Online Linear Programming: Applications and Extensions

Yinyu Ye

Stanford University and CUHKSZ (sabbatical leave)

(Joint work with many)

June 20, 2023

Organization

- Online Linear Programming for Auction Markets
- Online Linear Programming for Bandit Markets and More
- Online Mechanism for Price-Posting Markets
- Conclusion/Takeaway

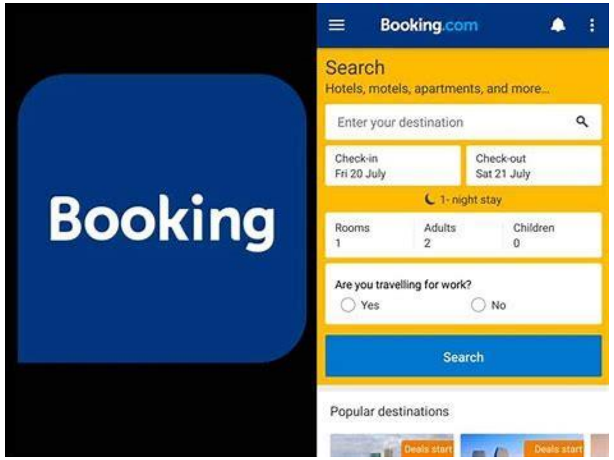
Linear Programming and LP Giants won Nobel Prize…

$$\max \quad \sum r_j x_j$$

$$\text{s.t.} \quad \sum_j \boldsymbol{a}_j x_j \le \boldsymbol{b},$$

$$0 \le x_j \le 1 \quad \forall\, j = 1, \ldots, n$$

Online Resource Allocation & Revenue Management via Combinatorial Auction



- m type of resources; T customers
- Decision maker needs to decide whether and how much resources are allocated to each customer/auctioner
- Resources are limited!
- **Online setting**:
  - Customers arrive sequentially and the decision needs to be made instantly upon the customer arrival: **Sell or No-sell**?

$$\max \sum_{t=1}^{T} r_t x_t$$

$$\text{s.t.} \sum_{t=1}^{T} a_{it} x_t \le b_i, \quad i = 1, \ldots, m$$

$$0 \le x_t \le 1 \quad \text{or} \quad x_t \in \{0, 1\}, \quad t = 1, \ldots, T$$

Performance of online algorithm measured with respect to regret from the offline linear objective

[Agrawal et al. 2010, 2014], [Kesselheim et al 2014]

[Li/Ye, 2019], [Li et al. 2020],

Online Auction Market: An Illustration Example

Bid #	$100	$30	….	…	…	Inventory	
Decision	X1=?	X2=?					
Pants	1	0	….	…	…	100	
Shoes	1	0				50	
T-Shirts	0	1				500	
Jackets	0	0				200	
Hats	1	1	…	…	…	1000	

Regret-Ratio for Online Algorithm/Mechanism

$$\begin{aligned} \mathrm{OPT}(A,\pi) = \max \quad & \sum_k \pi_k x_k \\ \text{s.t.} \quad & \sum_k a_{ik} x_k \le b_i \quad \forall\, i \in S \\ & 0 \le x_k \le 1 \quad \forall\, k \in N \end{aligned}$$

- We know the total number of customers, say *n;*
- Assume customers arrive in a random order or with i.i.d distributions.
- For a given online algorithm/decision-policy/mechanism

$$Z(A,\pi) = E_\sigma[\ \sum_{1}^{n} \pi_k x_k\ ]$$

$$R(A,\pi) = 1 - \frac{Z(A,\pi)}{OPT(A,\pi)}$$

$$R = \sup_{(A,\pi)} R(A,\pi)$$

Impossibility Result on Regret-Ratio

Theorem: There is no online algorithm/decision-policy/mechanism such that

$$R \leq \mathrm{O}\left(\sqrt{\log(m)/B} \right), \quad B = \min{}_{\mathrm{i}} b_i .$$

Corollary: If $B \leq \log(m)/\varepsilon^2$, then it is impossible to have a decision policy/mechanism such that $R \leq \mathrm{O}(\varepsilon)$.

Agrawal, Wang and Y, “A Dynamic Near-Optimal Algorithm for Online Linear Programming,” 2010.

Possibility Result on Regret-Ratio

Theorem: There is an online algorithm/decision-policy/mechanism such that

$$R \le \mathrm{O}\left(\sqrt{m\log(n)/B} \right), \quad B = \min_i b_i .$$

Corollary: If $B > m\log(n)/\varepsilon^2$, then there is an online algorithm/decision-policy/mechanism such that $R \le \mathrm{O}(\varepsilon)$.

Agrawal, Wang and Y, "A Dynamic Near-Optimal Algorithm for Online Linear Programming," 2010.

Theorem: If $B > \log(mn)/\varepsilon^2$, then there is an online algorithm/decision-policy/mechanism such that $R \le \mathrm{O}(\varepsilon)$.

Kesselheim et al. "Primal Beat the Dual…," 2014, …

Online Algorithm and Price-Mechanism: Learning-while-Doing

- **Learn "ideal" itemized-prices**
- **Use the prices to price each bid**
- **Accept if it is an over bid, and reject otherwise**

Bid #	$100	$30	….	…	…	Inventory	Price?
Decision	x1	x2					
Pants	1	0	….	…	…	100	45
Shoes	1	0				50	45
T-Shirts	0	1				500	10
Jackets	0	0				200	55
Hats	1	1	…	…	…	1000	15

Such ideal prices exist and they are shadow/dual prices of the offline LP

How to Learn "Shadow Prices" Online

For a given ε, solve the sample LP at t=εn, 2εn, 4εn, …; and use the new shadow prices for the decision in the coming period.

$$\begin{aligned} \max \quad & \sum_{k=1}^{t} \pi_k x_k \\ \text{s.t.} \quad & \sum_{k=1}^{t} a_{ik} x_k \le (1-h_t)\frac{t}{n} b_i \quad \forall\, i \in S \\ & 0 \le x_k \le 1 \qquad\qquad \forall\, k \in N \end{aligned}$$

Dual Convergence and the SGD Method

(Li/Y OR 2022, LI/Sun NeurIPS 2020)

Primal LP

$$\begin{aligned} \max \quad & \boldsymbol{\pi}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{1} \end{aligned}$$

Dual LP

$$\begin{aligned} \min \quad & \boldsymbol{b}^\top \boldsymbol{p} + \boldsymbol{1}^\top \boldsymbol{s} \\ \text{s.t.} \quad & \boldsymbol{A}^\top \boldsymbol{p} + \boldsymbol{s} \geq \boldsymbol{\pi} \\ & \boldsymbol{p} \geq \boldsymbol{0}, \boldsymbol{s} \geq \boldsymbol{0} \end{aligned}$$

