10/16/2022

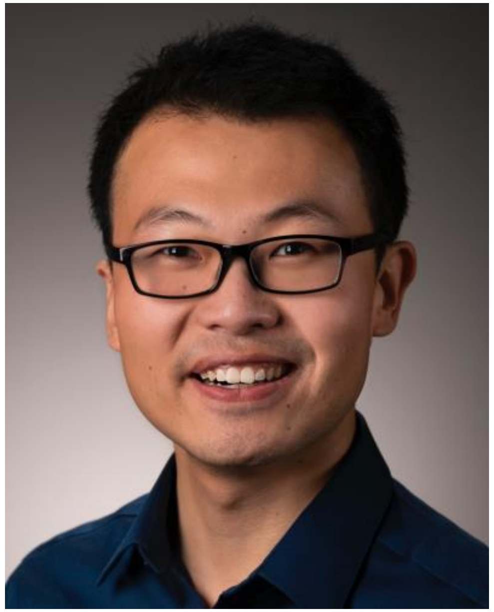
Linear Contextual Dynamic Pricing

Jianyu Xu

University of California, Santa Barbara

Outline

- Problem Settings
- A Linear Valuation (LV) Problem
- A Linear Policy (LP) Problem
- Open Problem



These are joint works with Prof. Yu-Xiang Wang

References:

[XW21] Xu, Jianyu, and Yu-Xiang Wang. "Logarithmic regret in feature-based dynamic pricing." *in NeurIPS 2021* [spotlight presentation].

[XW22] Xu, Jianyu, and Yu-Xiang Wang. "Towards Agnostic Feature-based Dynamic Pricing: Linear Policies vs Linear Valuation with Unknown Noise.“ *in AISTATS 2022* [oral presentation].

Problem Setting

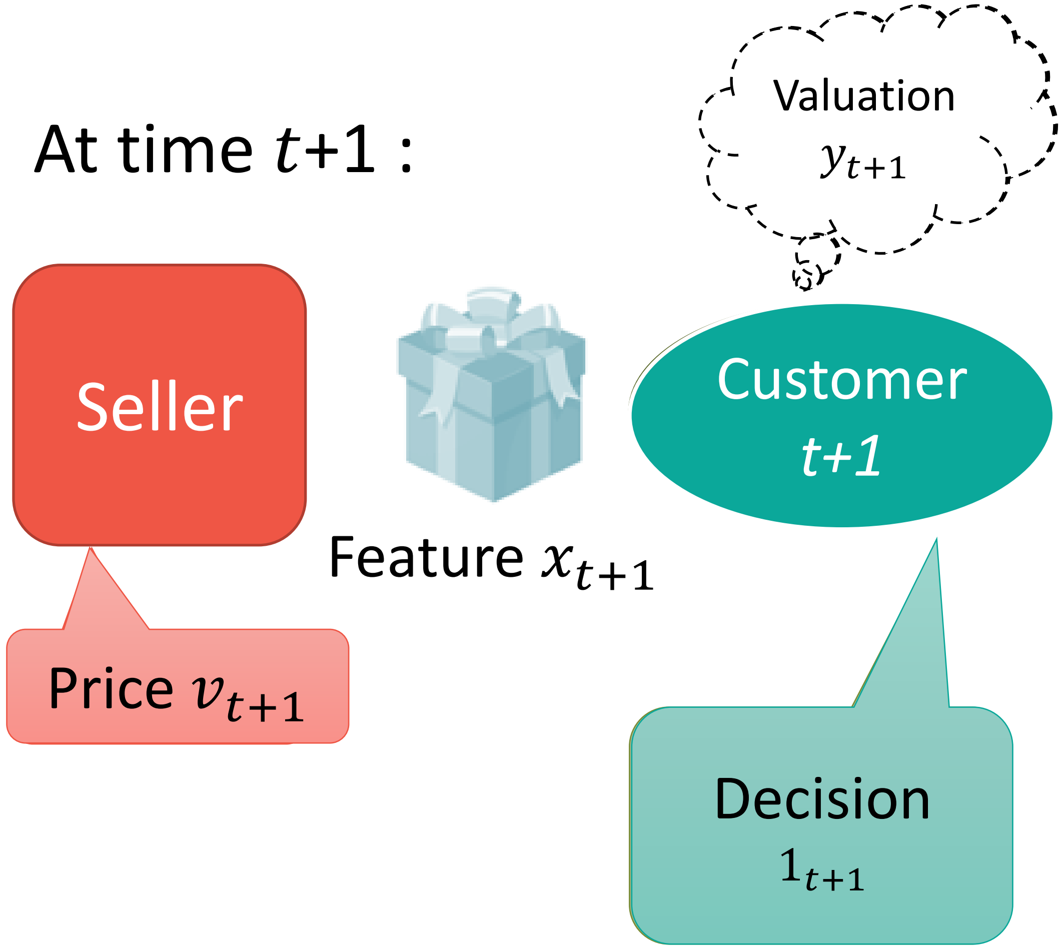
- Online-fashion contextual pricing:

