathfrak{n}} \in (0, 1)$; we denote by $\tilde{\mathfrak{n}}$ the one among them with the smallest depth $t^{\mathfrak{n}}$. So, the problem of strategic regret estimation reduces to the previously considered case of $0 < p_1 < 1$ and resolves by replacing p_1 with $p^{\tilde{\mathfrak{n}}}$ in our reasoning. The only one thing left to be proven is that the optimal buyer strategy will either pass trough the pricing of $\mathfrak{T}(\tilde{\mathfrak{n}})$, or will have a linear regret.

Let $\mathfrak{n}_1 \to \ldots \to \mathfrak{n}_{\tilde{t}}$ be the path from the root $\mathfrak{n}_1 = \mathfrak{e}(\mathfrak{T}(\mathcal{A}))$ to the node $\mathfrak{n}_{\tilde{t}} = \tilde{\mathfrak{n}}$. If, for some $t = 1, \ldots, \tilde{t} - 1$, we have $p^{\mathfrak{n}_t} = p^{\mathfrak{n}_{t+1}}$, then $p^{\mathfrak{r}(\mathfrak{n}_t)} = p^{\mathfrak{l}(\mathfrak{n}_t)}$, since, otherwise, by the regularity of $\mathcal{A}$ (see Definition A.1), we would have:

- either $\mathfrak{R}(\mathfrak{n}_t) = \mathfrak{R}(\mathfrak{n}_{t+1})$ for $p^{\mathfrak{n}_t} = 0$ ($\mathfrak{L}(\mathfrak{n}_t) = \mathfrak{L}(\mathfrak{n}_{t+1})$ for $p^{\mathfrak{n}_t} = 1$), that contradicts to the definition of $\tilde{\mathfrak{n}}$ with the smallest depth;
- or $p^{\mathfrak{m}} = p^{\mathfrak{n}_t}\ \forall \mathfrak{m} \in \mathfrak{T}(\mathfrak{n}_{t+1})$, that contradicts to the existence of $\tilde{\mathfrak{n}}$ with $p^{\tilde{\mathfrak{n}}} \in (0, 1)$.

Hence, in this case of $p^{\mathfrak{n}_t} = p^{\mathfrak{n}_{t+1}}$, by regularity of $\mathcal{A}$, we have that:

- either the buyer decision at the node $\mathfrak{n}_t$ does not affect the further pricing: $\mathfrak{R}(\mathfrak{n}_t) = \mathfrak{L}(\mathfrak{n}_t)$, i.e., the optimal buyer strategy may not pass exactly through the edge $\mathfrak{n}_t \to \mathfrak{n}_{t+1}$, but, if the buyer select the other edge from the node $\mathfrak{n}_t$, he will face the subtree which is price equivalent to the subtree $\mathfrak{T}(\mathfrak{n}_{t+1})$;
- or $\mathfrak{n}_{t+1} = \mathfrak{r}(\mathfrak{n}_t)$ for $p^{\mathfrak{n}_t} = 0$ ($\mathfrak{n}_{t+1} = \mathfrak{l}(\mathfrak{n}_t)$ for $p^{\mathfrak{n}_t} = 1$) and $p^{\mathfrak{m}} = p^{\mathfrak{n}_t}\ \forall \mathfrak{m} \in \mathfrak{L}(\mathfrak{n}_t)$ ($\forall \mathfrak{m} \in \mathfrak{R}(\mathfrak{n}_t)$, resp.); thus, if the optimal strategy passes through the alternative node $\mathfrak{l}(\mathfrak{n}_t)$ ($\mathfrak{r}(\mathfrak{n}_t)$, resp.), then the seller will get a linear regret.

If $p^{\mathfrak{n}_{t+1}} \neq p^{\mathfrak{n}_t} = 0$, $t = 1, .., \tilde{t} - 1$, then, again by the regularity of $\mathcal{A}$, any sub-strategy in the left subtree $\mathfrak{L}(\mathfrak{n}_t)$ (a path starting from $\mathfrak{l}(\mathfrak{n}_t)$, i.e., from the alternative to the choice of the right child $\mathfrak{n}_{t+1}$ decision) has one of the following forms:

- (a) there is no any acceptance;
- there is an acceptance and after the first acceptance the buyer either
  - (b) will receive pricing of the tree $\mathfrak{R}(\mathfrak{n}_t)$; or
  - (c) will always receive the price 0.

