FAST RATE LEARNING IN STOCHASTIC FIRST PRICE BIDDING

joint work with Olivier Cappé and Aurélien Garivier

numberly
1000mercis group

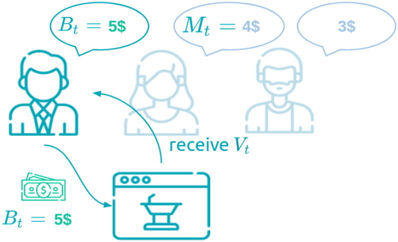
CONTEXT AND MODEL DEFINITION

Context: First price auctions have been largely adopted in the field of **programmatic advertising**, where they have progressively replaced second-price auctions.
Unlike second price auctions, first price auctions are noticeably **NOT truthful**.
There does not exist a close form for the optimal bid in first price auctions

Objective: designing an **online learning algorithm for bidding in repeated first price auctions**, in the case where the buyer plays against a *stationary stochastic environment*.

For $t = 1 \dots T$, the bidder of interest

1. Submits her bid B_t for the item of **unknown value** V_t. $\{V_t\}_{t \geq 1}$ i.i.d in $[0, 1]$ of expectation v ;
2. Observes the maximum of the other bids : M_t. $\{M_t\}_{t \geq 1}$ is i.i.d in $[0, 1]$ (with cdf F);
3. If $M_t \leq B_t$, she **observes and receives** V_t, and **pays** B_t. Otherwise, she loses the auction and does not observe V_t.



Utility The utility is $U_t(b) = \mathbb{E}[(V_t - b)\mathbb{1}\{b \geq M_t\}] = (v - b)F(b)$.

Regret

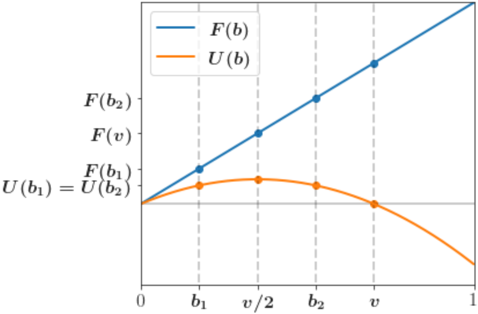
$$R_T := \max_{b \in [0,1]} \sum_{t=1}^{T} U_t(b) - \sum_{t=1}^{T} \mathbb{E}[U_t(B_t)].$$

We study two different settings:

- **Known** F. F is known in advance. This setting is close to the second-price setting, since v is the only parameter to be estimated. However, the utility function can be far more complex in first price auctions.
- **Unknown** F. F is unknown, and needs to be estimated. This setting bears similarities with the posted price one.

Exploration/Exploitation Trade-Off where Exploitation consists in bidding close to a certain b^* (not necessarily unique) and Exploration consists in bidding high enough (bidding 1 means observing everything).

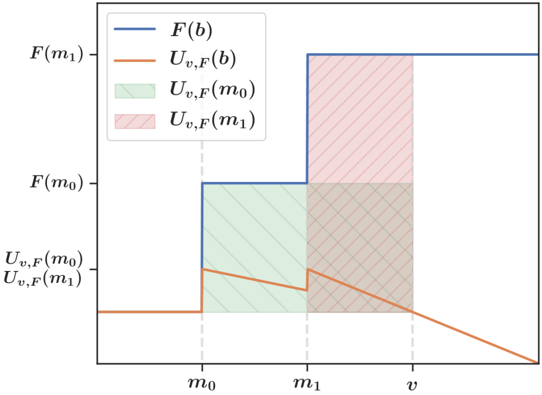
Example:
M_t uniform,
$V_t \sim$ Bernoulli(v).
$U_{v,F}(b) =$
$\frac{1}{4}v^2 - (v/2 - b)^2$.



