Technical Report
TTI-TR-2008-3

The Price of Truthfulness for Pay-Per-Click Auctions

Nikhil Devanur
Microsoft Research
Sham M. Kakade
Toyota Technological Institute at Chicago

ABSTRACT

We analyze the problem of designing a truthful pay-per-click auction where the click-through-rates (CTR) of the bidders are unknown to the auction. Such an auction faces the classic explore/exploit dilemma: while gathering information about the click through rates of advertisers, the mechanism may loose revenue; however, this gleaned information may prove valuable in the future for a more profitable allocation. In this sense, such mechanisms are prime candidates to be designed using multi-armed bandit techniques. However, a naive application of multi-armed bandit algorithms would not take into account the *strategic* considerations of the players — players might manipulate their bids (which determine the auction's revenue) in a way as to maximize their own utility. Hence, we consider the natural restriction that the auction be truthful.

The revenue that we could hope to achieve is the expected revenue of a Vickerey auction that knows the true CTRs, and we define the *2nd price regret* to be the difference between the expected revenue of the auction and this Vickerey revenue. This work sharply characterizes what regret is achievable, under a truthful restriction. We show that this truthful restriction imposes *statistical* limits on the achievable regret — the achievable regret is $\Theta^*(T^{2/3})$, while for traditional bandit algorithms (without the truthful restriction) the achievable regret is $\Theta^*(T^{1/2})$ (where T is the number of rounds). We term the extra $T^{1/6}$ factor, the 'price of truthfulness'.

1 Introduction

Pay-per-click auctions are the workhorse auction mechanism for web-advertising. In this paradigm, advertisers are charged only when their displayed ad is 'clicked' on (see Lahaie et al. [2007] for a survey). In contrast, more traditional 'pay-per-impression' schemes charge advertisers each time their ad is displayed. Such mechanisms are appealing from an advertisers standpoint as the advertiser now only has to gauge how much they value someone actually viewing their website (after a click) vs. just looking at their ad (one may expect the former to be easier as it is closer to the outcome that the advertiser desires). From a mechanism design standpoint, we clearly desire a mechanism which elicits advertisers preferences in a manner that is profitable.

A central underlying issue in these pay-per-click auctions is estimating which advertisers tend to get clicked on more often. Naturally, whenever the mechanism displays an ad which is not clicked, the mechanisms receives no profit. However, the mechanism does obtain information which is potentially important in estimating how often that advertiser gets clicked (the 'click through rate' of an advertiser). In this sense, the mechanism faces the classic explore/exploit tradeoff: while gathering information about the click through rate of an advertiser, the mechanism may loose revenue; however, this gleaned information may prove valuable in the future for a more profitable allocation.

The seminal work of Robbins [1952] introduced a formalism for studying this exploration/exploitation tradeoff, which is now referred to as the *multi-armed bandit* problem. In this foundational paradigm, at each time step a decision maker chooses one of n decisions or 'arms' (e.g. treatments, job schedules, manufacturing processes, etc) and receives some feedback loss only for the chosen decision. In the most unadorned model, it is assumed that the cost for each decision is independently sampled from some fixed underlying (and unknown) distribution (that is different for each decision). The goal of the decision maker is to minimize the average loss over some time horizon. This *stochastic* multi-armed bandit problem and a long line of successor bandit problems have been extensively studied in the statistics community (see, e.g., [Auer et al., 2002]), with close attention paid to obtaining sharp convergence rates.

In our setting, we can model this pay-per-click auction as a multi-armed bandit problem as follows: Say we have n advertisers and say advertiser i is willing to pay up to v_i per click (the advertisers value, which, for now, say is constant), then at each round the mechanism chooses which advertiser to allocate to (i.e. it decides which arm to pull) and then observes if that ad was clicked on. In this idealization, the mechanism simply has one 'slot' to allocate each round. If each advertiser had some click through ρ_i (the i.i.d. probability that i's ad will be clicked if displayed), then the maximal revenue the mechanism could hope to achieve on average would be $\max_i \rho_i v_i$, if i actually paid out v_i per click. Hypothetically, let us assume that i actually paid out v_i per click, but the mechanisms does not know ρ_i — the exploration/exploitation tradeoff is in estimating ρ_i accurately vs. using these estimates to obtain revenue. Hence, if we run one of the proficient bandit algorithms (say the upper confidence algorithm of Auer et al. [2002]), then the mechanism can guarantee that the difference between its revenue after T rounds and the maximal possible revenue of $T \max_i \rho_i v_i$ would be no more than $O^*(\sqrt{nT})$ (this difference is known as the 'regret'). What this argument does not take into account is the strategic motivations of the advertisers. With strategic considerations in mind, any mechanism only receives the advertisers purported value, their 'bid' b_i (and so the advertiser knows that i is only willing to pay up to b_i). Here, it is no longer clear which of our classic multi-armed bandit algorithms are appropriate, since an advertiser (with knowledge of the mechanism) might find it to be more profitable to bid a value $b_i \neq v_i$.

The focus of this paper is to understand this exploration-exploitation tradeoff in a strategic setting. The difficulty now is that our bandit mechanism must now be truthful, so as to disallow advertisers from manipulating the system. We are particularly concerned with what is achievable, under these constraints. Our results show that for pay-per-click auctions this truthful restriction places fundamental restrictions on what is *statistically* achievable.