- An equivalent form of the dual LP can be written as (by plugging s into the objective function above):

$$\min_{\boldsymbol{p} \geq \boldsymbol{0}} \quad \boldsymbol{b}^\top \boldsymbol{p} + \sum_{t=1}^{T} \left(\pi_t - \boldsymbol{a}_t^\top \boldsymbol{p}\right)^+$$

where at time t, we observe the t-th term in the above summation

- Idea: Perform online (stochastic) gradient descent to optimize the above form
- Theorem: With a step size $1/\sqrt{T}$, the algorithm achieves a regret bound of $\mathrm{m}\sqrt{T}$ (m being the number of constraints)

Action-History-Dependent Analysis
(Li/Y OR 2022)

- Instead of online gradient descent, we can learn the dual price more accurately and adaptively by solving the following problem at time t

$$\min_{\boldsymbol{p} \geq 0} \quad \boldsymbol{b}_t^\top \boldsymbol{p} + \sum_{j=1}^{t} \left(\pi_t - \boldsymbol{a}_t^\top \boldsymbol{p}\right)^+$$

- Compared to the previous problem, we replace $\boldsymbol{b}$ by the average remaining resource $\boldsymbol{b}_t$ (more adaptively), and solve the optimization problem (more accurately)
- Denoted the optimal solution by $\boldsymbol{p}_t^*$ – the adaptively learned **dual price**.
- The decision rule becomes $x_t = \begin{cases} 1, & \pi_t > \boldsymbol{a}_t^\top \boldsymbol{p}_t^* \\ 0, & \pi_t < \boldsymbol{a}_t^\top \boldsymbol{p}_t^* \end{cases}$

Action-History-Dependent Analysis II

- If $\{(\boldsymbol{a}_t, \pi_t)\}_{t=1}^{T}$ follow a distribution that is independent and identically distributed (stationary), we have

Results

- 1) The estimation $\boldsymbol{p}_t^*$ converges to $\boldsymbol{p}^*$. i.e. $\mathbb{E}\left[\sum_{t=1}^{T} \|\boldsymbol{p}_t^* - \boldsymbol{p}^*\|_2^2\right] = O(\log T)$

- 2) The regret is of the order $O(\log T)$. More specifically, if we denote Π our dual-based decision rule, we have

$$\mathbb{E}\left[\mathrm{OPT}(A, \pi)\right] - \mathbb{E}^{\Pi}\left[\sum_{t=1}^{T} \pi_t x_t\right] = O(\log T)$$

Improved OLP analysis I (Chen et al OR 2022)

- Now let's assume that $\{(\boldsymbol{a}_t, \pi_t)\}_{t=1}^T$ come from a distribution that has finite (with a total of J) categories, and $P\left((\boldsymbol{a}_t, \pi_t) = (\boldsymbol{c}_j, \mu_j)\right) = p_j$.

- For this case, we construct the fluid approximation LP

Primal LP

$$\begin{aligned} \max \quad & \boldsymbol{\pi}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{1} \end{aligned}$$

⟺

Fluid approximation LP

$$\begin{aligned} \max \quad & \sum_{j=1}^J p_j \mu_j y_j \\ \text{s.t.} \quad & p_j \boldsymbol{c}_j y_j \leq \boldsymbol{b}/T \\ & 0 \leq y_j \leq 1 \ \text{ for all } j \end{aligned}$$

- For the decision rule, if $(\boldsymbol{a}_t, \pi_t) = (\boldsymbol{c}_j, \mu_j)$, the optimal decision is $x_t = y_j^*$

Improved OLP analysis II

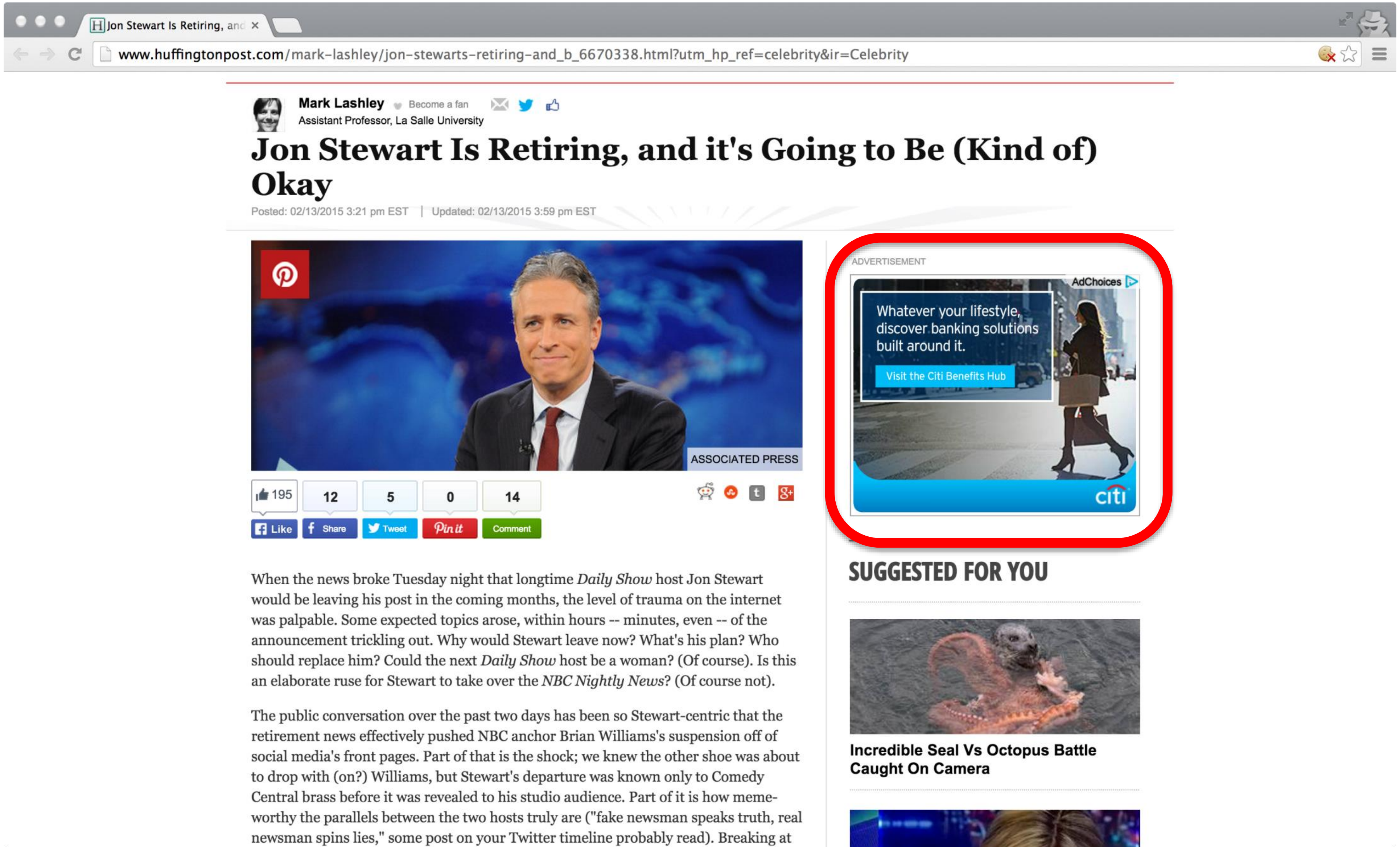
- At time t, we replace $\boldsymbol{b}$ and p_j by $\boldsymbol{b}_t$, the remaining resource, and $\hat{p}_j$, the **sample estimation** of p_j.
- Next, we solve the fluid approximation LP with updated parameters.
- Our decision at t will be based on the solution $\boldsymbol{y}_t^* = (y_{1,t}^*, \cdots, y_{J,t}^*)$.

