Fast Rate Learning in Stochastic First Price Bidding

Juliette Achddou [1] Olivier Cappé [2] Aurélien Garivier [3]

[1]DIENS, PSL, INRIA, 1000mercis Group, [2]DIENS, PSL, INRIA, CNRS [3]UMPA,CNRS, INRIA, ENS Lyon

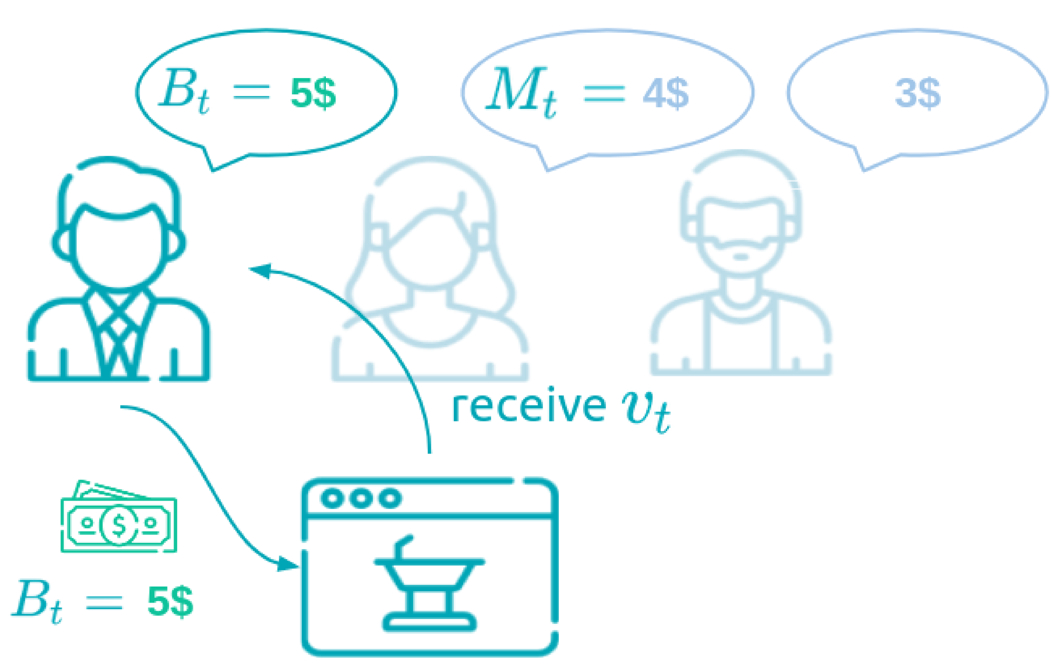
Objectives

Context: First price auctions have been largely adopted in the field of **programmatic advertising**, where they have progressively replaced second-price auctions.
Objective: designing an **online learning algorithm for bidding in first price auctions**, in the case where the buyer plays against a *stationary stochastic environment*.

Model

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 $[0,1]$;
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.



The (pseudo-) regret is defined by

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

where the utility is

$$U_{v,F}(b) := \mathbb{E}\big[(V_t - b)\mathbb{1}\{M_t \leq b\}\big] = (v-b)F(b).$$