1.1 Summary

The most immediate question is what is it reasonable to compare to? Certainly, $T \max_i \rho_i v_i$ is not reasonable, since even in a one shot full information setting (where ρ_i is known) this revenue not attainable, without knowledge of the actual values. In a single shot ($T = 1$) full information setting, what is reasonable to obtain is the revenue $\mathrm{smax}_i \rho_i v_i$, where smax is the operator which takes the *second* largest value. Hence, in a T round setting, the natural revenue for the mechanism to seek to obtain is $T \mathrm{smax}_i \rho_i v_i$. If ρ_i were known, it straightforward to see that such revenue could *always* be obtained (without knowledge of the true values) with a truthful mechanism.

In this paper, we introduce the notion of 2nd-price regret: the difference between the mechanisms revenue and $T \mathrm{smax}_i \rho_i v_i$. This quantity is the natural generalization of the notion of regret to a setting where truthfulness is imposed. Analogous to the usual bandit setting, the goal of the mechanism is obtain a sublinear (in T) 2nd-price regret, but the mechanism now has the added constraint of being truthful (ensuring that advertisers are not manipulating the mechanism, in a rudimentary sense).

This paper sharply characterizes this 2nd price regret. Our first result shows that a rather simple explore/exploit (truthful) strategy achieves sublinear 2nd price regret. This strawman mechanism simply explores for a certain number of rounds (charging nothing). After this exploration phase, this mechanism exploits by allocating the slot to the estimated highest revenue bidder for the remainder of the rounds (the estimated highest revenue bidder is determined with the empirical click through rate, which is estimated from the exploration phase). This bidder is charged a quantity analogous to the second price (the quantity charged is the second highest expected revenue), and this price is also determined by empirical estimates of the click through rate. The 2nd price regret achieved by mechanism is $O^*(n^{1/3}T^{2/3})$.

The immediate question is can we do better? In the traditional bandit settings, such explicit explore/exploit schemes perform unfavorably, and more sophisticated schemes achieve regret of $O^*(\sqrt{nT})$ (see Auer et al. [2002]). These mechanisms typically do not make a distinction between exploiting or exploring — they implicitly make this tradeoff. Roughly speaking, one of the difficulties in using these more sophisticated algorithms for pay-per-click auctions is determining how to charge — truthful mechanisms often determine prices based on properties of the non-winning bidders (thus sampling the highest bidder too often might lead to not having enough accuracy for charging him appropriately).

The main technical result in this paper is a lower bound which this formalizes intuition, showing that any mechanism must have regret $\Omega(T^{2/3})$. Roughly speaking, the proof technique shows that any pay-per-click auction must have the property that it behaves as an 'explore/exploit' algorithm, where when it explores, it must charge zero, and when it exploits, it cannot use this information for setting future prices.

The proof techniques go through the results on truthful pricing (see Myerson [1981], Hartline and Karlin [2007]), which (generally) characterize how to truthfully price any allocation scheme. The additional constraint we use on this truthful pricing scheme is an informational one — the auction must only use information from the observed allocations. We expect our proof technique to be more generally useful for other mechanisms, since information gathering in a strategic setting is somewhat generic. Our technique shows how to obtain restrictions on the pricing scheme, based on *both* truthfulness and bandit feedback.

We characterize this loss in revenue (in comparison to a bandit setting) as 'the price of truthfulness'. This (multiplicative) gap between the regret achievable is $O^*(T^{1/6})$.

1.2 Related Work

Gonen and Pavlov [2007] consider the same problem but the goal was simply to maximize social welfare. They work in a related framework, where the advertisers place a single bid at the start of the auction, which stands for the full T rounds. However, contrary to their claims, their auction is not truthful, even for a single slot[1].

Nazerzadeh et al. [2008] consider a similar problem, where the goal is to design a truthful Pay-per-acquisition auction — the key difference being that the bidders report whether an acquisition happened or not. Their auction employs an explore/exploit approach similar to our upper bound. In this work, they do not consider nor analyze the optimal achievable rate. We expect that our techniques also imply lower bounds on what is statistically achievable in their setting.

[1] We have confirmed this through personal communication. For the allocation given by their auction, there is a unique pricing that would make it truthful, but this price depends on the clicks that are not observed by the auction (which is what our lower bound techniques imply). In fact, this is one of the insights used in proving our lower bounds.

2 The Model

Here we define the model for a single-slot Pay-per-click (PPC) auction. We consider a repeated auction, where a single slot is auctioned in each of T time steps. There are n advertisers, each of whom values a 'click', while the auction can only assign 'impressions'. The auction proceeds as follows. At each round t, each advertiser bids a value b_i^t, which is their purported value of i per click at time t. Then, the auction assigns an impression to one of the n advertisers, e.g. the auction decides which ad will be displayed. We let x^t be this allocation vector, and say $x_i^t = 1$ iff the allocation is to advertiser i at time t (and $x_j^t = 0$ for all $j \neq i$, since only one advertiser is allocated). After this allocation, the auction then observes the event c_i^t which is equal to 1 if the item was clicked on and 0 otherwise. Crucially, the auction observes the click outcome only for allocated advertiser, i.e. c_i^t is observed iff $x_i^t = 1$. Also at the end of the round, the auction charges the advertiser i the amount p_i^t only if i was clicked. The revenue of the auction is $A = \sum_{i,t} p_i^t$.

Note the allocation x_i^t is a function of the bids and the observed clicks for $\tau < t$. Let $C = (c_i^t : i = 1..n, t = 1..T)$ be all click events, observed and otherwise. For the ease of notation, we only include those arguments of x_i^t that are relevant for the discussion (for example, if we write $x^t(b_i^t)$, then we may be explicitly considering the functional dependence on b_i^t, but one should keep in mind the implicit dependence on the other bids and the click history).

