Logarithmic Regret in Feature-based Dynamic Pricing

Jianyu Xu and Yu-Xiang Wang
University of California, Santa Barbara

Outline

- Problem Setup
- Summary of Results
  - EMLP algorithm
  - ONSP algorithm
  - Numerical Results
  - Lower Bounds
- Conclusions
  - Open Problem

Dynamic Pricing

Single-product Pricing

Feature-based Pricing

Problem Setting

- Online-fashion sales with a *linear-noisy valuation* model:

For $t = 1,2,\dots,T$:
- Feature $x_t \in \mathbb{R}^d$ is revealed;
- Customer generate a valuation $y_t = x_t^\top \theta^* + N_t$ secretly (with a fixed θ^*);
- Seller (we) propose a price v_t;
- We get a reward $r_t = v_t \cdot 1_t$ where $1_t = 1[v_t \le y_t]$ is customer's decision.

- Noise distribution is *known* to us.
- Comparing with *bandits* feedback:
  - Boolean-censored
  - Half-space information

Performance Metric: Regret

In this setting, a *regret* is defined as:

$$\sum_{t=1}^{n} \max_{v_t^*} \mathbb{E}_{N_t \sim \mathbb{D}}[v_t^* \cdot 1(v_t^* \leq x_t^\top \theta^* + N_t) | \theta^*] - \sum_{t=1}^{n} \mathbb{E}_{N_t \sim \mathbb{D}}[v_t \cdot 1(v_t \leq x_t^\top \theta^* + N_t)]$$

Max expected reward of a seller knowing θ^* in advance.

Expected reward of our algorithm.

Our Contribution

- We achieve the optimal $O(d \log T)$ regret.
  - in both **stochastic** and **adversarial** settings.

x_t's are drawn from **any** independent and identical **distributions**

x_t's are selected by an oblivious **adversary** prior to all sales starting

- Why a logarithmic regret?
  - Key: knowing noise distribution $\mathbb{D}$.
  - We also prove an $\Omega(\sqrt{T})$ lower bound for now knowing $\mathbb{D}$.

Results on Linear Feature-based Pricing

Regret	Noise-free	Distribution-known Noise		Distribution-unknown Noise	
		Stochastic features	Adversarial features	Parametric log-concave family	Agnostic
Upper Bound	$O(d \log\log T)$ [PLS 18]	$O\left(\min\left\{\frac{d \log T}{\lambda_{min}^2}, d\sqrt{T}\right\}\right)$ [JN19] $O(d \log T)$ [This paper]	$O\left(d^{\frac{1}{3}}T^{\frac{2}{3}}\right)$ [CLPL16] $O(d \log T)$ [This paper]	$O\left(d\sqrt{T}\right)$ [JN19] $O\left(poly(\log T)\right)$ [CLPL16, KLPS20 (for $\sigma = O\left(\frac{1}{T}\right)$ noise)]	Open
Lower Bound	$\Omega(d \log\log T)$ [KL03]	$\Omega(d \log T)$ [JN19]		$\Omega\left(\sqrt{T}\right)$ [JN19, This paper (for Gaussian noise)]	$\Omega\left(T^{\frac{2}{3}}\right)$ [KL03]

EMLP: Epoch-based Max-Likelihood Pricing

- Key to solve this pricing problem: **learning $\boldsymbol{\theta}^*$.**
  - Since noise distribution $\mathbb{D}$ is known.
- How to learn θ^* while exploiting $\hat{\theta}$?
  - A *Max Likelihood Estimator* (MLE)
  - A *doubling-epoch* design

For $k = 1,2,\dots,\log T$:

Epoch k. For $t = 1,2,\dots,\tau_k = 2^{k-1}$:

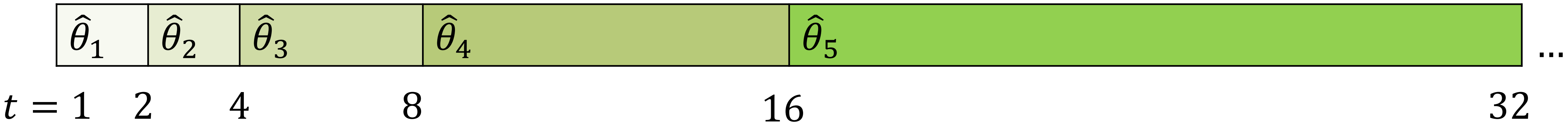
Set $v_t = \mathrm{argmax}_v\, v \cdot \left(1 - F\left(v - x_t^\top \hat{\theta}_k\right)\right)$;