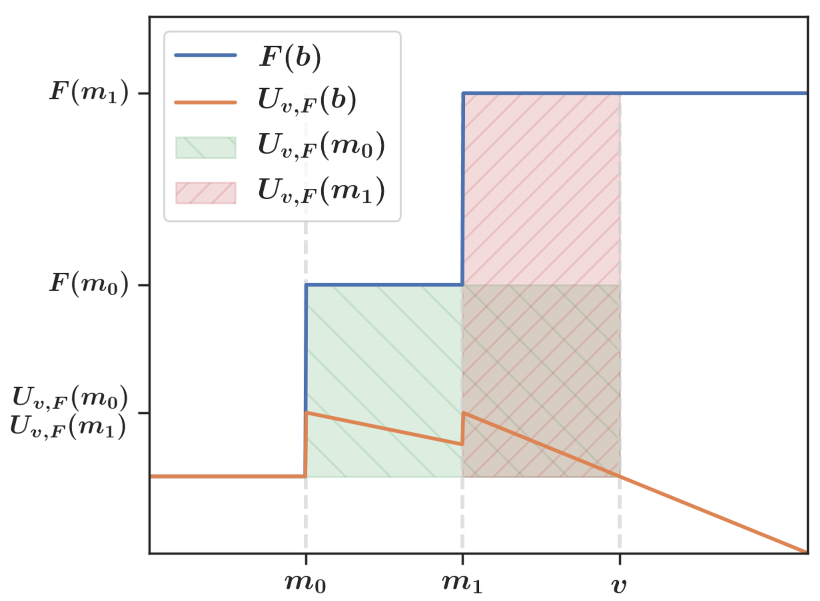
Some Intuition

Exploration/Exploitation Trade-Off where Exploitation consists in bidding close to a certain optimum and Exploitation consists in bidding high enough (bidding 1 means observing everything).

Properties of First Price Auctions

Difficulties: 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

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



General Lower Bound

Any strategy, whether it assumes knowledge of F or not, must satisfy

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

Regular Case. 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 $\forall b \in [b^*_{v,F} - \Delta, b^*_{v,F} + \Delta]$,
  $c(b^*_{v,F} - b)^2 \leq U_{v,F}(b^*_{v,F}) - U_{v,F}(b) \leq C(b^*_{v,F} - b)^2$

Estimation method: We estimate $U_{v,F}$ thanks to the average of V_t and to the empirical c.d.f.

$$\hat{V}_t := \frac{1}{N_t}\sum_{s=1}^{t-1} V_s \mathbb{1}\{M_s \leq B_s\},$$

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

The Algorithm and its Analysis

UCBid1+ Algorithm

Initially set $B_1 = 1$;
for $t \geq 2$, **do**

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

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

Regret Upper Bound

In all generality, when $\gamma > 2$

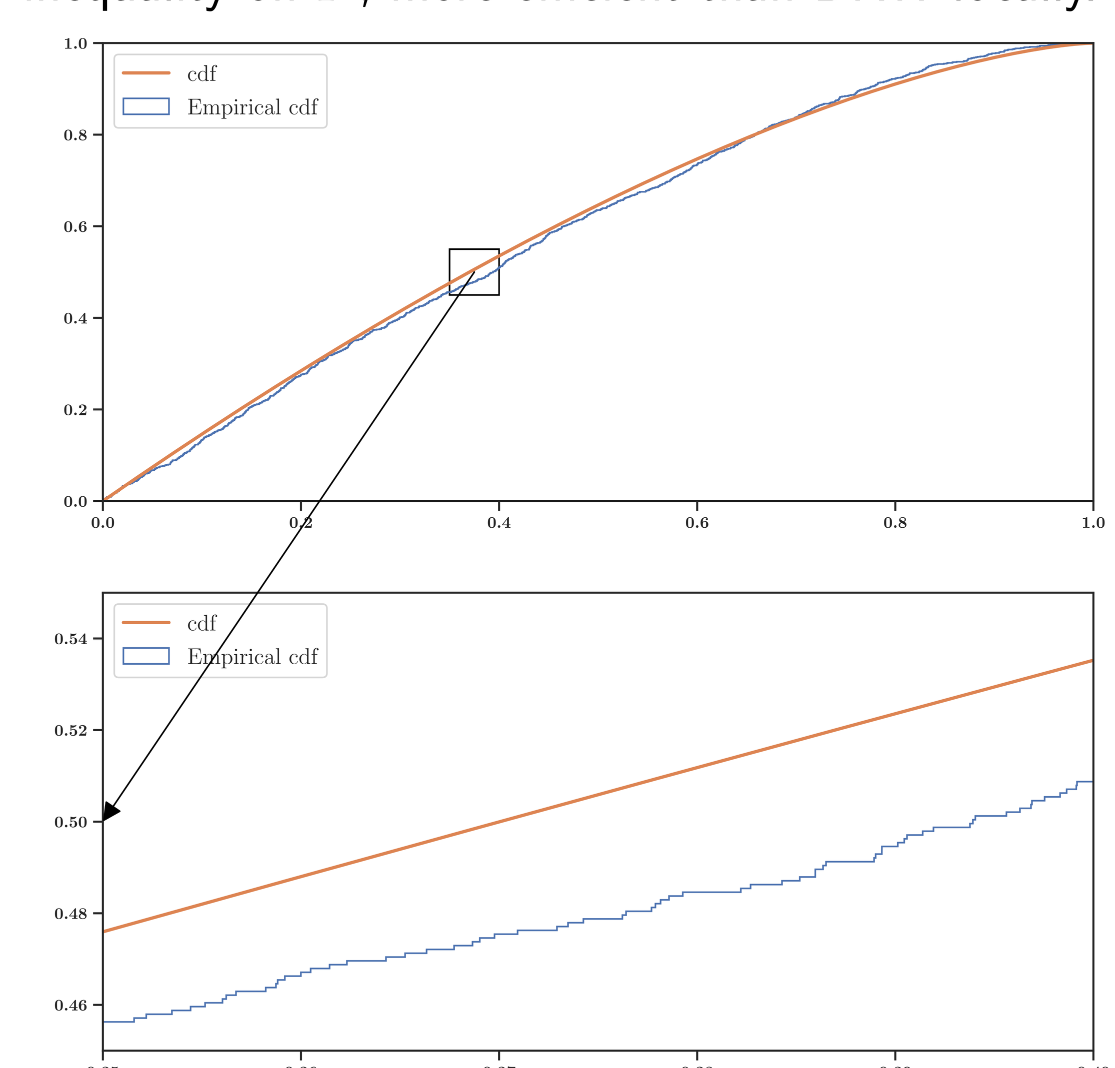
$$R_T^{v,F} \leq C_{v,F}\sqrt{\frac{\gamma v}{U_{v,F}(b^*_{v,F})}}\sqrt{T\log T} + O(\log T),$$