We assume advertiser i's 'true value' for a click at time t is v_i^t, which is private information. Then i derives a benefit of $\sum_t v_i^t c_i^t x_i^t$ from the auction. Hence, the utility of i is $\sum_t (v_i^t c_i^t x_i^t - p_i^t)$. An auction is *truthful* for a given sequence $C \in \{0,1\}^{n \times T}$, if bidding $v_i^t = b_i^t$ is a dominant strategy for all bidders: if for all possible bids of other advertisers $\{b_{-i}^t\}$, the utility of i is maximized when i bids $b_i^t = v_i^t$ for all t. As the auction depends on the advertisers previous bids, an advertiser could potentially try to manipulate their current bid in order to improve their *future* utility — this notion of truthfulness prevents such manipulation. An auction is *always truthful* if it is truthful for all $C \in \{0,1\}^{n \times T}$.

We work in a stochastic setting where the event that $c_i^t = 1$ is assumed to be i.i.d, with click probability ρ_i. This ρ_i is commonly referred to as the click-through rate (CTR) and is assumed to be constant throughout the auction. The auction has no knowledge of the CTRs of the advertisers prior to the auction.

Subject to the constraint of being always truthful, the goal of the auction is to maximize its revenue. Define $\mathrm{smax}_i\{u_i\}$ to be the second largest element of a set of numbers $\{u_i\}_i$. The benchmark we use to evaluate the revenue of the auction is as follows:

Definition 1. *Let*

$$OPT = \sum_{t=1}^{T} \mathrm{smax}_i\{\rho_i b_i^t\} .$$

It is the expected revenue of the Vickerey auction that knows the true ρ_i's. Let 2-Regret $:= OPT - \mathbb{E}_C[A]$ *be the expected* 2nd price regret *of the auction.*

We provide sharp upper and lower bounds for this quantity. We actually prove a lower bound for a stronger model: the static bid model. The key differences in this model are

- the bidders have $v_i^t = v_i$ for all t, and are only allowed to submit one bid, at the start.
- the auction could decide the payments of the bidders at the end of all the rounds.

Note that such auctions are more powerful and potentially have a lower regret. We show an identical lower bound for this case, which is thus a stronger statement. The proof is more technically demanding but follows a similar line of argument. Clearly, our upper bound holds in this model.

3 Main Results

Our first result is on the existence of an algorithm with sublinear (in T) 2nd-price regret.

Theorem 2. *Let $b_{max} = \max_{i,t} b_i^t$. There exists an always truthful PPC auction with 2-Regret this is $O(b_{max} n^{1/3} T^{2/3} \sqrt{\log(nT)})$.*

In the next section, we specify this mechanism and proof. The mechanism is essentially the strawman auction, which first explores for a certain number of rounds and then exploits. Here, we show such an auction is also always truthful.

For the n-arm multi-armed bandit mechanism, such algorithms typically also achieve a regret of the same order. However, in the n-arm bandit setting, there are sharper algorithms achieving regret of $O^*(\sqrt{nT})$ (see for example Auer et al. [2002]). Our second result (our main technical contribution) shows that such an improvement is not possible.

Theorem 3. *For every always truthful PPC auction (with $n = 2$), there exists a set of bids bounded in $[0, 1]$ and ρ_i such that 2-Regret $= \Omega(T^{1/3})$.*

In comparison to the multi-armed bandit problem, the requirement of truthfulness degrades the achievable statistical rate. In particular, the regret is larger by an additional $T^{1/6}$ factor, which we term 'price of truthfulness'.

In Section 3.3 we extend this lower bound to the static bid case, where the bidders submit a single value at the start, and the auction only charges at the end of the T rounds (rather than instantaneously).

3.1 Upper Bound Analysis

The algorithm is quite simple. For the first τ steps, the auction explores. By this we mean that the algorithm allocates the item to each bidder for $\lfloor \tau/n \rfloor$ steps (and it does so unadaptively in some deterministic order). All prices are 0 during this exploration phase. After this exploration phase is over, let $\hat{\rho}_i$ be the empirical estimate of the click through rate. With probability greater than $1 - \delta$, we have that the following upper bound holds for all i:

$$\rho_i \leq \hat{\rho}_i + \sqrt{2 \left\lfloor \frac{n}{\tau} \right\rfloor \log \frac{n}{\delta}} := \hat{\rho}_i^+$$

where we have defined $\hat{\rho}_i^+$ to be this upper bound. For the remainder of the timesteps, i.e. for $t > \tau$ (which is the exploitation phase), the auction allocates the item to the bidder i^* at time t which maximizes $\hat{\rho}_i^+ b_i^t$, i.e. the allocation is at time t is

$$x_{i^*}^t = 1 \text{ where } i^* = \arg\max_i \hat{\rho}_i^+ b_i^t$$

and the price charged to i^* at time t is:

$$p_i^t = \frac{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}{\hat{\rho}_{i^*}^+}$$

where smax is the second maximum operator.

It is straightforward to see that the auction is instantaneously truthful (i.e. an advertiser's revenue at any given round cannot be improved by changing the bid at that round). However, the proof also consists of showing that the auction is truthful over the T steps (in addition to proving the claimed regret bound).