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

- Feature $x_t \in \mathbb{R}^d$ is revealed;
- A customer generates a valuation y_t secretly;
- The seller (**we**) proposes a price v_t;
- Customer makes a decision $1_t = 1[v_t \leq y_t]$;
- We receive a reward $r_t = v_t \cdot 1_t$.

- We never observe customer's valuation!

At time t+1 :



Compare with Contextual Bandits

Contextual bandits:

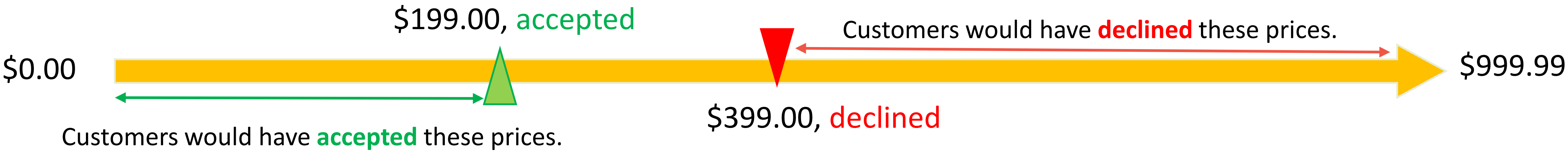
Input: Action set $\mathcal{A}$

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

- Context $x_t \in \mathbb{R}^d$ is revealed;
- Agent (we) takes an action $a_t \in \mathcal{A}$;
- Environment returns a reward r_t only on a_t.

In Comparison, pricing has …

- Similarities:
  - Online-learning process
  - Partial-information feedback
  - Interactive decisions
- Differences:
  - Infinite Actions & Non-continuous Reward
  - Half-space feedback



Performance Metric: Regret

A *regret* is defined as:

$$\sum_{t=1}^{n} \max_{v_t^*} \mathbb{E}[v_t^* \cdot 1(v_t^* \leq y_t) | x_t] - \sum_{t=1}^{n} \mathbb{E}[v_t \cdot 1(v_t \leq y_t) | x_t]$$

Max expected reward of a seller knowing optimal price in advance.

Expected reward of our algorithm.

