
Online Pricing with Strategic and Patient Buyers (Supplementary Material)

Michal Feldman

Tel-Aviv University and MSR Herzliya
michal.feldman@cs.tau.ac.il

Tomer Koren

Google Brain
tkoren@google.com

Roi Livni

Princeton University
rlivni@cs.princeton.edu

Yishay Mansour

Tel-Aviv University
mansour@tau.ac.il

Aviv Zohar

Hebrew University of Jerusalem
avivz@cs.huji.ac.il

A Proofs for Lower Bound

In this section we prove Theorem 2. We follow the two step proof idea presented in the main text. For simplicity we begin by assuming $\hat{\tau} = 1$. The general case is dealt with in Section A.3.1

A.1 Reduction from switching cost

A.1.1 proof of Lemma 4

By definition we have that:

$$\mathbb{E}_{\bar{\mathbf{b}}_{1:T}} \text{Regret}_T(A; \bar{\mathbf{b}}_{1:T}) = \mathbb{E} \left[\max_p \sum_{t=1}^T R(p, \dots, p, \bar{\mathbf{b}}_1, \dots, \bar{\mathbf{b}}_t) - \sum_{t=1}^T R(p_{t-1}, p_t, p_{t+1}, \bar{\mathbf{b}}_1, \dots, \bar{\mathbf{b}}_t) \right]$$

By definition of R we can rewrite the terms as:

$$\mathbb{E} [\text{Regret}_T(A; \bar{\mathbf{b}}_{1:T})] = \mathbb{E} \left[\max_p \sum_{t=1}^T \beta(p, p; \bar{\mathbf{b}}_t) - \sum_{t=1}^T \beta(p_t, p_{t+1}; \bar{\mathbf{b}}_t) \right].$$

Let p^* be the maximizer of the non noisy sequence of buyers:

$$p^* = \arg \max_{p^*} \sum_{t=1}^T \beta(p^*, p^*; \mathbf{b}_t) = \sum p^* \mathbb{1}(v_t \leq p^*)$$

Then

$$\text{Regret}_T(A; \bar{\mathbf{b}}_{1:T}) \geq \mathbb{E} \left[\sum_{t=1}^T \beta(p^*, p^*; \bar{\mathbf{b}}_t) - \sum_{t=1}^T \beta(p_t, p_{t+1}; \bar{\mathbf{b}}_t) \right]$$

Note that p^* , p_t and p_{t+1} ² are independent on whether $\bar{\mathbf{b}}_t = \mathbf{b}_t$ or $\bar{\mathbf{b}}_t = \mathbf{z}_t$ hence

$$\begin{aligned} & \mathbb{E} \left[\sum_{t=1}^T \beta(p^*, p^*; \bar{\mathbf{b}}_t) - \sum_{t=1}^T \beta(p_t, p_{t+1}; \bar{\mathbf{b}}_t) \right] = \\ & \frac{1}{2} \mathbb{E} \left[\sum_{t=1}^T \beta(p^*, p^*; \mathbf{b}_t) - \beta(p_t, p_{t+1}; \mathbf{b}_t) \right] + \frac{1}{2} \mathbb{E} \left[\sum_{t=1}^T \beta(p^*, p^*; \mathbf{z}_t) - \beta(p_t, p_{t+1}; \mathbf{z}_t) \right] \end{aligned}$$

²Recall that the seller needs to publish one price a head in advance

Recall that if the value of \mathbf{z}_t is 1/2 her patience is $\tau = 0$ and we have $\beta(p_t, p_{t+1}; \mathbf{z}_t) = \beta(p_t, p_t; \mathbf{z}_t)$ and if \mathbf{z}_t has value 1 then $\beta(p_t, p_{t+1}; \mathbf{z}_t) = \beta(p_t, p_t; \mathbf{z}_t) - \frac{1}{2}\mathbb{1}(p_t \geq p_{t+1})$. Where

$$\mathbb{1}(p_t \geq p_{t+1}) = \begin{cases} 1 & p_t > p_{t+1} \\ 0 & \text{else} \end{cases}.$$