Proof. We first provide the proof of truthfulness. Consider a set of positive weights w_i. First, note that we could construct a truthful auction with this vector w_i in the following manner: let the winner at time t be $i^* = \arg\max_i w_i b_i^t$ and charge i^* the amount $\frac{\mathrm{smax}_i w_i b_i^t}{w_{i^*}}$ this time. It is straightforward to verify that this auction is truthful for any click sequence and for any duration T. Now observe that the weights used by the auction are $w_i = \rho_i^+$ which are *not* a function of the bids. Hence, the auction is truthful since during the exploitation phase the auction is truthful (for any set of weights).

Now we bound the regret of the auction. Note that for all t after the exploration phase (all $t > \tau$), $\mathbb{E}\left[c_{i^*}^t\right] = \rho_{i^*}$. Hence, the expected revenue at time t is just $\frac{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}{\hat{\rho}_{i^*}^+}\rho_{i^*}$. First note by construction,

$$\frac{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}{\hat{\rho}_{i^*}^+} \leq b_{i^*}^t \leq b_{\text{max}}$$

and also note that with probability greater than $1 - \delta$:

$$\frac{\mathrm{smax}_i \rho_i b_i^t}{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t} \leq 1$$

since $\rho_i \leq \hat{\rho}_i^+$ (with probability greater than $1 - \delta$). Using these facts, the instantaneous regret is bounded follows:

$$\begin{aligned}
\mathrm{smax}_i \rho_i b_i^t - \frac{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}{\hat{\rho}_{i^*}^+}\rho_{i^*} &= \frac{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}{\hat{\rho}_{i^*}^+}\left(\frac{\mathrm{smax}_i \rho_i b_i^t}{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}\hat{\rho}_{i^*}^+ - \rho_{i^*}\right) \\
&\leq b_{\text{max}}\left(\frac{\mathrm{smax}_i \rho_i b_i^t}{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}\hat{\rho}_{i^*}^+ - \rho_{i^*}\right) \\
&\leq b_{\text{max}}\left(\hat{\rho}_{i^*}^+ - \rho_{i^*}\right) \\
&\leq b_{\text{max}}\sqrt{2\frac{n}{\tau}\log\frac{n}{\delta}}\ .
\end{aligned}$$

Hence, since there are $T - \tau$ exploitation phases and τ exploration phases (with no revenue), we have shown that the expected regret is:

$$\text{2-Regret} \leq b_{\text{max}}\left((T-\tau)\sqrt{2\frac{n}{\tau}\log\frac{n}{\delta}} + \tau + \delta T\right)$$

where the δT term comes from the failure probability. Choosing $\delta = 1/T$ and $\tau = n^{1/3}T^{2/3}\sqrt{\log(nT)}$ completes the proof. $\square$

3.2 Lower bound

We prove the lower bound in the setting where the values are constant for all times t. First, we characterize the restriction imposed on the allocation function by truthfulness.

3.2.1 Constraints from Truthful Pricing

The following theorem from Myerson [1981] (also see Hartline and Karlin [2007]) for characterizing truthful auctions will be extensively used. Since the values of the advertisers are assumed to be constant, one strategy the advertisers could take is to only consider changing their bid at the start of the auction — hence, the entire auction must be truthful with respect to the cumulative prices. Applying this theorem to the cumulative prices charged over the auction leads to the following pricing restriction:

Theorem 4. Truthful pricing rule: *Fix a click sequence. Let $y_i = \sum_t x_i^t c_i^t$ and let $p_i = \sum_t p_i^t$. If an auction x (which implies y) is truthful then*

1. *y_i is monotonically increasing in b_i*
2. *the price p_i charged to i is exactly*

$$p_i(\boldsymbol{b}) = b_i y_i(\boldsymbol{b}) - \int_{z=0}^{b_i} y_i(z, \boldsymbol{b}_{-i}) dz.$$

Also, let $y_i^t = x_i^t c_i^t$. Also define

$$p_i^t(\boldsymbol{b}) = b_i y_i^t(\boldsymbol{b}) - \int_{z=0}^{b_i} y_i^t(z, \boldsymbol{b}_{-i}) dz,$$

and note that $p_i = \sum_t p_i^t$. It is also straightforward to see that the truthful pricing rule also implies that these must be the instantaneous prices, and that instantaneously, the x_i^t must be monotonic in b_i^t. To see this, note that it could be the case that the current round is effectively the advertisers last round (say this advertisers true value drops to 0 for the remainder of the auction). Hence, the current round must also have an instantaneously truthful price.

Since the mechanism is always truthful, the allocation function has to be such that the prices can always be calculated exactly (with the observed clicks). Using this one can argue that the allocation function only depends on the clicks observed during certain 'non-competitive' time periods, which we now formalized.

3.2.2 Competitive Pricing

Recall, we only include those arguments of x_i^t that are relevant for the discussion (for example, if we write $x^t(b_i, b_{-i})$, then we may be explicitly considering the functional dependence on b_i and b_{-i}, but one should keep in mind the implicit dependence on the click history). From now on, we assume that there are only 2 bidders, 1 and 2. We also now restrict the bids to be constant for the duration of the auction.

A competitive round for bidder 1 is one in which there exists a high enough bid b_1 such that 1 can win. More formally,

Definition 5. *Say that a time τ is* competitive *(w.r.t bidder 1) if for all b_2, there exist b_1 so that $x_1^\tau(b_1, b_2) = 1$.*

We also consider the functional dependence on clicks:

Definition 6. *Say that the allocation x_1^t* depends on *c_2^τ if there exist b_1, b_2 such that $x_1^t(b_1, b_2, c_2^\tau) \neq x_1^t(b_1, b_2, 1 - c_2^\tau)$.*

Note that in order for x_1^t to have a functional dependence on c_2^τ, the auction must observe c_2^τ, in which case $x_2^\tau(b_1, b_2) = 1$.