Larger regret → Worse performance ！

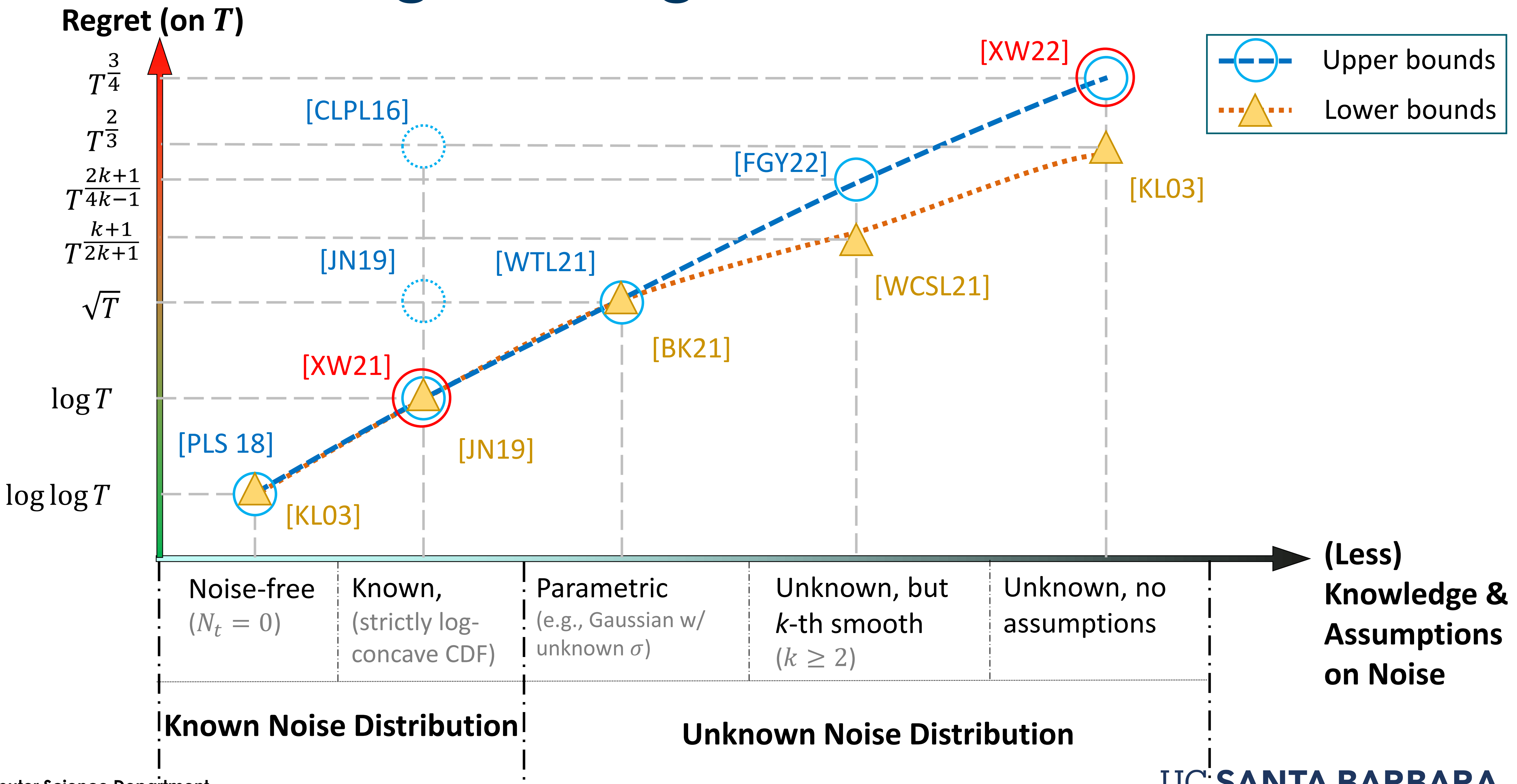
Linear Valuation (LV)

- Customer's valuation y_t is …
  - **Linear** (on feature x_t), and
  - **Noisy** (added as N_t)

$$y_t = x_t^\top \theta^* + N_t$$

- We assume …
  - θ^* is **fixed** and **unknown**;
  - N_t's are **i.i.d.**

LV: Knowledge vs Regret



LV: $O(d \log T)$ Regrets for *Known* Noise Distributions [XW21]

- 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** (i.i.d)

Algorithm: EMLP

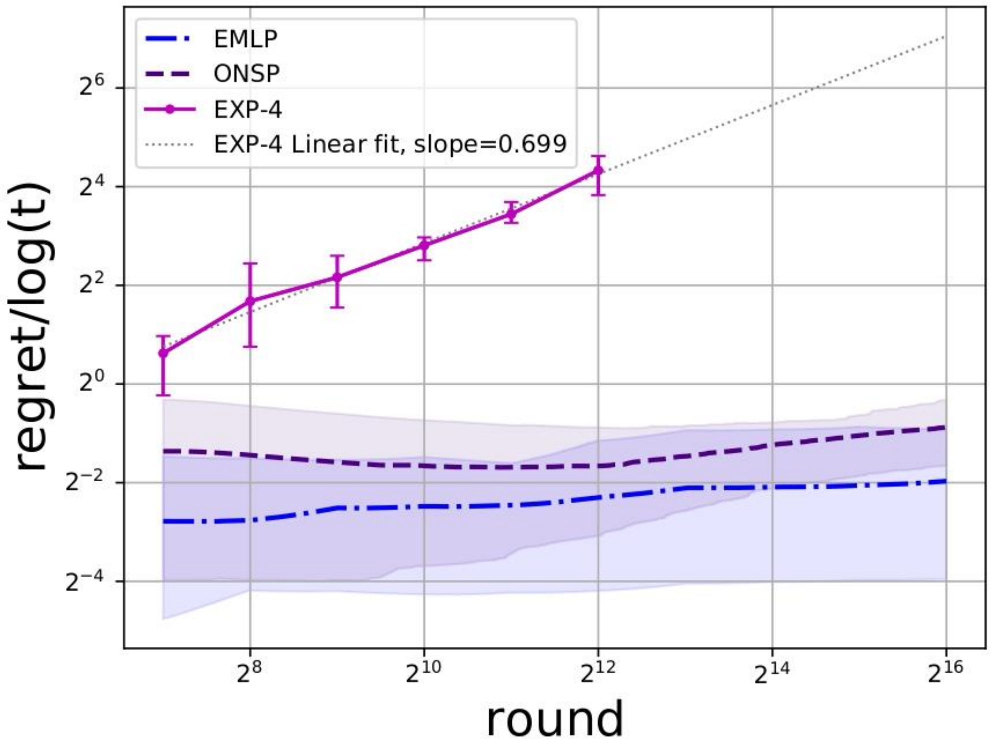
x_t's are selected by an oblivious **adversary** prior to all sales starting