Result

- The regret is of the order $O(1)$. More specifically, if we denote Π the decision rule above, we have

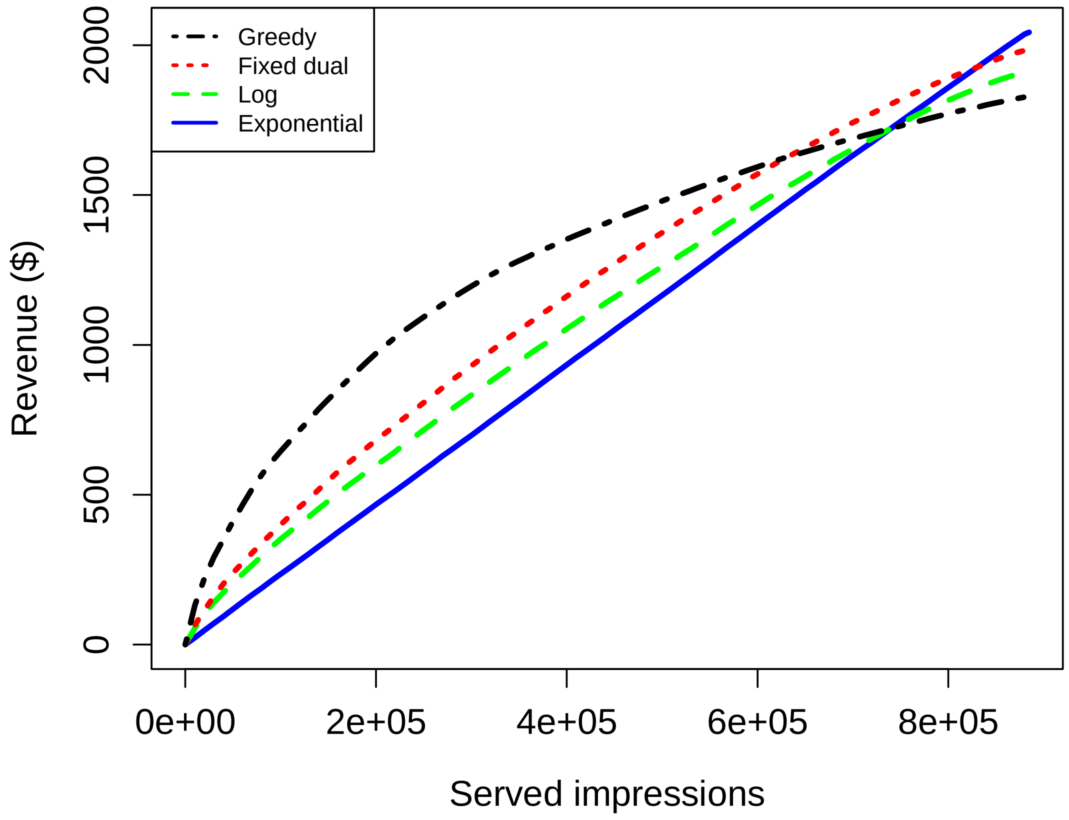
$$\mathbb{E}\left[\mathrm{OPT}(A,\pi)\right] - \mathbb{E}^{\Pi}\left[\sum_{t=1}^{T} \pi_t x_t\right] = O\left(1\right)$$

Application: Online Matching for Display Advertising



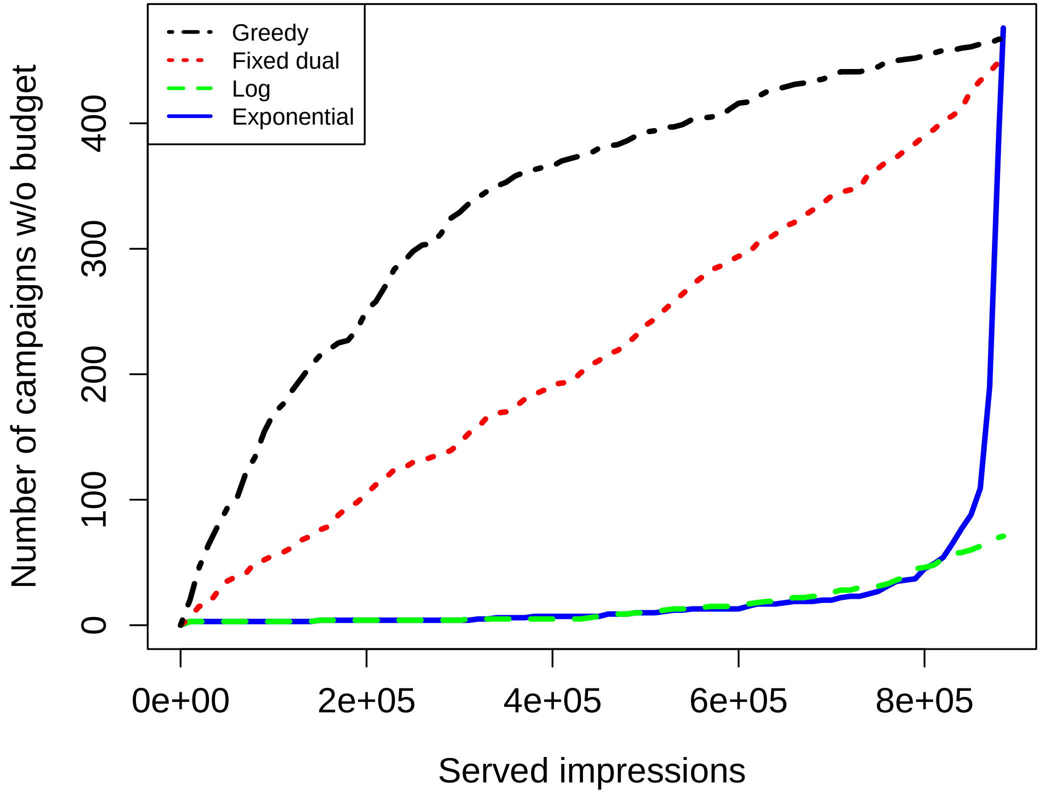
Revenues generated by different methods

- Total Revenue for impressions in T2 by Greedy and OLP with different allocation risk functions