Lemma 7. *If τ is competitive w.r.t bidder 1, then x_1^t does not depend on c_2^τ.*

Proof. Say τ is competitive and x_1^t depends on c_2^τ, for the bids b_1, b_2. Hence, the auction must observe c_2^τ, so clearly $x_2^\tau(b_1, b_2) = 1$. Since τ is competitive, there exist $b_1' > b_1$ so that $x_1^\tau(b_1', b_2) = 1$.

Since the auction is instantaneously truthful, x_1^t is monotone, and the mechanism has to calculate p_1^t. Now we will argue that $p_1^t(b_1', b_2, c_2^\tau) \neq p_1^t(b_1', b_2, 1 - c_2^\tau)$, which is a contradiction, since the mechanism does not observe c_2^τ at bids b_1', b_2, as $x_2^\tau(b_1', b_2) = 0$.

Consider the case $x_1^t(b_1, b_2, c_2^\tau) = 1$ and $x_1^t(b_1, b_2, 1 - c_2^\tau) = 0$. Also, by monotonicity, we have that $x_1^t(b_1', b_2, c_2^\tau) = 1$. We must also have $x_1^t(b_1', b_2, c_2^\tau) = x_1^t(b_1', b_2, 1 - c_2^\tau) = 1$ since the auction does not observe c_2^τ at these bids, i.e. $x_2^\tau(b_1', b_2) = 0$. Note that since $x_1^t \in \{0, 1\}$, by the truthful pricing rule we have for all bids b, $p_1^t(b, b_2) = \min\{b' \leq b : x_1^t(b', b_2) = 1\}$. Hence, $p_1^t(b_1', b_2, c_2^\tau) \leq b_1$. Since, $x_1^t(b_1, b_2, 1 - c_2^\tau) = 0$ and $x_1^t(b_1', b_2, 1 - c_2^\tau) = 1$, we also have that $p_1^t(b_1', b_2, 1 - c_2^\tau) > b_1$, which is a contradiction. The other case is identical ($x_1^t(b_1, b_2, c_2^\tau) = 0$ and $x_1^t(b_1, b_2, 1 - c_2^\tau) = 1$), and we are done. □

Lemma 8. *(Scale Invariance) We may assume w.l.o.g that the allocation is scale invariant, i.e., for all $\lambda > 0, x(\boldsymbol{b}) = x(\lambda \boldsymbol{b})$.*

Proof. We will show that given any allocation x that is not scale invariant, there is a scale invariant allocation x' with smaller regret. For all $\boldsymbol{b}$, define $\lambda(\boldsymbol{b}) := \arg\max_{\lambda > 0} A(\lambda \boldsymbol{b})/\lambda$ and let $x'(\boldsymbol{b}) := x(\lambda \boldsymbol{b})$. Note that $\lambda()$ is scale invariant, and so is x'. It is easy to see that the revenue $A'(\boldsymbol{b}) = A(\lambda \boldsymbol{b})/\lambda$, and hence by definition of λ, $A'(\boldsymbol{b}) \geq A(\boldsymbol{b})$. □

Corollary 9. *If τ is not competitive w.r.t. bidder 1, then $p_1^\tau \equiv p_2^\tau \equiv 0$.*

Proof. Since τ is not competitive, $\exists b_2 : \forall b_1$, $x_1^\tau(b_1, b_2) = 0$. By scale invariance, it follows that $\forall b_1' b_2'$, $x_1^\tau(b_1', b_2') = x_1^\tau(b_1' b_2 / b_2', b_2) = 0$. Thus $p_1^\tau \equiv 0$.

If $x_2^\tau(b_1, b_2) = 0$, then $p_2^\tau(b_1, b_2) = 0$. If $x_2^\tau(b_1, b_2) = 1$, then since τ is not competitive, for all $b_1' > b_1$, $x_2^\tau(b_1', b_2) = 1$. Because of scale invariance, it follows that for all $0 < b_2' \leq b_2$, $x_2^\tau(b_1, b_2') = x_2^\tau(b_1 b_2 / b_2', b_2) = 1$. Now $p_2^\tau(b_1, b_2) = 0$ by truthful pricing. □

Lemma 10. *Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be probability distributions on $\{0, 1\}^T$ generated by i.i.d samples w.p. $1/2 + \delta$ and $1/2 - \delta$ respectively. Then for all functions $\chi : \{0, 1\}^T \to \{0, 1\}$ that can be represented as decision trees of depth $n = o(1/\delta^2)$, either $\sum_{c \in \{0,1\}^T} \mathbb{P}_1(c)\chi(c)$ or $\sum_{c \in \{0,1\}^T} \mathbb{P}_2(c)(1 - \chi(c))$ is $\Omega(1)$.*