Algorithm: ONSP

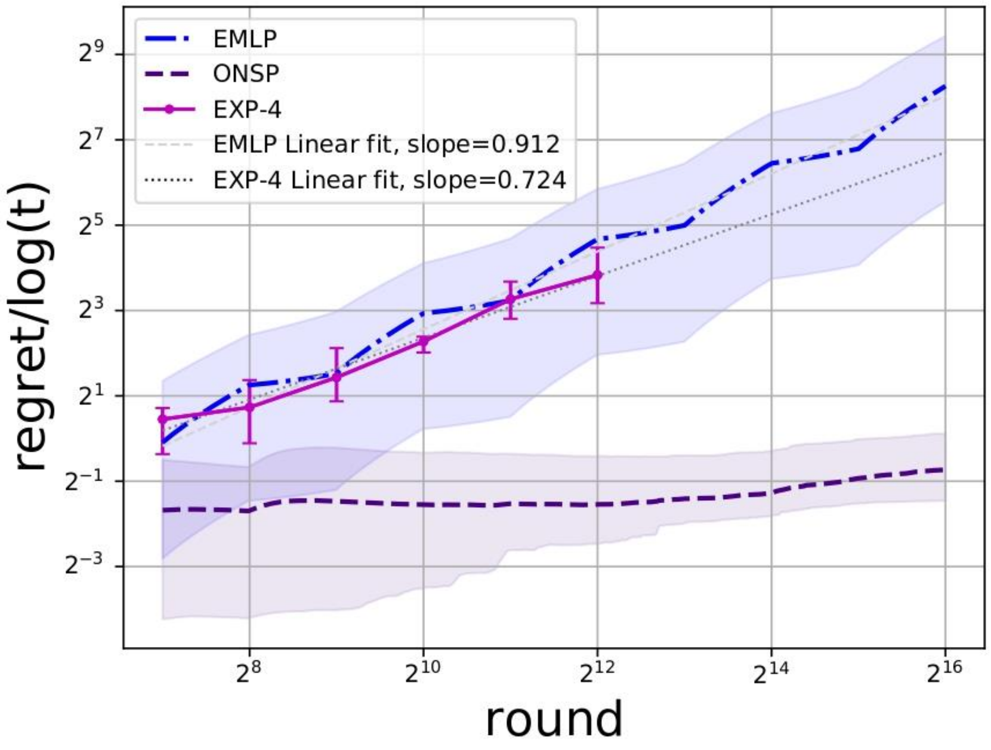
- Why a logarithmic regret?
  - Key: knowing noise distribution $\mathbb{D}$.
  - Prove an $\Omega(\sqrt{T})$ lower bound for $\mathcal{N}(0, \sigma^2)$ noise with unknown σ.