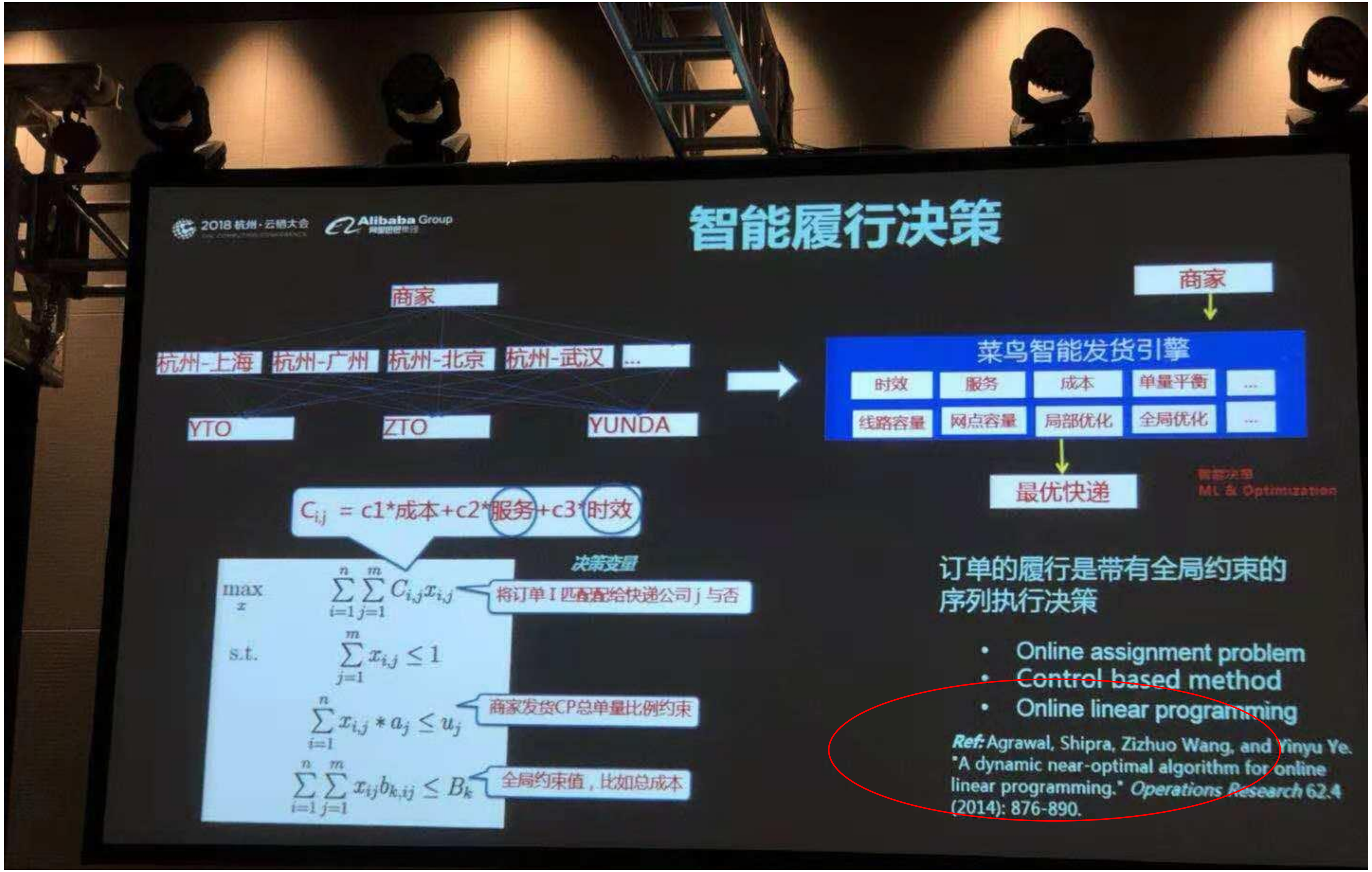


of Out-of-Budget Advertisers

- Greedy exhausts budget of many advertisers early.
- Log penalty keeps advertisers in budget but it is very conservative.
- Exponential penalty Keeps advertisers in budget until almost the end of the timeframe.



阿里巴巴在2019年云栖大会上提到在智能履行决策上使用OLP的算法



$$\max_x \sum_{i=1}^{n}\sum_{j=1}^{m} C_{i,j}x_{i,j}$$
$$\text{s.t.} \quad \sum_{j=1}^{m} x_{i,j} \le 1$$
$$\sum_{i=1}^{n} x_{i,j} * a_j \le u_j$$
$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij} b_{k,ij} \le B_k$$

阿里巴巴团队在2020年CIKM会议论文Online Electronic Coupon Allocation based on Real-Time User Intent Detection上提到他们设计的发红包的机制也使用了OLP的方法[2]

Spending Money Wisely: Online Electronic Coupon Allocation based on Real-Time User Intent Detection

Liangwei Li*
Liucheng Sun*
leon.llw@alibaba-inc.com
liucheng.slc@alibaba-inc.com
Alibaba Group
Hangzhou, Zhejiang

Chenwei Weng
wengchenwei.pt@alibaba-inc.com
Alibaba Group
Hangzhou, Zhejiang

Chengfu Huo
chengfu.huocf@alibaba-inc.com
Alibaba Group
Hangzhou, Zhejiang

Weijun Ren
afei@alibaba-inc.com
Alibaba Group
Hangzhou, Zhejiang

$$\begin{aligned} \max & \sum_{i=1}^{M}\sum_{j=1}^{N} v_{ij}x_{ij} \\ s.t. & \sum_{i=1}^{M}\sum_{j=1}^{N} c_j x_{ij} \le B, \\ & \sum_{j}^{N} x_{ij} \le 1, \quad \forall i \\ & x_{ij} \ge 0, \quad \forall i, j \end{aligned} \tag{5}$$

3.3 MCKP-Allocation

We adopt the primal-dual framework proposed by [2] to solve the problem defined in Equation 5. Let α and β_j be the associated dual variables respectively. After obtaining the dual variables, we can solve the problem in an online fashion. Precisely, according to the principle of the primal-dual framework, we have the following allocation rule:

$$x_{ij} = \begin{cases} 1, & \text{where } j = \arg\max_i (v_{ij} - \alpha c_j) \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

Online learning algorithms can also be applied to more general programming

- n energy suppliers with privately known convex cost functions c_i
- Customer demand d for energy
- How to find equilibrium prices to match supply and demand without information on cost functions?
- **[Jalota, Sun, Azizan, 2023]** develop online learning algorithms with sub-linear regret:
  - O(log log T) for static cost functions and demands
  - O($\sqrt{T}$ log log T) for static costs, varying demands
  - O($T^{2/3}$) for varying costs and finite function class

$$C^* = \min_{x_i \geq 0, \forall i \in [n]} \sum_{i=1}^{n} c_i(x_i),$$
$$\text{s.t.} \quad \sum_{i=1}^{n} x_i = d,$$