PROPERTIES OF FIRST PRICE AUCTIONS

Unlike in second price auctions, the maximizer of the utility is **not available in close form**. More generally, there could be **multiple maximizers**, or arbitrarily close maxima. Thus, we define

$$b^*_{v,F} = \max\left\{ \arg\max_{b\in[0,1]} U_{v,f}(b) \right\}.$$



Theorem

Let $\mathcal{C}$ denote the class of cumulative distribution functions on $[0, 1]$. *Any strategy, whether it assumes knowledge of F or not, must satisfy*

$$\liminf_{T\to\infty} \frac{\max_{v\in[0,1], F\in\mathcal{C}} R_T^{v,F}}{\sqrt{T}} \geq \frac{1}{64}.$$

Theorem

Under general regularity assumptions on F (see paper):

- *there exists one unique maximizer $b^*_{v,F}$ of the utility*
- $\psi_F : v \mapsto b^*_{v,F}$ *is Lipschitz continuous with a Lipschitz constant 1.*
- *there exist two constants c and C such that* $\forall b \in [b^*_{v,F} - \Delta, b^*_{v,F} + \Delta]$,

$$c(b^*_{v,F} - b)^2 \leq U_{v,F}(b^*_{v,F}) - U_{v,F}(b) \leq C(b^*_{v,F} - b)^2$$

- $F(b^*_{v,F})$ *can not be arbitrarily small*

These assumptions include large classes of distributions (like the majority of Beta distributions)

KNOWN BID DISTRIBUTION

Estimation method:

We estimate $U_{v,F}$ thanks to the average $\hat{V}_t$

$$\hat{V}_t := \frac{1}{t-1} \sum_{s=1}^{t-1} \mathbb{1}\{M_s < b\} V_s.$$

Algorithm (UCBid1)

Initially set $B_1 = 1$ *and, for* $t \geq 2$, *bid according to*

$$B_t = \max \left\{ \underset{b \in [0,1]}{\arg\max} (\hat{V}_t + \epsilon_t - b) F(b) \right\}.$$

where $\epsilon_t := \sqrt{\gamma \log(t-1)/2N_t}$.

Theorem

When $\gamma > 1$, the regret of UCBid1 is upper-bounded as

$$R_T^{v,F} \leq \frac{\sqrt{2\gamma}}{F(b^*_{v,F})} \sqrt{\log T} \sqrt{T} + O(\log T) \ .$$

While if F is regular then

$$R_T^{v,F} \leq \frac{\gamma \lambda \bar{\beta}^2}{F(b^*_{v,F}) \beta} \log^2(T) + O(\log T).$$

where β and $\bar{\beta}$ are constants depending only on F.

(+ parametric lower bound confirms that you can not do much better when being optimistic)

UNKNOWN BID DISTRIBUTION

Estimation method:

We estimate $U_{v,F}$ thanks to $\hat{V}_t$ and to the empirical c.d.f.

$$\hat{F}_t(b) := \frac{1}{t-1}\sum_{s=1}^{t-1} \mathbb{1}\{M_s < b\}.$$

Intuition: We do not add any optimistic bonus to the estimate $\hat{F}_t$: it is not necessary to be optimistic about F since the observation M_t drawn according to F is observed at each time step whatever the bid submitted.

Algorithm (UCBid1+)

Submit a bid equal to 1 *in the first round, then bid:*

$$B_t = \max\{\arg\max_{b\in[0,1]}(\bar{V}_t + \epsilon_t - b)\hat{F}_t(b)\},$$

Theorem

In all generality, when $\gamma > 2$

$$R_T^{v,F} \leq C_{v,F}\sqrt{\frac{\gamma v}{U_{v,F}(b^*_{v,F})}}\sqrt{T\log T} + O(\log T),$$

While in the regular case

$$R_T^{v,F} \leq O\left(T^{1/3+\epsilon}\right),$$

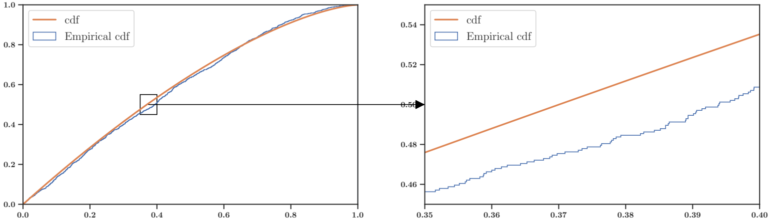
for any $\epsilon > 0$.