Numerical Result – LV with Known Gaussian Noise

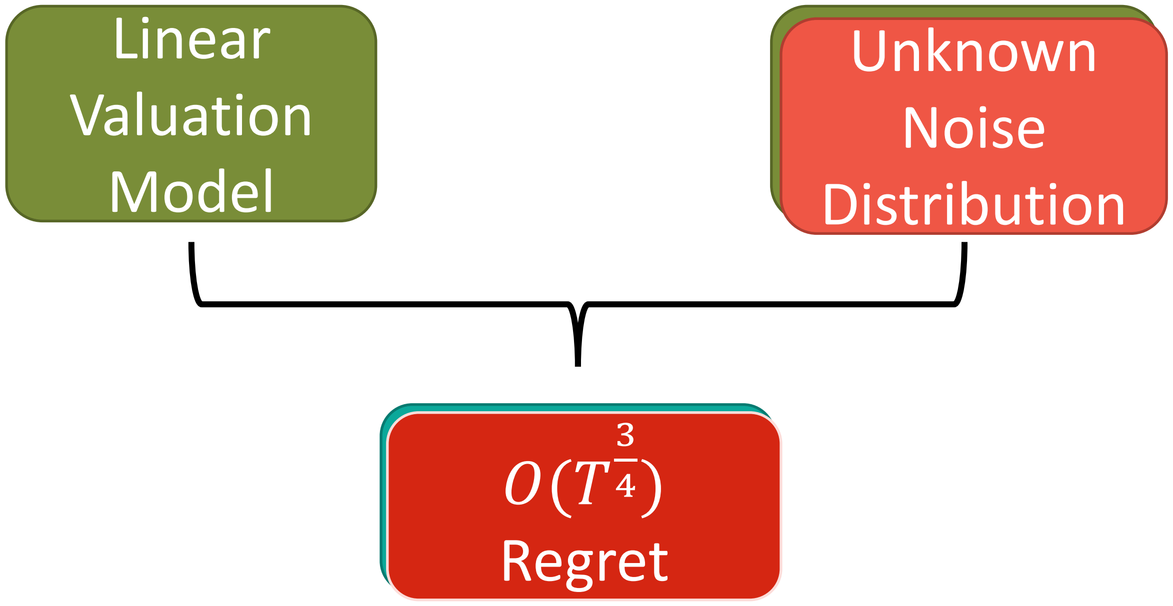
Stochastic x_t's



Adversarial x_t's



LV: $\tilde{O}(T^{\frac{3}{4}})$ Regret for *Unknown* Noise Distributions [XW22]



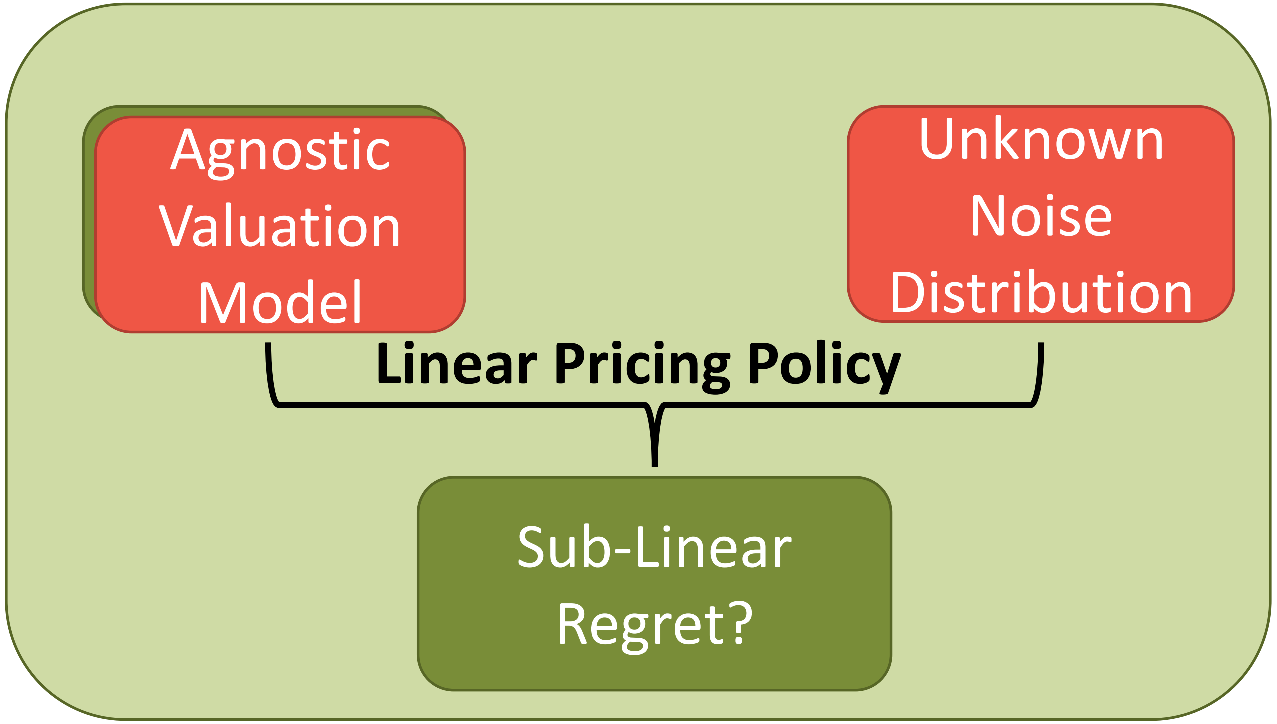
In [XW22], we …

- Propose a contextual-bandit-based algorithm: **D2-EXP4**.
- Half-space feedback⇒ **Half-Lipschitz**
- Achieve $\tilde{O}\left(T^{\frac{3}{4}} + d^{\frac{1}{2}}T^{\frac{5}{8}}\right)$ regret.
- Prove an $\tilde{\Omega}\left(T^{\frac{2}{3}}\right)$ regret lower bound.
  - Works even for Lipschitz noise CDF.