Online Learning for Equilibrium Pricing in Electricity Markets under Incomplete Information

Organization

- Online Linear Programming for Auction Markets
- Online Linear Programming for Bandit Markets and More
- Online Mechanism for Price-Posting Markets
- Conclusion/Takeaway

The Online Algorithm can be Applied to Bandits with Knapsack (BwK) Applications

- For the previous problem, the decision maker first wait and observe the customer order/arm and then decide whether to accept/play it or not.
- An alternative setting is that the decision maker first decides which order/arm (s)he may accept/play, and then receive a random resource consumption vector **a**$_j$ and yield a random reward π_j of the pulled arm.
- Known as the Bandits with Knapsacks, and it is a tradeoff exploration v.s. exploitation

$$\max \quad \sum \pi_j x_j \quad \text{s.t.} \quad \sum_j \boldsymbol{a}_j x_j \le \boldsymbol{b}\,, \quad x_j \quad \ge 0 \qquad \forall\ j = 1,\ldots,J$$

- The decision variable x_j represents the total-times of pulling the j-th arm.
- We have developed a two-phase algorithm
  - Phase I: Distinguish the optimal super-basic variables/arms from the optimal non-basic variables/arms with as fewer number of plays as possible
  - Phase II: Use the arms in the optimal face to exhaust the resource through an adaptive procedure and achieve fairness
- The algorithm achieves a problem dependent regret that bears a logarithmic dependence on the horizon T. Also, it identifies a number of LP-related parameters as the bottleneck or condition-numbers for the problem
  - Minimum non-zero reduced cost
  - Minimum singular-values of the optimal basis matrix.
- **First algorithm** to achieve the $\boldsymbol{O(\log T)}$ regret bound [Li, Sun & Y 2021 ICML] (**https://proceedings.mlr.press/v139/li21s.html**)

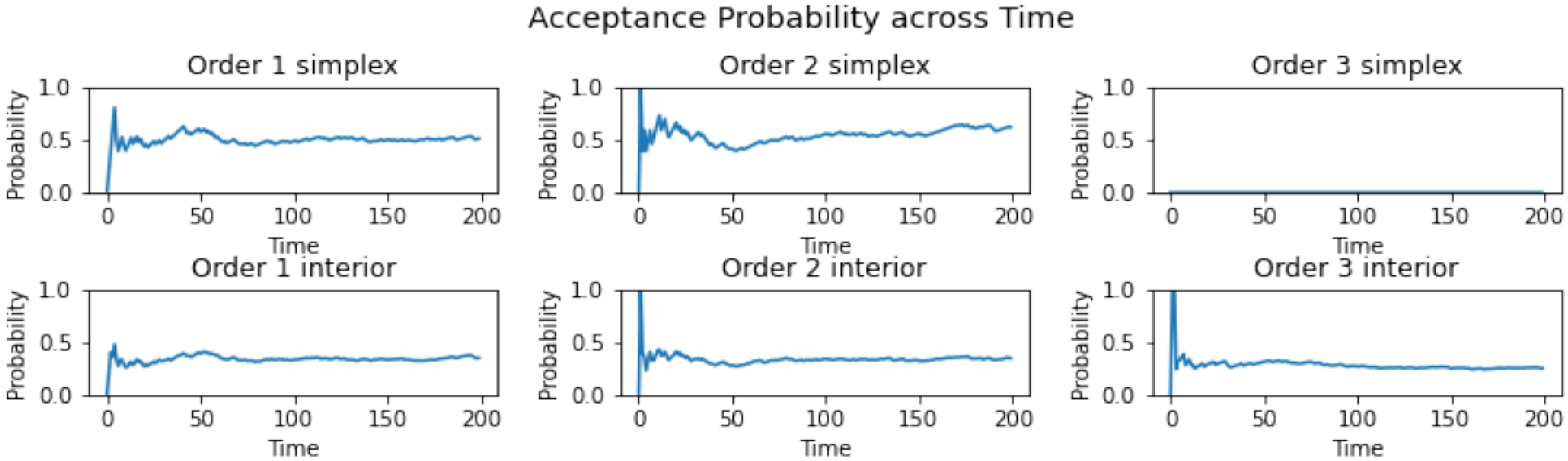
Fairness: there are many settings when we need to fairly allocate shared resources to users

Public Good Allocation

Vaccine Allocation

A Motivation Example

- Consider an allocation problem: there exists three types of orders/customers, where the first two types have the reward/resource characteristics that are considered equivalent from the system.
- The following plots show the acceptance fraction/probability of the three types across time by two different online algorithms: the simplex and interior-point methods (Jasin 2015, Chen et al 2021).



Fairness Desiderata

- Technically, Non-Uniqueness/Degeneracy degrades the quality of online algorithm since the learning “targets” are ambiguous – no ground-truth.

- More importantly, Individual Fairness needs to be achieved: similar customers should be treated similarly. Since the optimal object value depends on the total resources spent, not on the resources spent on which groups, some individual or group may be ignored by a particular online algorithm/allocation-rule.

- Also, Time Fairness: The algorithm may tend to accept mainly the first half (or the second half of the orders), which is unfair or unideal…

Fair OLP Model and Algorithm

$$\max \sum_{j=1}^{J} p_j \mu_j y_j \quad \text{s.t.} \quad \sum_{j=1}^{J} p_j c_j y_j \leq b/T, \quad y_j \in [0,1]$$

- We define $\boldsymbol{y}^*$ the fair offline optimal solution of the LP problem as the analytical center of the optimal solution set, which represents an "average" of all the optimal corner solutions – their product is maximized.
- The fair solution $\boldsymbol{y}^*$ will treat individuals fairly, based on their similar reward and resource consumption.
- An online interior-point learning algorithm would use the data points up to time t and solve the sample-based linear program to decide fair $\boldsymbol{y}_t$.
- We give provable time and individual fairness guarantees.

Fairness-Performance Measure

- Let $\boldsymbol{y}_t$ be the allocation rule at time t which encodes the accepting probabilities under the online algorithm π. Then we define the cumulative unfairness of the online algorithm π as

$$UF_T(\pi) = E[\sum_{t=1}^{T} ||\boldsymbol{y}_t - \boldsymbol{y}^*||_2^2]$$

- Intuition: If $UF_T(\pi)$ is sub-linear, we know Time Fairness is satisfied since the deviation of the online solution cannot be large. Moreover, Individual Fairness is satisfied because we know $UF_T(\pi)$ being sub-linear implies $\boldsymbol{y}_t$ converging to $\boldsymbol{y}^*$.
- Let j_t denote the incoming customer type at time t, the Revenue Regret is defined as
- $$\text{Reg}_T(\pi) = E[\sum_{t=1}^{T} r_t(y_{j_t}^* - y_{t,j_t})]$$

Regret measures the performance loss compared to the optimal policy.