Proof. Assume that the decision tree is a complete binary tree. Each leaf of the tree is represented by a string $x \in \{0,1\}^n$. Let $\chi(x)$ be the output of the decision tree at leaf x. Then for any probability distribution on $c \in \{0,1\}^T$

$$\sum_{c\in\{0,1\}^T} \mathbb{P}(c)\chi(c) = \sum_{x\in\{0,1\}^n} \mathbb{P}(x)\chi(x)$$

where $\mathbb{P}(x)$ is the probability that decision tree reaches leaf x. $\mathbb{P}_1(x) = (1/2+\delta)^{|x|}(1/2-\delta)^{n-|x|}$ where $|x|$ is the number of 1's in x. Similarly $\mathbb{P}_2(x) = (1/2-\delta)^{|x|}(1/2+\delta)^{n-|x|}$. $\mathbb{P}_1(x) \geq \mathbb{P}_2(x)$ if and only if $|x| \geq n/2$. $\mathbb{P}_1(x)\chi(x) + \mathbb{P}_2(x)(1-\chi(x)) \geq \min\{\mathbb{P}_1(x), \mathbb{P}_2(x)\} = \mathbb{P}_2(x)$ if $|x| \geq n/2$. Therefore

$$\begin{aligned}\sum_{x\in\{0,1\}^n} \mathbb{P}_1(x)\chi(x) + \mathbb{P}_2(x)(1-\chi(x)) &\geq \sum_{x\in\{0,1\}^n:|x|\geq n/2} \mathbb{P}_2(x) \\ &= \mathbb{P}_2[|x| \geq n/2] = \Omega(1).\end{aligned}$$

□

Proof of Theorem 11. We consider two instances. In Instance 1, we have $(\rho_1, b_1) = (1, 1/2)$ and $(\rho_2, b_2) = (1/2+\delta, 1)$. In Instance 2, again we have $(\rho_1, b_1) = (1, 1/2)$ but now have $(\rho_2, b_2) = (1/2-\delta, 1)$. We set $\delta = T^{-1/3}$. We will show that the regret of any truthful mechanism is $\Omega(T^{2/3})$ for either of the two instances.

Because of Corollary 9, the number of non-competitive rounds is $o(T^{2/3})$ with probability $1 - o(1)$, else our 2-Regret would be $\Omega(T^{2/3})$. Hence it is enough to prove that given that the number of non-competitive rounds, say n, is $o(T^{2/3})$, the 2-Regret is $\Omega(T^{2/3})$.

Recall that 2-Regret $= OPT - \mathbb{E}[p_1 + p_2]$, and $p_i(\boldsymbol{b}) = b_i y_i(\boldsymbol{b}) - \int_{z=0}^{b_i} y_i(z, \boldsymbol{b}_{-i})dz$. We will show that $OPT - \mathbb{E}[y_1 b_1 + y_2 b_2] = \Omega(T^{2/3})$, which is enough since the integrals are positive[2]. Note that $\mathbb{E}_{c_2^t}[y_2^t | c_2^1 \ldots c_2^{t-1}] = \rho_2 x_2^t$. Thus $\mathbb{E}_{c_2^1 \ldots c_2^t}[y_2^t] = \rho_2 \mathbb{E}_{c_2^1 \ldots c_2^t}[x_2^t]$. Hence it is enough to argue that $OPT - \mathbb{E}_C[\rho_1 b_1 x_1 + \rho_2 b_2 x_2] = \Omega(T^{2/3})$. Call this latter quantity for a particular C, the *loss* for that C.

Claim 11. *For all click sequences C, if $x_1(1/2, 1, C) \geq T/2$ (resp. $x_1(1/2, 1, C) \leq T/2$) then the loss for C is at least $\delta T/2$ for Instance 1 (resp. Instance 2).*

Proof. Consider Instance 1 and suppose $x_1(1/2, 1, C) \geq T/2$. Then $x_2 \leq T/2$. Therefore $\rho_1 b_1 x_1 + \rho_2 b_2 x_2 = x_1/2 + (1/2+\delta)x_2 = 1/2(x_1 + x_2) + \delta x_2 \leq T/2 + \delta T/2 = (1/2+\delta)T - \delta T/2 = OPT - \delta T/2$. Hence loss for C is $\geq \delta T/2$. The case when $x_1(1/2, 1, C) \leq T/2$ is similar. □

Let χ be a function of C that is 1 if the loss for that click sequence is $\geq \delta T/2$ for Instance 1, and is 0 otherwise (loss is $\geq \delta T/2$ for Instance 2, as guaranteed by Claim 11). Since χ only depends on x_1, which in turn only depends on the clicks in the non-competitive rounds (Lemma 7), χ can be represented as a boolean decision tree of depth $n = o(1/\delta^2)$. From Lemma 10, either $\sum_{c\in\{0,1\}^n} \mathbb{P}_1(c)\chi(c) = \Omega(1)$ or $\sum_{c\in\{0,1\}^n} \mathbb{P}_2(c)(1-\chi(c)) = \Omega(1)$. If the former holds, this says that the probability that the loss is $\Omega(\delta T)$=$\Omega(T^{2/3})$ for Instance 1 is $\Omega(1)$. Thus the expected loss is $\Omega(T^{2/3})$. The other case implies an expected loss of $\Omega(T^{2/3})$ on Instance 2. □

[2] In fact, we can also show that the integrals themselves are $\Omega(T^{2/3})$, with a slightly different argument.

3.3 Lower bound for the stronger model

In this section, we prove a stronger version of the lower bound. The key differences in the setting we consider here are that

- the auction could charge at the end of all the rounds, and
- the bidders are only allowed to submit one bid, at the start.

Note that such auctions are more powerful and potentially have a lower regret. We show, however, that the regret is still $\Omega(T^{2/3})$.