We do **NOT** have …

- Knowledge on noise models (besides being bounded).
- Assumptions on noise distributional functions.

Linear Policy (LP): Compromise for Agnostic [XW22]



No assumptions on valuation y_t at all!

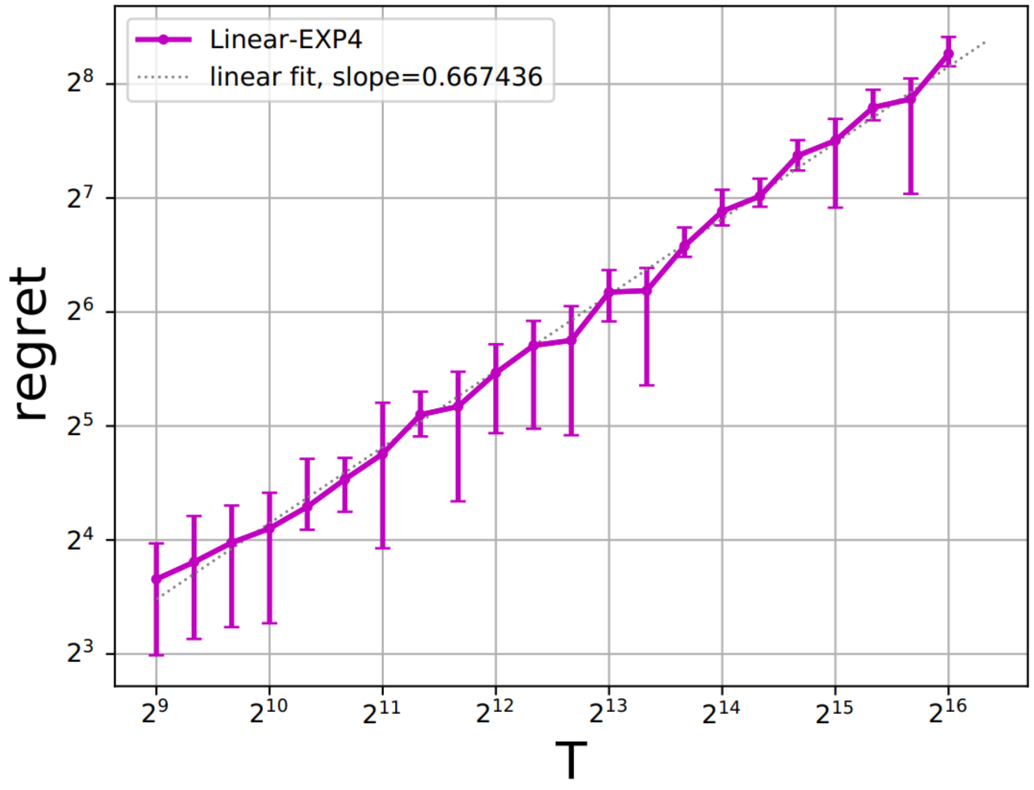
- y_t can be chosen *arbitrarily*.
- *Not possible* to learn the $x_t \to y_t$ model:
  - even *not necessarily exists*!

What if we *only* compete with **linear** pricing policies?