Our Result

- We develop an algorithm [Chen, Li & Y (2021)] that achieve

$$UF_T(\pi) = O(\log T)$$

$$Reg_T(\pi) \text{ Bounded independent of } T$$

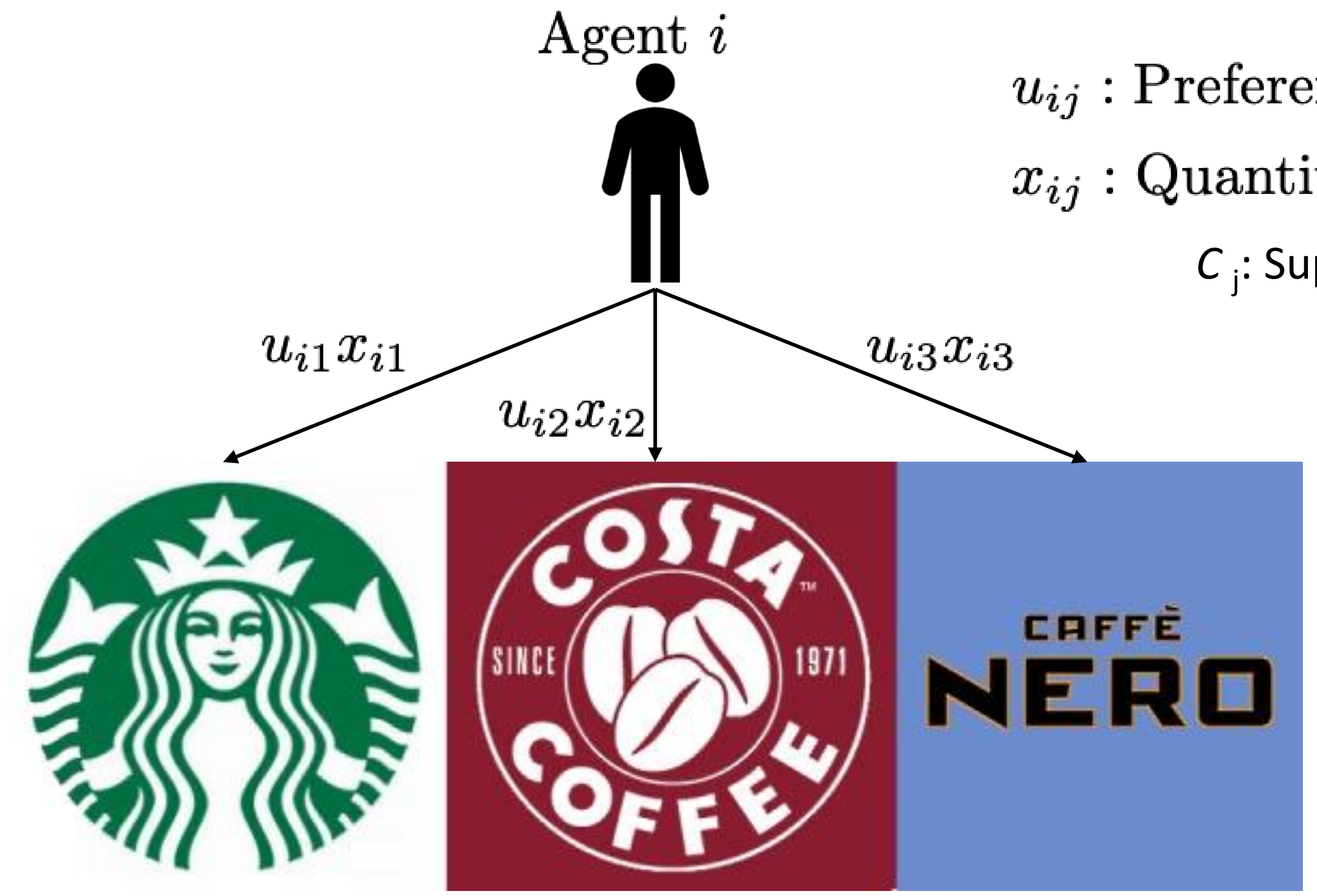
- Key ideas in algorithm design:
  - At each time t, we use interior-point method to obtain the sample analytic-center solution and randomly make decision based on sample solution $\boldsymbol{y}_t$.
  - We also adjust the right-hand-side resource of the LP to ensure the depletion of binding resources and non-binding resources does not affect the fairness.
  - This state of the art result removes typical non-degeneracy or non-uniqueness assumption in the OLP literature.

(Chen et al. arXiv:2110.14621 2021)

Organization

- Online Linear Programming for Auction Markets
- Online Linear Programming for Bandit Markets and More
- Online Mechanism for Price-Posting Markets
- Conclusion/Takeaway

One of Price-Posting Markets to the Fisher Market



p_1, C_1 p_2, C_2 p_3, C_3

$M =$ Total Number of Goods

u_{ij} : Preference of Agent i for one unit of good j

x_{ij} : Quantity of good j purchased by person i

C_j: Supply Amount, p_j : Price of Good j

w_i : Budget of Agent i

Individual Optimization Problem:

$$\max_{\mathbf{x}_i} \sum_j u_{ij} x_{ij}$$

$$\text{s.t. } \mathbf{p}^T \mathbf{x}_i \leq w_i$$

$$\mathbf{x}_i \geq \mathbf{0}$$

Prices are posted and known to all agents so that they have freedom to choose

Are there Prices to Clear Market?
Yes, and they can be derived from the Eisenberg-Gale optimization problem

Individual Optimization Problem:

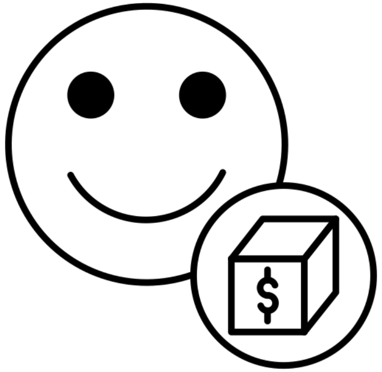
$$\max_{\mathbf{x}_i} \sum_j u_{ij} x_{ij}$$
$$\text{s.t. } \mathbf{p}^T \mathbf{x}_i \leq w_i$$
$$\mathbf{x}_i \geq \mathbf{0}$$

$\Longleftrightarrow$

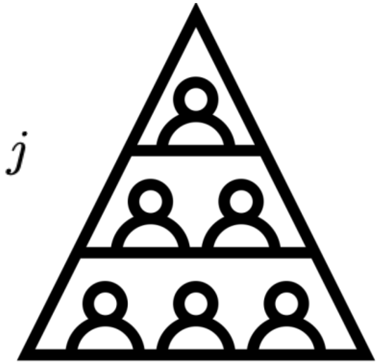
Social Optimization Problem:

$$\max_{\mathbf{x}_i, \forall i \in [N]} \sum_i w_i \log \left(\sum_j u_{ij} x_{ij} \right)$$
$$\text{s.t. } \sum_i x_{ij} \leq c_j, \forall j \in [M]$$
$$x_{ij} \geq 0, \forall i, j$$