Taken together, and exploiting the fact that p_t and p_{t+1} are both independent of \mathbf{z}_t we have that $\mathbb{E}(\beta(p_t, p_{t+1}; \mathbf{z}_t)) = \mathbb{E}(\beta(\mathbf{z}_t; p_t, p_t)) - \frac{1}{4}\mathbb{1}(p_t \geq p_{t+1})$. Hence

$$E [\text{Regret}_T(A; \bar{\mathbf{b}}_{1:T})] \geq \frac{1}{2}\mathbb{E} \left[\sum_{t=1}^T \beta(p^*, p^*; \mathbf{b}_t) - \beta(p_t, p_{t+1}; \mathbf{b}_t) \right] + \frac{1}{2}\mathbb{E} \left[\sum_{t=1}^T \beta(p^*, p^*; \mathbf{z}_t) - \beta(p_t, p_t; \mathbf{z}_t) + \frac{1}{4}\mathbb{1}(p_t \geq p_{t+1}) \right]$$

Recall that \mathbf{b}_t has patience 0 hence we can write $\beta(p_t, p_{t+1}; \mathbf{b}_t) = \beta(p_t; \mathbf{b}_t)$.

Finally note that for any fixed price $\mathbb{E}(\beta(p, p; \mathbf{z}_t)) = \frac{1}{2}$. Since both p^* and p_t are independent of \mathbf{z}_t we have: $\mathbb{E}(\beta(p^*, p^*; \mathbf{z}_t)) = \mathbb{E}(\beta(p_t, p_t; \mathbf{z}_t))$, and we obtain the desired result,

$$\mathbb{E}_{\bar{\mathbf{b}}_{1:T}} [\text{Regret}_T(A; \bar{\mathbf{b}}_{1:T})] \geq \frac{1}{2}\mathbb{E} \left[\sum_{t=1}^T \beta(p^*; \mathbf{b}_t) - \beta(p_t; \mathbf{b}_t) \right] + \frac{1}{8}\mathbb{1}(p_t \geq p_{t+1}).$$

A.2 Pricing with switching cost

The aim of this section is to prove Theorem 6. Our proof relies on the following technical Lemma which we leave her proof to Section A.2.2:

Lemma 7. *Let \mathcal{F} be the class of pairs of transformations from $\{0, 1\}$ to revenues in $\{0, \frac{1}{2}, 1\}$ i.e.*

$$\mathcal{F} = \left\{ \mathbf{f} = (f^{(1)}, f^{(2)}) : f^{(i)} : \{0, 1\} \rightarrow \left\{0, \frac{1}{2}, 1\right\}, i = 1, 2 \right\}.$$

and let \mathcal{V} be the class of transformations from $\{0, 1\}^2$ to values in $\{0, \frac{1}{2}, 1\}$

$$\mathcal{V} = \left\{ v : v : \{0, 1\}^2 \rightarrow \left\{0, \frac{1}{2}, 1\right\} \right\}$$

There exist a distribution D over $\mathcal{F} \times \mathcal{V}$, that can be efficiently implemented and has the following properties:

1. For every pairs of bits $(a_1, a_2) \in \{0, 1\}^2$, if \mathbf{b} is a buyer with value $v = v(a_1, a_2)$ and patience 0 then:

$$\beta(1; \mathbf{b}) = f^{(1)}(a_1) \text{ and } \beta\left(\frac{1}{2}; \mathbf{b}\right) = f^{(2)}(a_2),$$

always.

2. For both $i = 1, 2$ we have

$$\mathbb{E}_{f \sim D} \left[f^{(i)}(x) \right] = \frac{1}{2} - \frac{1}{4}x.$$

A.2.1 Proof of Theorem 6

Let A be some seller against non-strategic buyers with bounded S_c -Regret $_T$. we will first construct an algorithm A' for the 2-action MAB problem with bounded S_c -Regret $_T$. Namely, we will have that for every sequence of losses ℓ_1, \dots, ℓ_T we can construct a sequence of non-strategic buyers such that:

$$S_c\text{-Regret}_T(A'; \ell_{1:T}) \leq S_c\text{-Regret}_T(A; \mathbf{b}_{1:T}).$$

In the reduction we are considering, Algorithm A' receives at each iteration t the price posted by algorithm A at step t and at each iteration algorithm A' chooses the feedback algorithm A observes. The algorithm is depicted in Algorithm 3.