Observe feedback 1_t;

Construct negative log-likelihood $l_t(\theta)$:

$$l_t(\theta) := -\mathbb{1}_t \cdot \log\left(1 - F(v_t - x_t^\top \theta)\right) - (1 - \mathbb{1}_t)\log\left(F(v_t - x_t^\top \theta)\right)$$

MLE: $\hat{\theta}_{k+1} = \mathrm{argmin}_\theta\, \hat{L}_k(\theta)$, where $\hat{L}_k(\theta) = \frac{1}{\tau_k}\sum_{t=1}^{\tau_k} l_t(\theta)$;



EMLP: Regret Analysis

Regret $\leq$ $\left(x^{\top}(\theta^* - \hat{\theta})\right)^2$ $\leq$ $\mathbb{E}[\hat{L}(\hat{\theta}) - \hat{L}(\theta^*)]$ $\leq$ $d/(\tau_k + 1)$

Smoothness

Strong Convexity w.r.t. xx^T

Fast rate of a Probit Model

summation

$O(d \log T)$

ONSP: Online-Newton-Step Pricing

- Notice:
  $Regret_t \leq \mathbb{E}\left[l_t(\hat{\theta}) - l_t(\theta^*)|x_t\right]$.
  - Why not directly reduce $\hat{L}(\hat{\theta})$?
  - Without x_t's concentration
    -- allowing **adversarial** sequences.
- Exp-concave $\hat{L}$ $\Rightarrow$ fast rate:
  Online Newton Step (ONS).
  - Suitable for (oblivious) adversarial objective functions.
  - Update θ_t at each round.

For $t = 1,2,\ldots,T$:

Set $v_t = \text{argmax}_v\, v \cdot \left(1 - F(v - x_t^\top \theta_t)\right)$;

Observe feedback 1_t;

Construct $l_t(\theta)$ and $\nabla_t := \nabla l_t(\theta_t)$;

Online Newton Step:

- $A_t = A_{t-1} + \nabla_t \nabla_t^\top$; $(A_0 = I_d)$
- $\hat{\theta}_{t+1} = \theta_t - \frac{1}{\gamma} A_t^{-1} \nabla_t$
- Projection: $\theta_{t+1} = \Pi^{A_t}(\hat{\theta}_{t+1})$