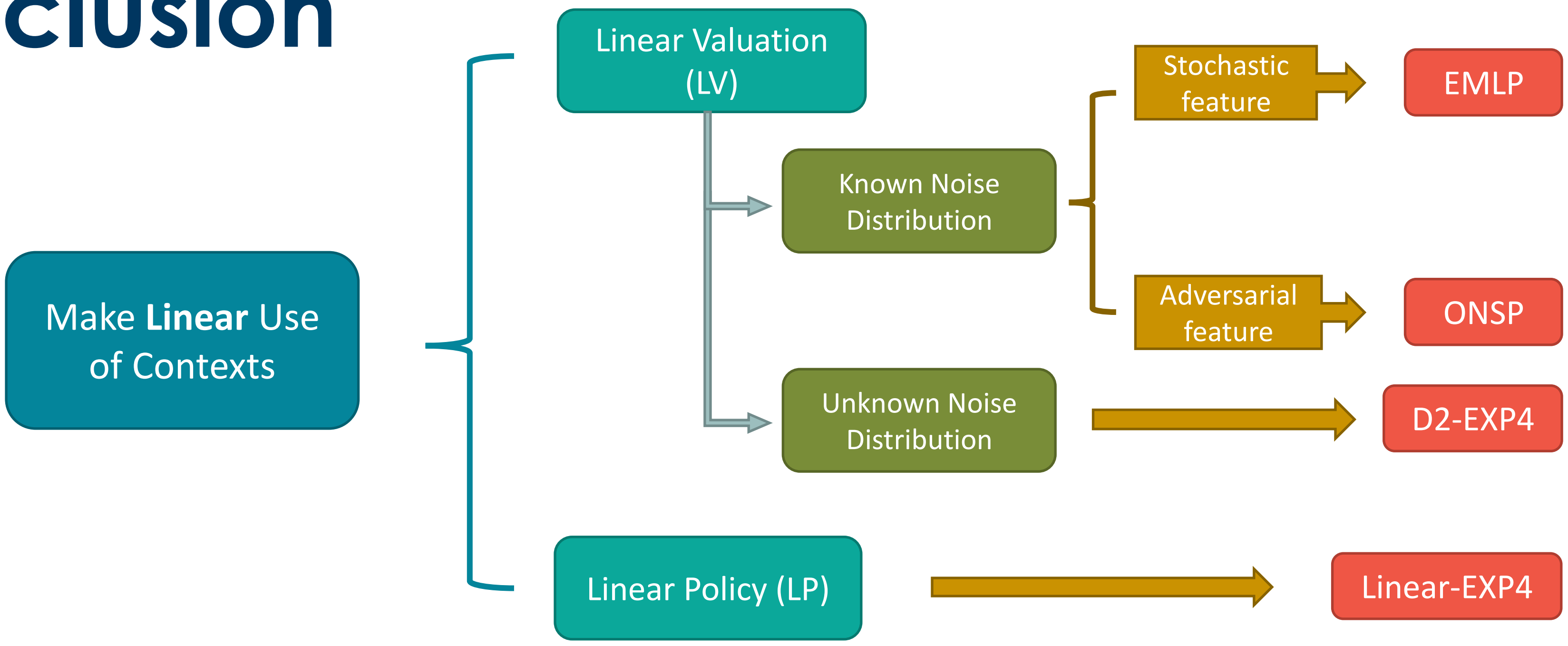
- $\eta^* = \arg\max_{\eta} \mathbb{E}[\sum_{t=1}^{T} r_t(v_t) | v_t = x_t^\top \eta]$
- Regret:= $\mathbb{E}[\sum_{t=1}^{T} r_t(x_t^\top \eta^*) - \sum_{t=1}^{T} r_t]$
- This is a **Linear Policy** (LP) problem!
- Can we achieve sub-linear regret?

Linear Policy (LP): $\tilde{O}\left(T^{\frac{2}{3}}d^{\frac{1}{3}}\right)$ Optimal Regret [XW22]

- We design an algorithm: **Linear-EXP4**.
  - Idea: discretization on action & parameter spaces.
- We achieve a $\tilde{O}\left(T^{\frac{2}{3}}d^{\frac{1}{3}}\right)$ **optimal** regret.
- We prove a matching $\tilde{\Omega}\left(T^{\frac{2}{3}}d^{\frac{1}{3}}\right)$ regret lower bound.