Our algorithm A' work as follows: At the beginning of the iterations, the algorithm produces an IID sequence of pairs $(\mathbf{f}_1, v_1), \dots, (\mathbf{f}_T, v_T)$ as depicted in Lemma 7. At each iteration, the algorithm

receives from algorithm A a price $1/2$ or 1 . If the algorithm A posts price $\frac{1}{2}$ then algorithm A' chooses action 2, observes $\ell_t(2)$ and returns to algorithm A as feedback $f_t^{(2)}(\ell_t(2))$. By property 1 of Lemma 7 we have that A returns as feedback $f_t^{(2)}(\ell_t(2)) = \beta(\frac{1}{2}; \mathbf{b}_t)$. Similarly if algorithm A posts price 1 then algorithm A' chooses action 1, observes $\ell_t(1)$ and returns to algorithm A as feedback $f_t^{(1)}(\ell_t(1)) = \beta(1; \mathbf{b}_t)$.

Since algorithm A received at each iteration as feedback $\beta(p_t; \mathbf{b}_t)$ we have by assumption that:

$$\text{S}_c\text{-Regret}_T(A; \mathbf{b}_{1:T}) = \mathbb{E} \left(\max_{p^*} \sum_{t=1}^T \beta(p^*; \mathbf{b}_t) - \beta(p_t; \mathbf{b}_t) \right) + c\mathbb{E}(|\{p_t : p_t \neq p_{t+1}\}|)$$

Next, for every fixed action i and loss vectors ℓ_1, \dots, ℓ_T , we have by property 2 of Lemma 7:

$$\mathbb{E}_{\mathbf{f}} \left[\sum_{i=1}^T f_t^{(i)}(\ell_t(i)) \right] = \frac{T}{2} - \frac{1}{4} \sum_{i=1}^T \ell_t(i).$$

Thus, fix ℓ_1, \dots, ℓ_T and i^* and set $p^* = \begin{cases} 1 & i^* = 1 \\ \frac{1}{2} & i^* = 2 \end{cases}$:

$$\begin{aligned} \left(\sum_{t=1}^T (\ell_t(i_t)) - \sum_{t=1}^T (\ell_t(i^*)) \right) &= 4 \left(\mathbb{E}_{\mathbf{f}} \sum_{t=1}^T (f_t^{(i^*)}(\ell_t(i^*)) - f_t^{(i_t)}(\ell_t(i_t))) \right) = \\ 4\mathbb{E} \left(\sum_{t=1}^T \beta(p^*; \mathbf{b}_t) - \beta(p_t; \mathbf{b}_t) \right) &\leq 4\mathbb{E} \left(\max_{p^*} \sum_{t=1}^T \beta(p^*; \mathbf{b}_t) - \beta(p_t; \mathbf{b}_t) \right) \end{aligned}$$

Thus we have

$$\text{Regret}_T(A'; \ell_{1:T}) \leq \mathbb{E}_{\mathbf{b}_{1:T}} \left[4\text{S}_c\text{-Regret}_T(A; \mathbf{b}_{1:T}) - 4c\mathbb{E}(|\{p_t : p_{t+1} \neq p_t\}|) \right].$$

Next note that for every realization $\mathbf{b}_1, \dots, \mathbf{b}_T$ algorithm A and algorithm A' have the same number of switching taken together we obtain that for algorithm A' :

$$\text{S}_{4c}\text{-Regret}_T(A'; \ell_{1:T}) \leq \mathbb{E}_{\mathbf{b}_{1:T}} \left[\text{S}_c\text{-Regret}_T(A; \mathbf{b}_{1:T}) \right] \leq \text{S}_c\text{-Regret}_T(A).$$

Finally, since the result holds for every $\ell_{1:T}$ we obtain the desired result from Theorem 5.

Algorithm 3: Reduction from MAB with switching cost to pricing with switching cost

Input: T , A % A is an algorithm with bounded regret for pricing with switching cost;

Output: i_1, \dots, i_T ;

Draw IID $(\mathbf{f}_1, v_1), \dots, (\mathbf{f}_T, v_T) \sim D$ % see Lemma 7;

for $t=1, \dots, T$ **do**

 Receive from A a posted price p_t ;

if $p_t = 1$ **then**

 Set $i_t = 1$

else

 Set $i_t = 2$

 Play action i_t and receive as feedback $\ell_t(i_t)$;

 Return to A as feedback $f_t^{(i_t)}(\ell_t(i_t))$; %Note that $f_t^{(i_t)}(\ell_t(i_t)) = \beta(p_t; \mathbf{b}_t)$