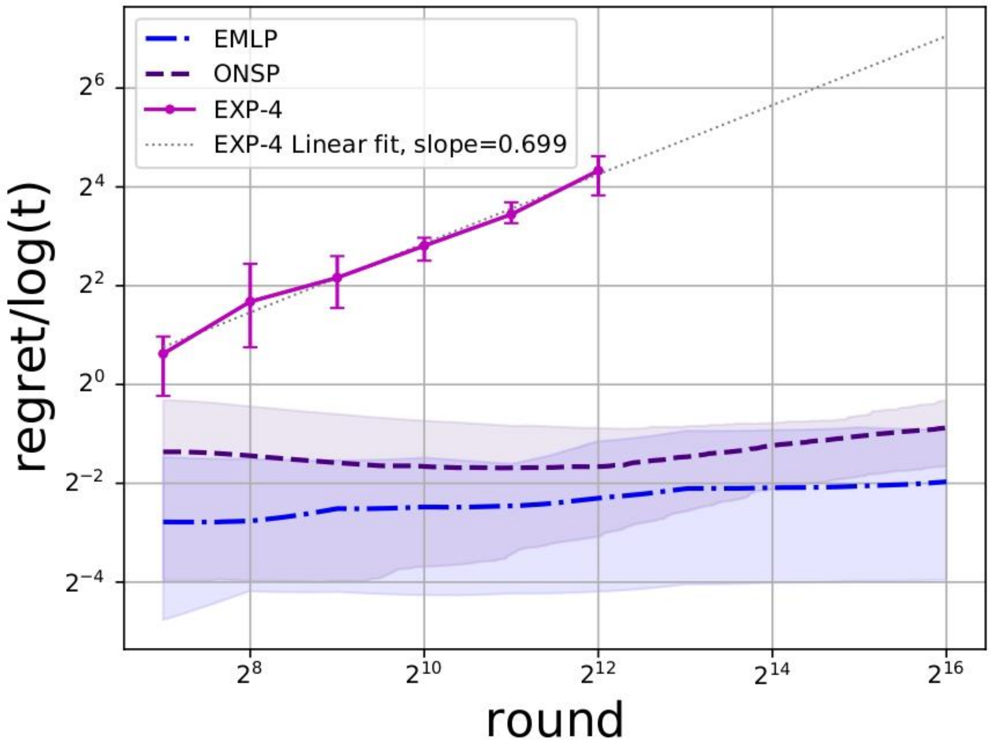
ONSP: Regret Analysis

Regret $\leq$ $\left(x^\top(\theta^* - \hat{\theta})\right)^2$ $\leq$ $\mathbb{E}[\hat{L}(\hat{\theta}) - \hat{L}(\theta^*)]$ $\leq$ (summation) $O(d \log T)$

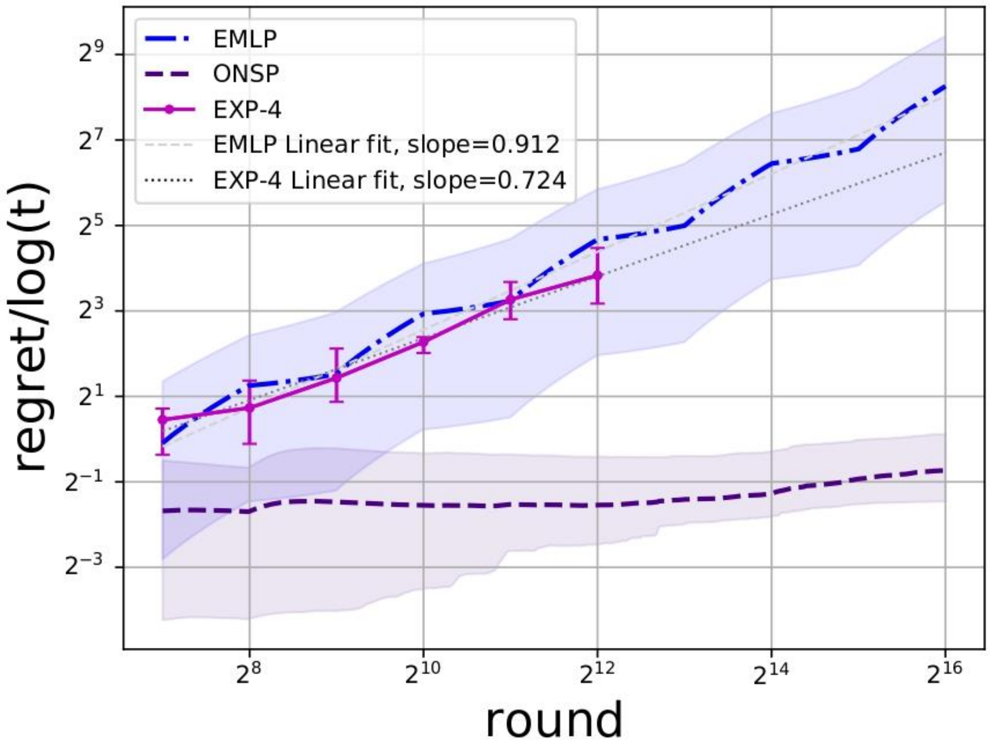
- Smoothness
- Strong Convexity w.r.t. xx^T
- Strong Convexity & Online Newton Step