While in the regular case

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

for any $\epsilon > 0$

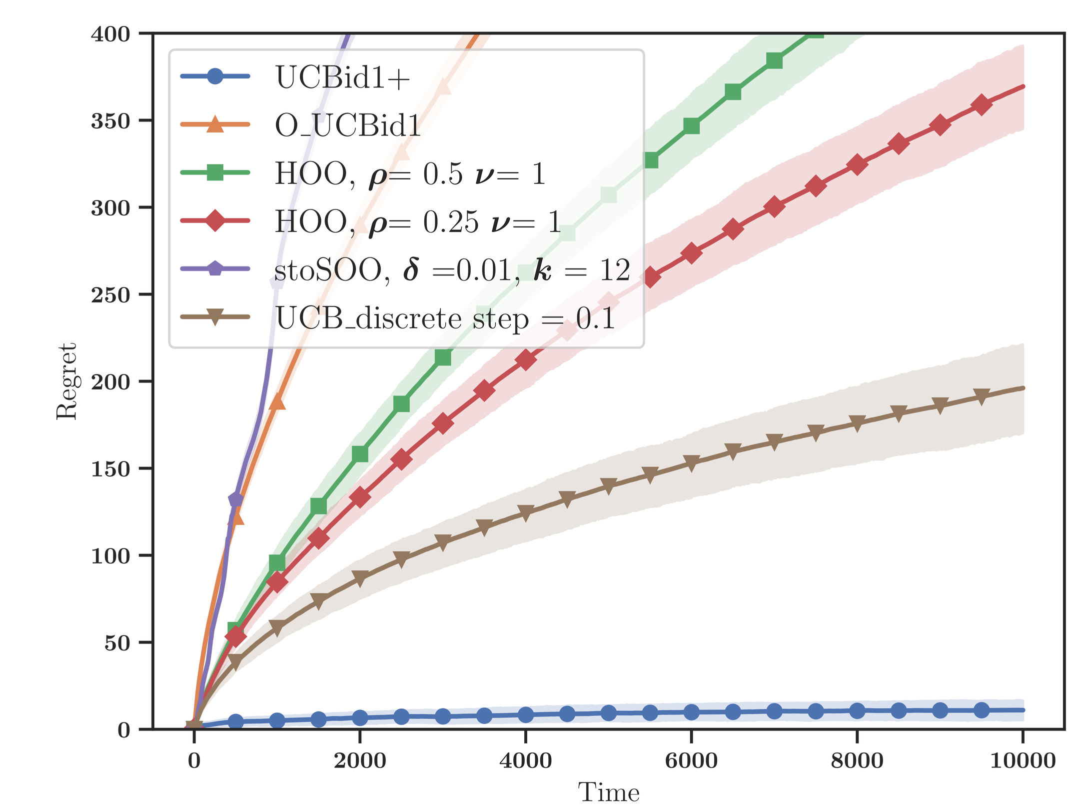
The proof relies on a new local concentration inequality on F, more efficient than DKW locally.