If the buyer uses a strategy from the cases (a) and (c), then the seller will get a linear regret. The case (b) means that the algorithm $\mathcal{A}$ will behave similarly whenever the buyer accepts the price 0: at the round $t^{\mathfrak{n}_t}$ or after several rejections. Hence, the strategic buyer will accept 0 at the round $t^{\mathfrak{n}_t}$ (i.e., the buyer follows the edge $\mathfrak{n}_t \to \mathfrak{n}_{t+1}$). The examination of the case $p^{\mathfrak{n}_{t+1}} \neq p^{\mathfrak{n}_t} = 1, t = 1, .., \tilde{t} - 1$ is similar. □

B Details on prePRRFES

We present the pseudo-code for the learning algorithm prePRRFES in Algorithm B.1.

Algorithm B.1 Pseudo-code of the prePRRFES

Input: $r \in \mathbb{N}$ and $g : \mathbb{Z}_+ \to \mathbb{Z}_+$
Initialize: $q := 0$, $p := 1/2$, $l := 0$
while the buyer plays **do**
  Offer the price q to the buyer
  if the buyer accepts the price **then**
    $q := p$
  else
    Offer the price q to the buyer for $r - 1$ rounds
    if the buyer accepts one of the prices **then**
      go to line B.1
    end if
    Offer the price q to the buyer for $g(l)$ rounds
    $l := l + 1$
  end if
  if $p < 1$ **then**
    $p := q + 2^{-2^l}$
  end if
end while

C Algorithm announcement and an example from practice

Announcement of an algorithm to the buyer in advance is also known as commitment for a pricing. Note that the studies [6, 1, 3] showed that the seller earns noticeably less revenue in settings without commitment than with it. In particular, when the seller is faced with one buyer and does not commit for a pricing algorithm, the buyer in perfect Bayesian equilibrium[3] rejects goods all rounds except a low number of last ones. These studies constitute economic arguments for the seller to commit for a pricing algorithm and to do the best to assure the buyer that the commitment will not be violated in practice.

The most popular global ad exchanges face with instances of our game, that can be described by the following example: an Internet user searches for an apartment for rent; an advertiser (with an ad about apartments) targets this user. An ad exchange (seller) tracks this user each time she visits web sites related to the rent intent. Each view of a web page with a vacant ad slot by this user is a round ($t = 1, 2, ..$) in a sequence of posted price auctions between the seller and the advertiser (buyer). The advertiser holds fixed valuation v for a view of this user of his ad about apartments until the user holds the rent intent. The discount rate γ is the probability that the user will still search for an apartment for rent at the next round.

In practice of ad exchanges, many thousands of instances of our game are performed each day. In this case, the buyer believes that the seller will follow the committed algorithm, since the seller does it all previous instances; on the other hand, the seller is incentivized to follow his commitment, since its violation will incur significant losses due to loss of trust from advertisers. So, the seller needs to assure the buyer that the commitment will hold, what is very difficult task: buyers may doubt that the announced algorithm is used by the seller [1] and, e.g., may try to test the commitment [4].

A right-consistent (RC) but non-WC algorithm, such as PRRFES, revises earlier rejected prices and may offer higher ones after, what can raise doubts in the buyer on commitment. In this study, we try to find an algorithm that is less confusing for the buyer. We believe that the full consistency (both right, and left) is less confusing, because this property is easier to check by the buyer (he is

[3]Perfect Bayesian equilibrium is the main equilibrium concept that is used to study repeated games with incomplete information.

able to check, when he plays thousands of instances of our game within a day, as in our example above). Therefore, we seek for a WC algorithm. However, we found an optimal algorithm with an even less confusing property than WC: it never decreases offered prices (the algorithm prePRRFES).