A.2.2 Proof of Lemma 7

We begin by constructing \mathbf{f} : We choose \mathbf{f} as follow:

- With probability $\frac{1}{4}$ we let: $f^{(1)}(\ell) = 1 - \ell$ and $f^{(2)}(\ell) = \frac{1}{2}$.
- With probability $\frac{1}{4}$ we let: $f^{(1)}(\ell) = 1$ and $f^{(2)}(\ell) = \frac{1}{2}$.
- With probability $\frac{1}{2}$ we let: $f^{(1)}(\ell) = 0$ and $f^{(2)}(\ell) = \frac{1}{2}(1 - \ell)$.

The random variable ν is then defined as a function of \mathbf{f} and for which 1 holds. To define ν note that for any feasible realization of \mathbf{f} and any two bits a_1, a_2 we have $(f^{(1)}(a_1), f^{(2)}(a_2)) \in \{(0, 0), (0, \frac{1}{2}), (1, \frac{1}{2})\}$. Thus we can define $\nu(a_1, a_2)$ as a function of \mathbf{f} as follows:

$$\nu(a_1, a_2; \mathbf{f}) = \begin{cases} 0 & (f^{(1)}(a_1), f^{(2)}(a_2)) = (0, 0) \\ \frac{1}{2} & (f^{(1)}(a_1), f^{(2)}(a_2)) = (0, \frac{1}{2}) \\ 1 & (f^{(1)}(a_1), f^{(2)}(a_2)) = (1, \frac{1}{2}) \end{cases}$$

A.3 Proof of Theorem 2 for $\hat{\tau} = 1$

Let A be some seller for buyers with patience at most $\hat{\tau} = 1$. Let A' be the algorithm whose existence follows from Lemma 4. A' is an algorithm against non strategic sellers, and by Theorem 6

$$S_{\frac{1}{12}}\text{-Regret}_T(A') = \Omega(T^{2/3}).$$

By Lemma 4 $\text{Regret}_T(A) \geq S_{\frac{1}{12}}\text{-Regret}_T(A')$.

A.3.1 Generalization to arbitrary $\hat{\tau}$

In this section we prove Theorem 2, in which the lower bound has dependence on $\hat{\tau}$, namely $\Omega(\hat{\tau}^{1/3}T^{2/3})$. For simplicity we will assume $\hat{\tau}$ is odd (we can always construct an adversary with $\hat{\tau} - 1$).

We will restrict ourselves to adversaries with the following properties:

1. The adversary divides the interval T to $\frac{\hat{\tau}+1}{2}$ blocks and at each block the adversary chooses a constant value for all buyers.
2. The adversary only chooses buyers with $\tau = 0$ or with patience up to the end of the next block (since the size of the blocks is $\frac{\hat{\tau}+1}{2}$ he can always do that).
3. If the patience of the buyer is not 0 then the buyer has maximal value $\nu = 1$.

To further simplify things, we will strengthen our seller and allow him to choose all the prices in the next block at the beginning of the block before (this only strengthen him since he can now delay the posting of some of the prices). Note that since the patience of all buyers is only one block ahead their revenue is well defined even for this type of buyer.

Next note, the given such buyers and sellers the seller can choose a fixed price throughout the block. Indeed, if the optimal seller chooses price $p_t, p_{t+1}, \dots, p_{t+\frac{\hat{\tau}+1}{2}}$ for some block: the expected revenue for those prices and for picking one of them randomly is the same (as the buyers are fixed within a block). Further, since the buyer has $\tau_t \neq 0$ only if his value is maximal, by choosing an expected price he only gains in terms of revenue from buyers and previous blocks.

Taken together, we reduced the problem to a setting of $\hat{\tau} = 1$ — both adversary and seller choose at each block a fixed buyer and price. However, the revenue per round is now multiplied by a factor of $\frac{\hat{\tau}+1}{2}$.

The only tackle is that we restrict ourselves to adversarial sequences where buyer has patience different then zero iff his value is maximal. Luckily one can see in our construction that this is indeed the case for our adversarial buyers.

Taken together with number of blocks we obtain:

$$\text{Regret}_T \geq \frac{\hat{\tau} + 1}{2} \left(\frac{2T}{\hat{\tau} + 1} \right)^{2/3} = \Omega(\hat{\tau}^{1/3}T^{2/3}).$$