Lemma

Local concentration inequality For any $a, b \in [0,1]$, if F is increasing,

$$\sup_{a \le x \le b} |\hat{F}_t(x) - F(x) - (\hat{F}_t(a) - F(a))| \le \sqrt{\frac{2(F(b)-F(a)) \log\left(\frac{e\sqrt{t}}{\eta\sqrt{2(F(b)-F(a))}}\right)}{t}} + \frac{\log(\frac{t}{2(F(b)-F(a)\eta^2})}{6t},$$

with probability $1-\eta$.



EXPERIMENTS

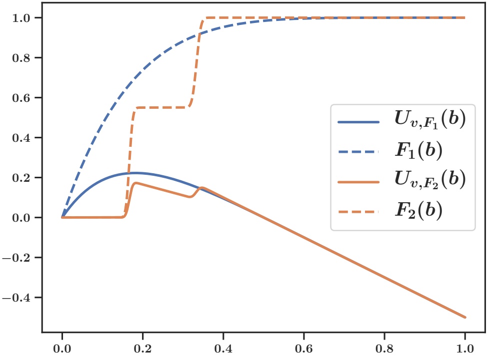
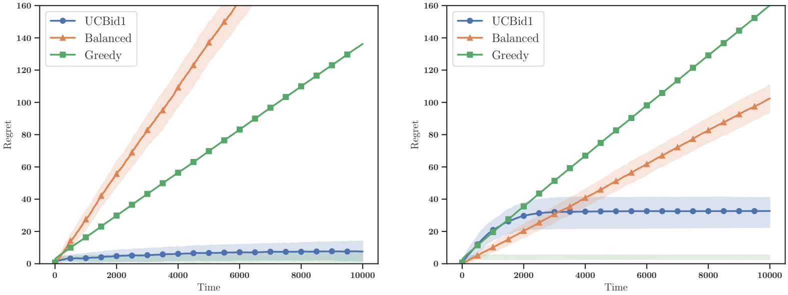
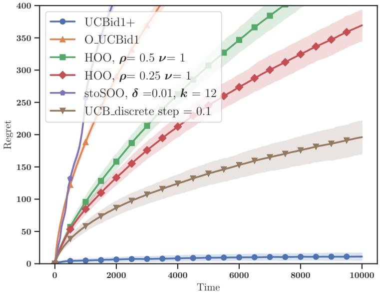


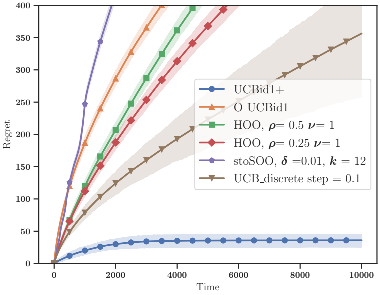
Figure: Two choices of F; associated utilities for $v = 1/2$.

Known Bid Distribution, 2 instances

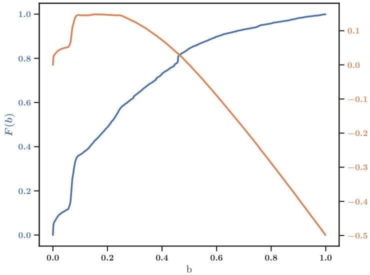


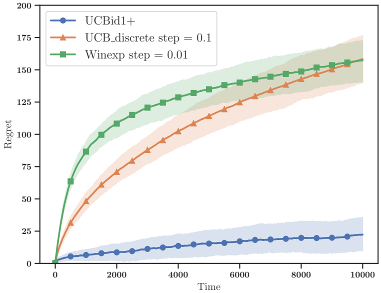
Unknown Bid Distribution, instance 2





Data from one advertising campaign





ADVERLINE xandr numberly 1000mercis group

Thanks