Numerical Result

Stochastic x_t's



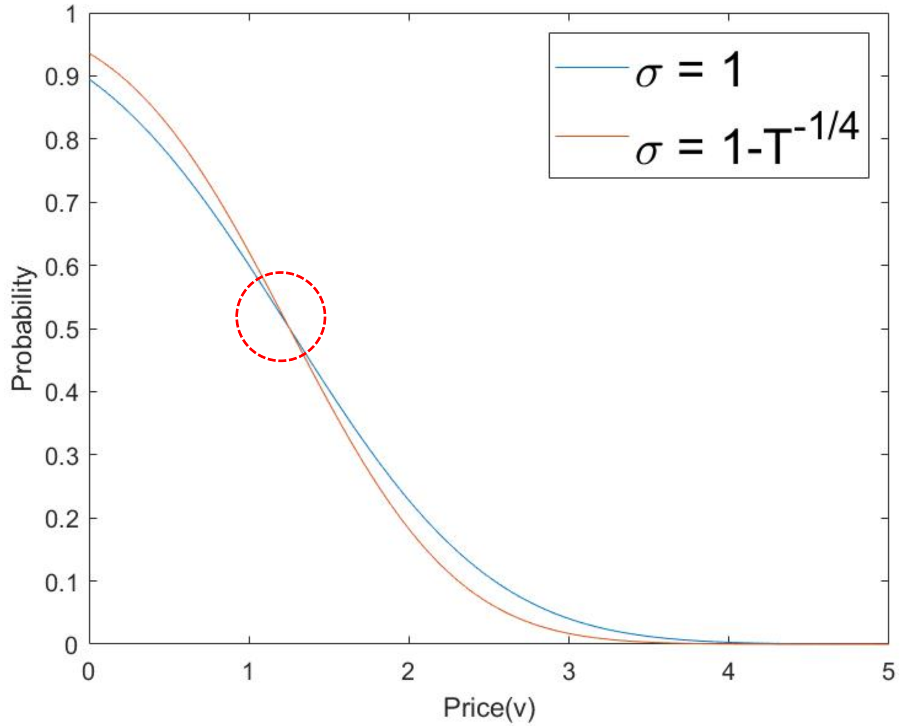
Adversarial x_t's



$\Omega(\sqrt{T})$ lower bound for $\mathcal{N}(0,\sigma)$ noise with unknown σ

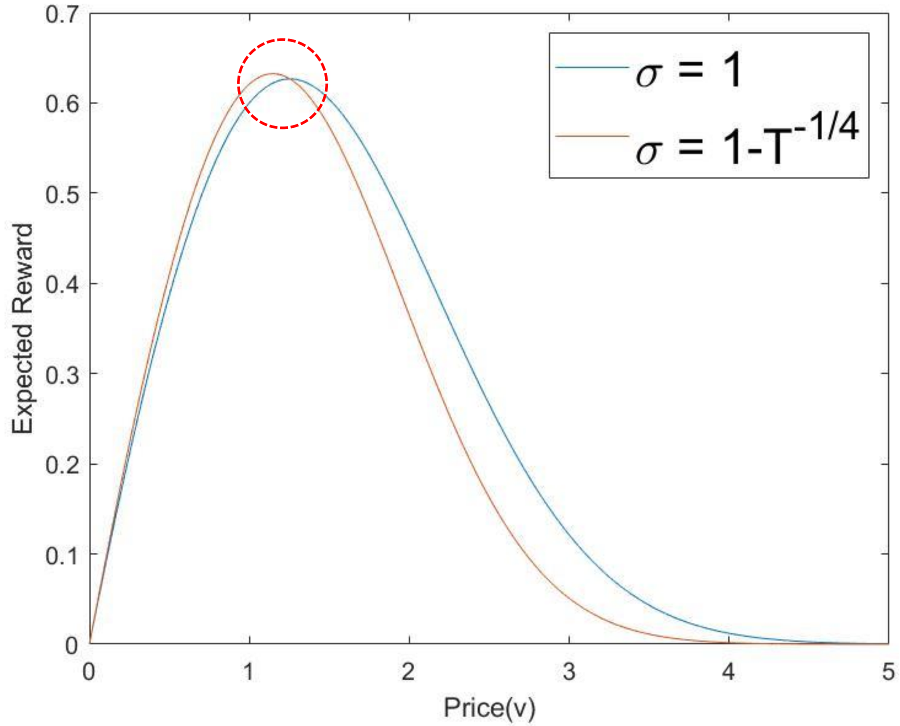
Expected Boolean feedback:

Hard to distinguish



Expected reward:

Have to distinguish

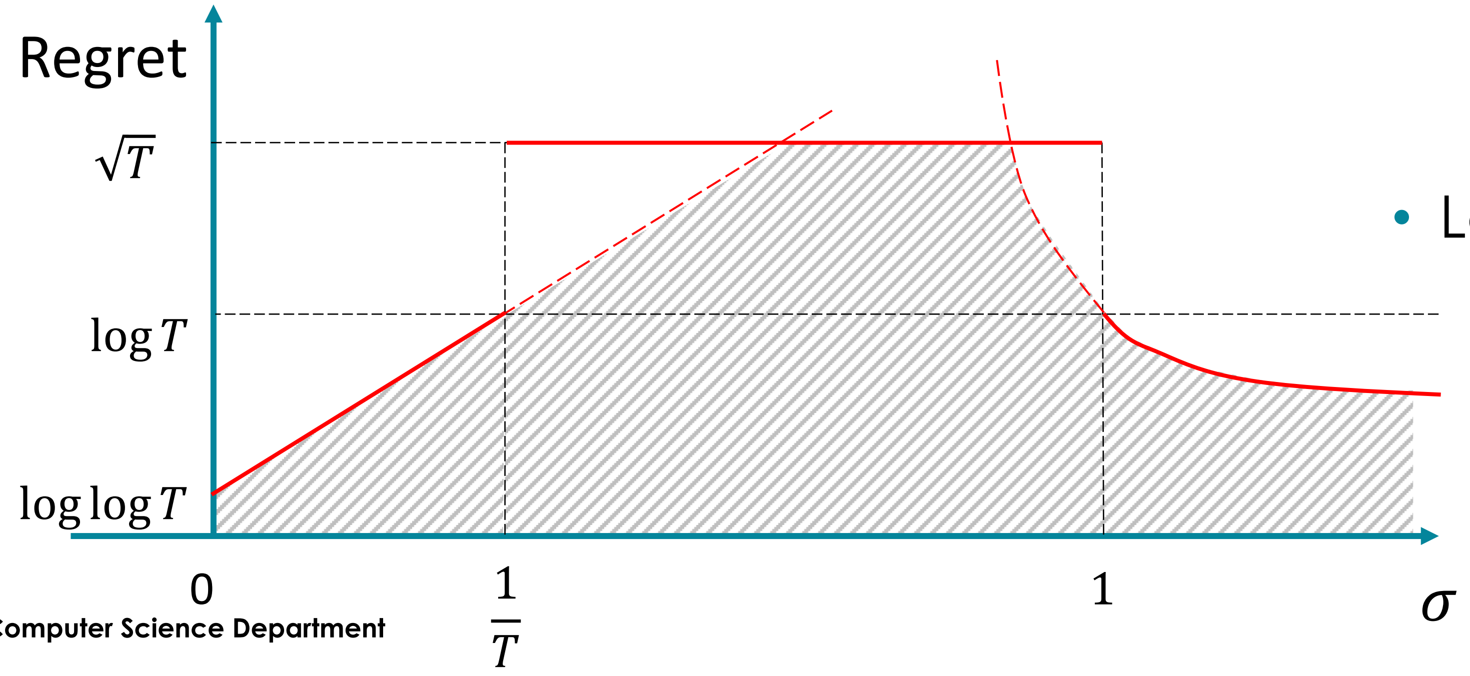


Conclusion

- Achieve an $O(d\log T)$ regret for stochastic/adversarial x_t's. Comparing with existing results, we:
  - Get rid of distributional assumptions.
  - Exponentially improve the regret bound.
- Pricing is exponentially easier than contextual bandits.
  - As long as we ***know*** the noise distribution $\mathbb{D}$.
  - If not, then the regret is still $\Omega(\sqrt{T})$.

An Open Problem: Regrets for different σ

Noise σ	Regret	Trend
0	$O(d \log\log T)$	/
$\tilde{O}(1/T)$	$O(d^2 \log T)$	Increasing w.r.t. σ.
$\big(\Omega(1/T), O(1)\big)$	$O(\sqrt{T})$	(Not matching.)
$\Theta(1)$	$O(d \log T)$	Decreasing w.r.t. σ.



- All existing algorithms for $\sigma = \Theta(1)$ suffer a "higher variance lower regret" phenomenon.
  - Very counter-intuitive.
  - Is this a necessity?
- No log-regret algorithm for "not-very-small" variance noise yet.
  - What is the minimax regret?
- Look forward to a unified algorithm!

UC SANTA BARBARA