D Examples of algorithms

D.1 Example: consistency VS weak consistency

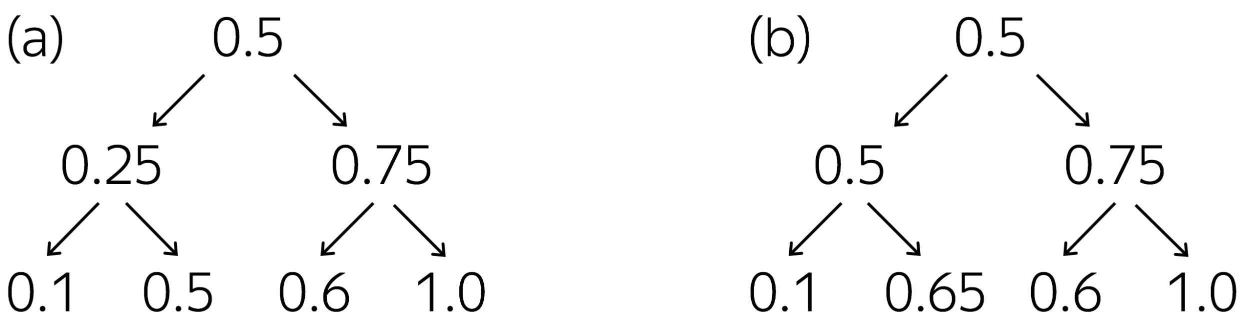

Figure D.1: Examples of algorithms (represented as binary trees for 3 rounds): (a) a consistent algorithm; (b) a weakly consistent algorithm, which is not consistent one (because 0.65 can be offered after rejection of 0.5 at the first round).

D.2 Example: source algorithm VS its `pre`-transformation

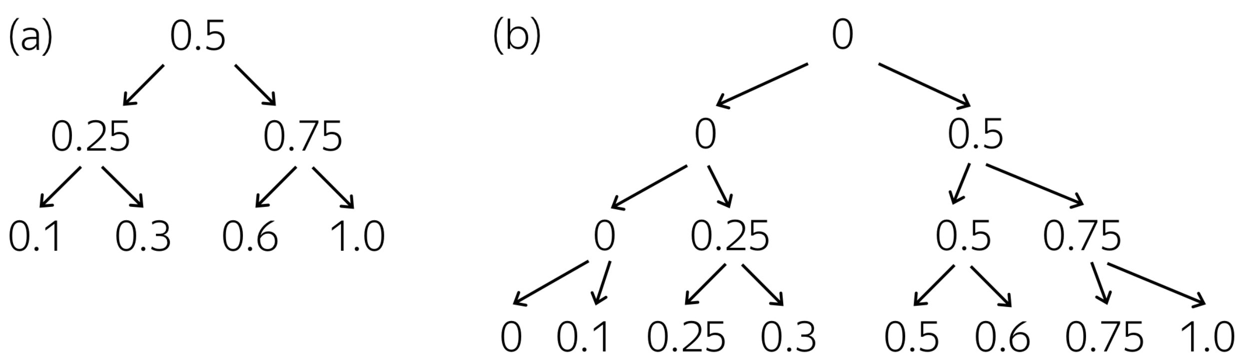

Figure D.2: Example of (a) a pricing algorithm $\mathcal{A}$ and (b) its `pre`-transformation $\texttt{pre}(0, \mathcal{A})$ (both algorithms are represented as binary trees).

References

[1] N. R. Devanur, Y. Peres, and B. Sivan. Perfect bayesian equilibria in repeated sales. In *SODA'2015*, pages 983–1002, 2015.

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

[3] N. Immorlica, B. Lucier, E. Pountourakis, and S. Taggart. Repeated sales with multiple strategic buyers. In *EC'2017*, pages 167–168, 2017.

[4] S. Lahaie, A. M. Medina, B. Sivan, and S. Vassilvitskii. Testing incentive compatibility in display ad auctions. In *WWW'2018*, 2018.

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

[6] K. M. Schmidt. Commitment through incomplete information in a simple repeated bargaining game. *Journal of Economic Theory*, 60(1):114–139, 1993.