We consider two instances. $\rho_1 = b_2 = 1$ in both the instances. Instance I1 is when $\rho_2 = 1/2 + \delta = b_1$ and Instance I2 is when $\rho_2 = 1/2 - \delta = b_1$, where $\delta = T^{-1/3}$. Also let $b^+ = 1/2 + \delta$ and $b^- = 1/2 - \delta$. We will show that the regret of any truthful mechanism is $\Omega(T^{2/3})$ for either of the two instances.

The loss: Recall that 2-$Regret = OPT - \mathbb{E}_C[p_1 + p_2]$, where $p_i(\boldsymbol{b}) = b_i y_i(\boldsymbol{b}) - \int_{z=0}^{b_i} y_i(z, \boldsymbol{b}_{-i})dz$. Note that[3] $\mathbb{E}_{c_2^t}[y_i^t | c_2^1 \dots c_2^{t-1}] = \rho_2 x_i^t$. Thus $\mathbb{E}_{c_2^1 \dots c_2^t}[y_i^t] = \rho_i \mathbb{E}_{c_2^1 \dots c_2^t}[x_i^t]$. Hence it is enough to argue that $OPT - \mathbb{E}_C[\rho_1 q_1 + \rho_2 q_2] = \Omega(T^{2/3})$, where $q_i(\boldsymbol{b}) = b_i x_i(\boldsymbol{b}) - \int_{z=0}^{b_i} x_i(z, \boldsymbol{b}_{-i})dz$. Let $loss(C) := OPT - (\rho_1 q_1 + \rho_2 q_2)$, keeping in mind that q_1 and q_2 are functions of C. Also say that C is *bad* for an instance if $loss(C)$ is $\Omega(T^{2/3})$. Note that for both I1 and I2, $OPT = \rho_1 b_1 T = \rho_2 b_2 T$. Therefore for any C, $loss(C)$ is non negative for both I1 and I2 because $\rho_1 q_1 + \rho_2 q_2 \leq \rho_1 b_1 x_1 + \rho_2 b_2 x_2 = \frac{OPT}{T}(x_1 + x_2) \leq OPT$.

Outline of the proof: We construct a boolean function of the click sequence C, $\chi : \{0,1\}^T \rightarrow \{0,1\}$ such that

- χ can be represented as a decision tree $\mathcal{T}$.
- If the depth of $\mathcal{T}$ on input C is $\Omega(T^{2/3})$ then C is bad for both I1 and I2.
- If $\chi(C) = 1$, then C is bad for I1, and if $\chi(C) = 0$, then C is bad for I2.

Given such a construction, it follows that the regret of any auction is $\Omega(T^{2/3})$ for either I1 or I2: Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be probability distributions on $\{0,1\}^T$ generated by i.i.d samples w.p. $1/2 + \delta$ and $1/2 - \delta$ respectively (as in Lemma 10). If the depth of $\mathcal{T}$ was $o(T^{2/3})$, then we could just apply Lemma 10. $\sum_{C \in \{0,1\}^T} \mathbb{P}_1(C)\chi(C)$ (resp. $\sum_{C \in \{0,1\}^T} \mathbb{P}_2(C)(1 - \chi(C))$) is the same as the probability that C is bad for I1 (and resp. for I2). If either of these probabilities is $\Omega(1)$, then so is the regret for that instance.

The problem is that for some inputs, the depth of $\mathcal{T}$ could be $\Omega(T^{2/3})$. Since all these inputs are bad for both instances, we may assume that both $\mathbb{P}_1$ and $\mathbb{P}_2$ of such instances is $o(1)$. We can apply Lemma 10 to the tree obtained by "pruning" $\mathcal{T}$ so that its depth is $o(T^{2/3})$, and the same conclusion would still hold.

Construction of $\mathcal{T}$: We construct $\mathcal{T}$ by looking at which clicks the auction observes when $b_1 = b^+(1 + \lambda)$ and $b_2 = 1$, where λ is a small constant > 0. We label the nodes of $\mathcal{T}$ by $t \in [T]$. (The same label may appear multiple times in $\mathcal{T}$, but occurs only once in any root to leaf path.) Let the root of $\mathcal{T}$ be the first t such that $x_2^t(b^+(1 + \lambda), 1) = 1$. Note that this has to be independent of C since the auction does not observe any clicks until it assigns an impression to advertiser 2. For each node t in the tree, the left (resp.

[3] We only need to consider c_2^t's, since in our instances, $\rho_1 = 1$.

right) child is the next time period t' such that $x_2^{t'}(b^+(1+\lambda), 1) = 1$, given that $c_2^t = 0$ (resp. $c_2^t = 1$). Again, as before, note that this is well defined, since t' only depends on the those c_2^t's that lie on the path from the root to t, as these are all the clicks that the auction observes.

For a given C, let $S = \{t : x_2^t(b^+(1+\lambda), 1) = 1\}$ be the set of nodes in the path from the root to the leaf the decision tree ends in, on input C. We now prove that *if* $|S| \geq \Omega(T^{2/3})$ *then both instances have* $loss(C) \geq \Omega(T^{2/3})$.

We introduce the following notation: $A_1[l, u](b_2) = \int_l^u x_1(z, b_2)dz$, and similarly A_2. Note that $loss(C) \geq \rho_1 A_1[0, b_1](b_2) + \rho_2 A_2[0, b_2](b_1)$.