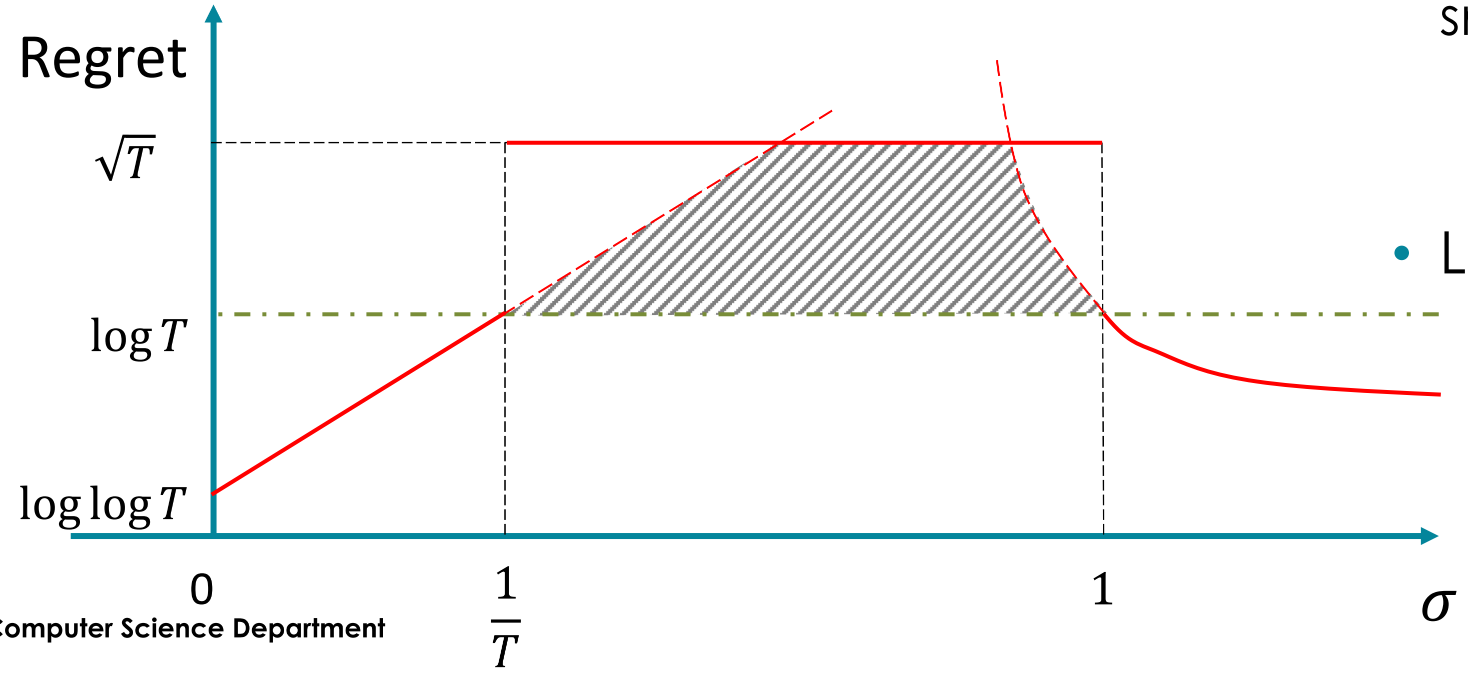
Conclusion



- For Linear Valuation problem ($y_t = x_t^\top \theta^* + N_t$), we achieve:
  - $O(d \log T)$ *optimal* regret for **known** noise distributions.
  - $\tilde{O}\left(T^{\frac{3}{4}} + T^{\frac{5}{8}} d^{\frac{1}{2}}\right)$ regret for **fully agnostic** noise distribution.
- For Linear Policy problem, we achieve $\tilde{O}\left(T^{\frac{2}{3}} d^{\frac{1}{3}}\right)$ *optimal* regret.

Open Problem: Regrets over different σ
(on LV with known noise distributions)

Noise σ	Regret	Trend
0	$O(d \log\log T)$	/
$\tilde{O}(1/T)$	$O(d^2 \log T)$	Increasing w.r.t. σ.
$(\Omega(1/T), O(1))$	$o(\sqrt{T})$	(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!

Reference

- [KL03] The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In FOCS-03.
- [CLPL16] Feature-based dynamic pricing. In EC-16.
- [PLS18] Contextual search via intrinsic volumes. In FOCS-18.
- [JN19] Dynamic pricing in high-dimensions. JMLR.
- [**X**W21] Logarithmic regret in feature-based dynamic pricing. In NeurIPS-21.
- [WCSL21] Multimodal dynamic pricing. Management Science.
- [LS21] Distribution-free contextual dynamic pricing. arXiv preprint.
- [FGY22] Policy optimization using semiparametric models for dynamic pricing. JASA.
- [WTL21] On Dynamic Pricing with Covariates. arXiv preprint arXiv:2112.13254.
- [BK21] Personalized dynamic pricing with machine learning: High-dimensional features and heterogeneous elasticity. Management Science.
- [**X**W22] Towards Agnostic Feature-based Dynamic Pricing: Linear Policies vs Linear Valuation with Unknown Noise. In AISTATS-22.

UC SANTA BARBARA