Remark: We also study the setting in which one knows F in advance. Then one can directly resort to F instead of $\hat{F}$, which leads to a significantly reduced regret ($O(\log^2 T)$) in the regular case.

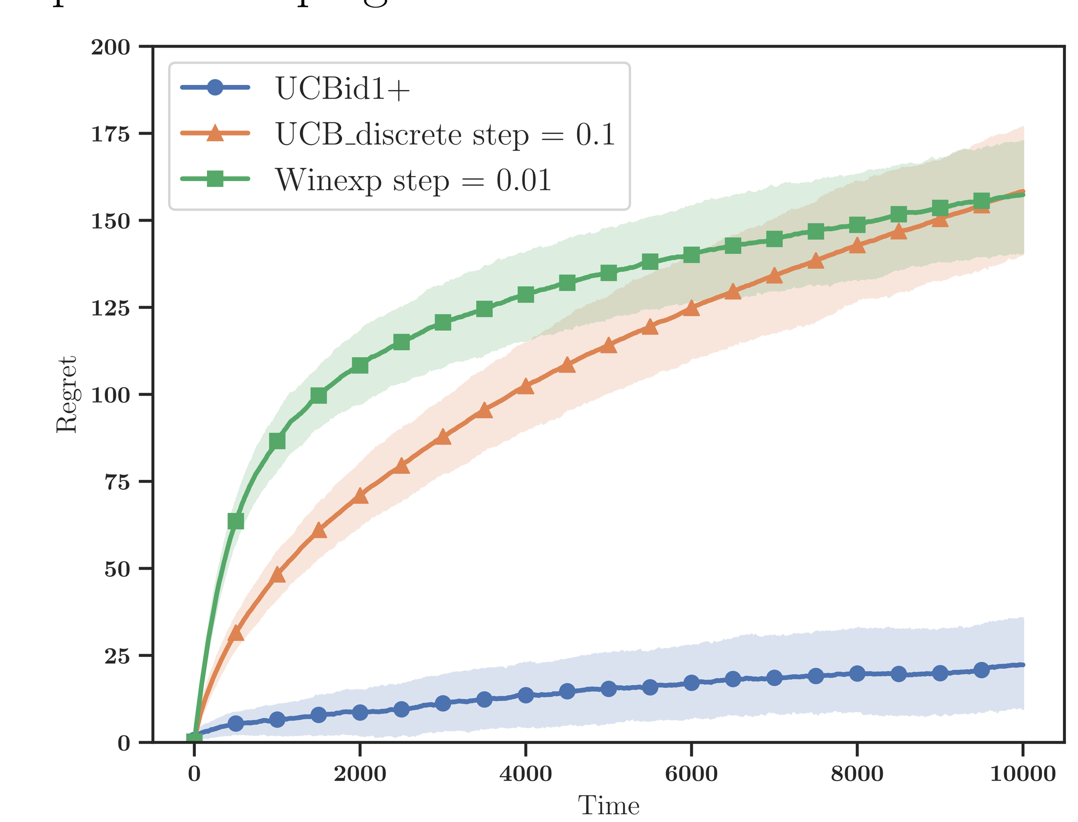
Simulations

With $M_t \sim Beta(1,6)$ and $v = 1/2$



Real Data Experiments

Real-world bidding dataset: 56607 bids that were made on a specific placement on Adverline's inventory on auctions that Numberly participated to, for a specific campaign.



Acknowledgements

We would like to thank Adverline for accepting to provide us with the bidding data on their inventories and Xandr for making this data transaction possible. Aurélien Garivier acknowledges the support of the Project IDEXLYON of the University of Lyon, in the framework of the Programme Investissements d'Avenir (ANR-16-IDEX-0005), and Chaire SeqALO (ANR-20-CHIA-0020).

ADVERLINE xandr numberly