Capacity Constraints



p_j : Price of Good j = Dual Variable of Constraint j



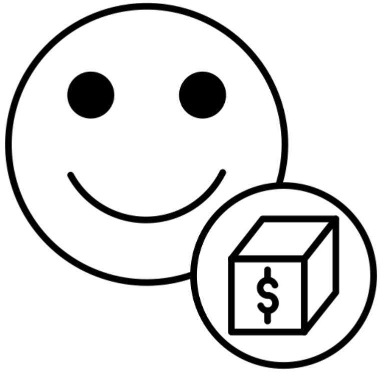
However, the applicability of Fisher markets is restricted to the "Perfect and Static Information Setting"

Individual Optimization Problem:

$$\max_{\mathbf{x}_i} \sum_j u_{ij} x_{ij}$$

$$\text{s.t. } \mathbf{p}^T \mathbf{x}_i \leq w_i$$

$$\mathbf{x}_i \geq \mathbf{0}$$



$\Longleftrightarrow$

Social Optimization Problem:

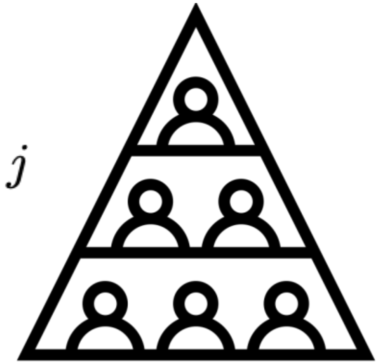
$$\max_{\mathbf{x}_i, \forall i \in [N]} \sum_i w_i \log \left(\sum_j u_{ij} x_{ij} \right)$$

$$\text{s.t. } \sum_i x_{ij} \leq c_j, \forall j \in [M]$$

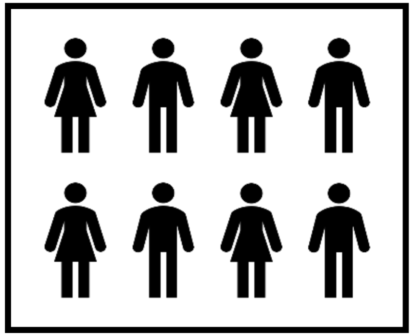
Capacity Constraints

$$x_{ij} \geq 0, \forall i, j$$

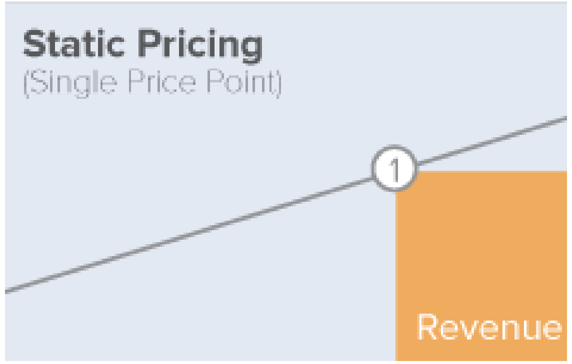
p_j : Price of Good j = Dual Variable of Constraint j



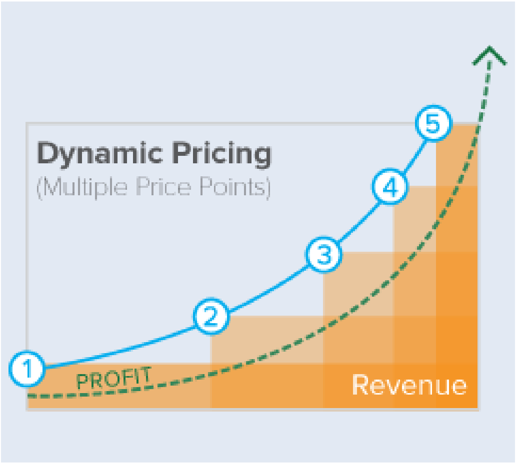
We study an online and incomplete information variant of Fisher markets



Buyers arrive sequentially with utility and budget parameters drawn i.i.d. from a distribution



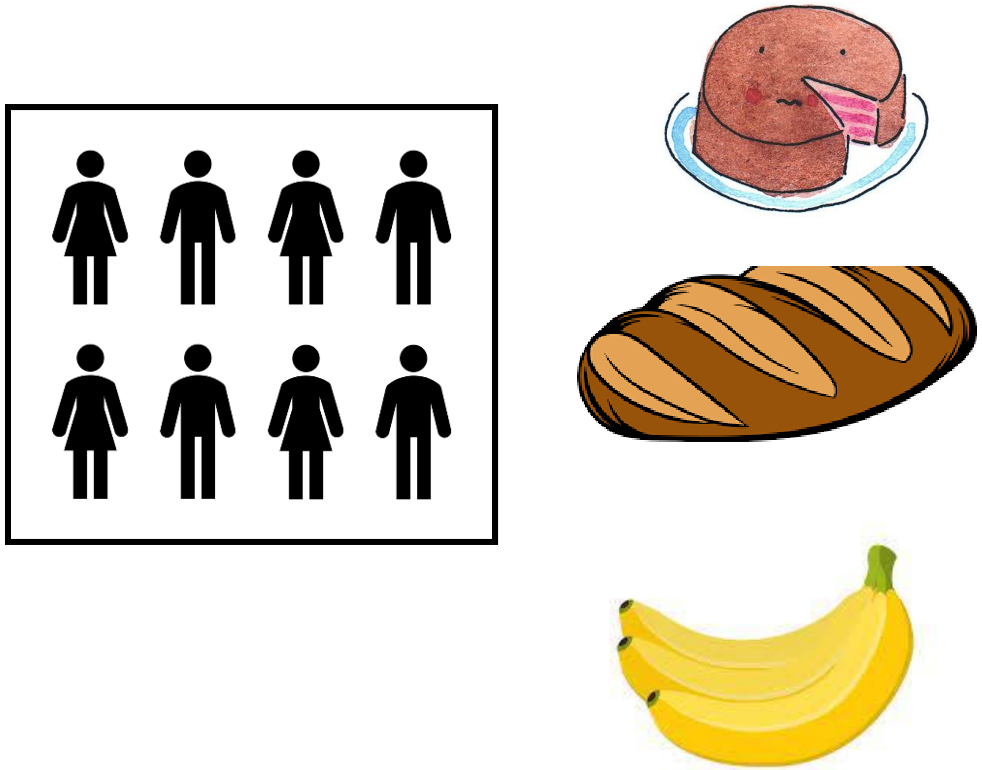
Establish performance limits of static pricing algorithms, including one that sets expected equilibrium prices