Consider I1. If $|S| \geq \Omega(T^{2/3})$, then $x_2(b^+(1+\lambda), 1) \geq \Omega(T^{2/3})$. By scale invariance, $x_2(b^+, 1/(1+\lambda)) \geq \Omega(T^{2/3})$. Thus, $loss(C) \geq \rho_2 A_2[0, 1](b^+) \geq \rho_2 x_2(b^+, 1/(1+\lambda))(1 - 1/(1+\lambda)) \geq \Omega(T^{2/3})$. The proof for I2 is similar.

Construction of χ*:* We now construct the function χ such that *it depends only on* $(c_2^t : t \in S)$*, and if* $\chi(C) = 1$*, then* C *is bad for I1, and if* $\chi(C) = 0$*, then* C *is bad for I2.* We show this by arguing that if C and C' agree on S, then either they are both bad for I1, or they are both bad for I2. The existence of χ as required follows from this.

Now consider I1. $loss(C) \geq A_1[0, b^-](1) + A_1[b^-, b^+](1) + b^+ A_2[0, 1](b^+)$. By scale invariance, $\int_0^1 x_2(b^+, z)dz = \int_0^1 x_2(b^+/z, 1)dz$. By change of variables, $t = b^+/z$, it is equal to

$$\begin{aligned} \int_{b^+}^{\infty} b^+ x_2(t, 1)/t^2 dt &\geq \int_{b^+}^{b^+(1+\lambda)} b^+ x_2(t, 1)/t^2 dt \\ &\geq 1/(b^+(1+\lambda)^2) \int_{b^+}^{b^+(1+\lambda)} x_2(t, 1)dt \\ &= 1/(b^+(1+\lambda)^2) A_1^c[b^+, b^+(1+\lambda)](1), \end{aligned}$$

where $A_1^c[l, u](b_2)$ is defined to be $= \int_l^u x_2(t, b_2)dt$. Thus $loss \geq A_1[0, b^-](1) + A_1[b^-, b^+](1) + 1/(1+\lambda)^2 A_1^c[b^+, b^+(1+\lambda)](1)$.

Now consider I2. As before, $loss \geq A_1[0, b^-](1) + b^- A_2[0, 1](b^-)$. Again, as before, by scale invariance and change of variables,

$$\int_0^1 x_2(b^-, z)dz = b^-/(b^+(1+\lambda))^2 A_1^c[b^-, b^+(1+\lambda)].$$

The loss is therefore $\geq A_1[0, b^-](1) + (b^-/b^+(1+\lambda))^2 \left(A_1^c[b^-, b^+](1) + A_1^c[b^+, b^+(1+\lambda)](1)\right)$.

Thus, if either $A_1[0, b^-]$ or $A_1^c[b^+, b^+(1+\lambda)]$ is $\Omega(T^{2/3})$ then C is bad for both I1 and I2. Also, we argue later that $(A_1 + A_1^c)[b^-, b^+](1) \geq \Omega(T^{2/3})$, which implies that every C is bad for either I1 or I2,

The proof is completed by showing that if C and C' agree on S, then $|A_1[b^-, b^+](C) - A_1[b^-, b^+](C')| = o(T^{2/3})$, which then implies that either they are both bad for I1, or they are both bad for I2.

This is because of the following: C and C' agree on S, which is the set of clicks observed when $b_1 = b^+(1+\lambda)$ and $b_2 = 1$. Hence, $x_1(b^+(1+\lambda), 1)$ and $p_1(b^+(1+\lambda), 1)$ should be the same for both C and C'. This implies that $A_1[0, b^+(1+\lambda)] = b_1 x_1 - p_1$ is the same for both C and C'. We may also assume that $A_1[0, b^-]$ and $A_1^c[b^+, b^+(1+\lambda)]$ are $o(T^{2/3})$ which implies $|A_1[0, b^-](C) - A_1[0, b^-](C')| = o(T^{2/3})$, and $|A_1[b^+, b^+(1+\lambda)](C) - A_1[b^+, b^+(1+\lambda)](C')| = o(T^{2/3})$. The conclusion follows from the fact that if $x + y + z = x' + y' + z'$, then $|x - x'| \leq |y - y'| + |z - z'|$.

All that remains is to argue that $(A_1 + A_1^c)[b^-, b^+](1) \geq \Omega(T^{2/3})$. It is sufficient to argue that $(x_1 + x_2)(b^-, 1) \geq \Omega(T)$. Suppose not, then for both I1 and I2, $loss \geq OPT - \rho_1 b_1 x_1 - \rho_2 b_2 x_2 \geq b^- T - b^+(x_1 + x_2) \geq \Omega(T) - o(T) \geq \Omega(T)$.

Q.E.D

References

P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.*, 47(2-3):235–256, 2002. ISSN 0885-6125.

Rica Gonen and Elan Pavlov. An incentive-compatible multi-armed bandit mechanism. In *PODC '07: Proceedings of the twenty-sixth annual ACM symposium on Principles of distributed computing*, pages 362–363, 2007. ISBN 978-1-59593-616-5.

Jason Hartline and Anna Karlin. *Algorithmic Game Theory*, chapter Profit Maximization in Mechanism Design. Cambridge University Press, 2007.

S. Lahaie, D. Pennock, A. Saberi, and R. Vohra. *Algorithmic Game Theory*, chapter Sponsored Search. Cambridge University Press, 2007.

R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.

Hamid Nazerzadeh, Amin Saberi, and Rakesh Vohra. Dynamic cost-per-action mechanisms and applications to online advertising. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 179–188, 2008. ISBN 978-1-60558-085-2.

H. Robbins. Some aspects of the sequential design of experiments. In *Bulletin of the American Mathematical Society*, volume 55, 1952.