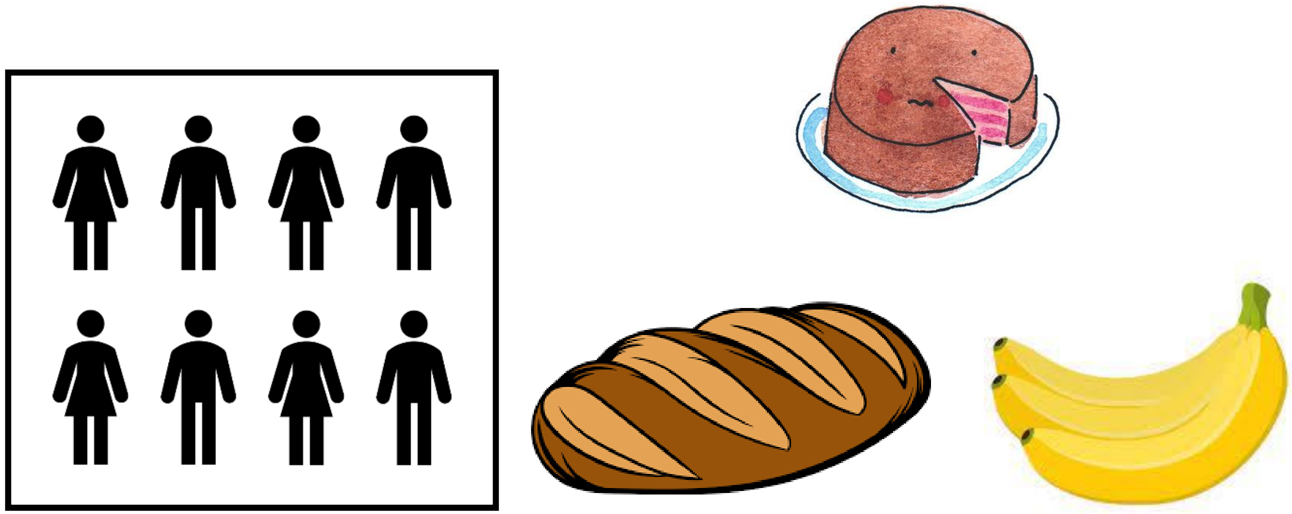


Develop an adaptive expected equilibrium pricing approach with strong performance guarantees

Develop a revealed preference algorithm with sub-linear regret and capacity violation

Prior work on online variants of Fisher markets have considered the setting of goods arriving sequentially





Prior Work: Goods Arrive Online
[Gorokh, Banerjee, Iyer, 2021]

This Work: Agents Arrive Online

Online for Geometric Objective: evaluate algorithms through the absolute regret of social welfare and capacity violation

Regret (Optimality Gap)

Difference in the Optimal Social Objective of the online policy $\boldsymbol{\pi}$ to that of the optimal offline social value

$$R_n(\boldsymbol{\pi}) = \sum_i w_i \log\left(\sum_j u_{ij} x^*_{ij}\right) - \sum_i w_i \log\left(\sum_j u_{ij} x_{ij}(\boldsymbol{\pi})\right)$$

Optimal Offline Objective

Objective of online policy

Prior Work on concave objectives [Lu, Balserio, Mirrkoni, 2020] assume non-negativity and boundedness of utilities, none of which are true for the log objective

Constraint Violation

Norm of the violation of capacity constraints of the online policy $\boldsymbol{\pi}$

$$V_j(\boldsymbol{\pi}) = \sum_j x_{ij}(\boldsymbol{\pi}) - c_j$$

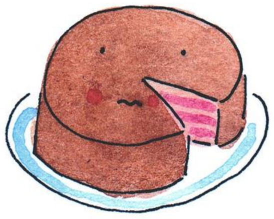
Violation of Capacity Constraint of good *j*

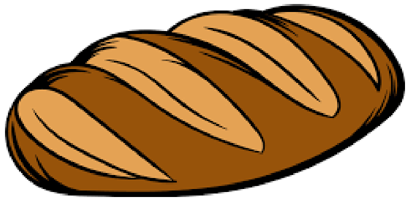
$$V_n(\boldsymbol{\pi}) = ||\mathbb{E}[V(\boldsymbol{\pi})^+]||_2$$

Norm of the expected constraint violation

Using the optimal expect prices, the capacity violation must be $\Omega(\sqrt{n})$, where n is the number of total agents

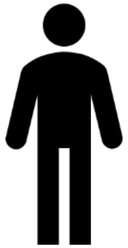
2 goods, each with a capacity of n





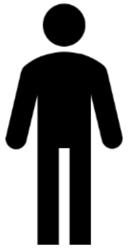
Two agent types specified by (Utility for Good 1, Utility for Good 2)

Type I: (1, 0)



Arrival Probability = 0.5

Type II: (0, 1)



Arrival Probability = 0.5

Expected Optimal Objective $\approx n\, log(2)$
Since Type I users receive two units of good one, while type two receive two units of good two

While $\frac{n}{2}$ users of Type I arrive in expectation, the realized arrivals of type I users deviates by $O(\sqrt{n})$

Theorem: More generally, any static pricing algorithm achieves either a regret or capacity violation of $\Omega(\sqrt{n})$

To set static expected equilibrium prices, we can solve the following deterministic problem

Assumption: The distribution from which the utility and budget parameters of users are drawn is discrete with finite support, where $\mathbb{P}((w_t, \mathbf{u}_t) = (\tilde{w}_k, \tilde{\mathbf{u}}_k)) = q_k$ for all $k \in [K]$

CE($\boldsymbol{d}$)

$$\max_{\mathbf{z}_k \in \mathbb{R}^m, \forall k \in [K]} \quad U(\mathbf{z}_1, ..., \mathbf{z}_K) = \sum_{k=1}^{K} q_k \tilde{w}_k \log \left(\sum_{j=1}^{m} \tilde{u}_{kj} z_{kj} \right),$$

$$\text{s.t.} \quad \sum_{k=1}^{K} z_{kj} q_k \leq d_j, \quad \forall j \in [m],$$

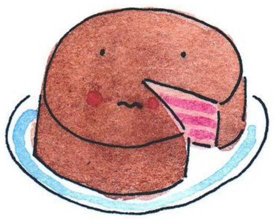
$$z_{kj} \geq 0, \quad \forall k \in [K], j \in [m],$$

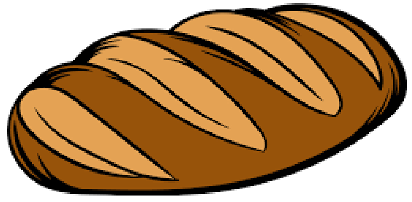
Average resource capacity per user

Dual variables of the capacity constraints are the static expected equilibrium prices

Example: For two-good counterexample, $K = 2$, $(\tilde{w}_1, \tilde{\boldsymbol{u}}_1) = (1, (1, 0))$, $(\tilde{w}_2, \tilde{\boldsymbol{u}}_2) = (1, (0, 1))$, $q_1 = q_2 = 0.5$
Static expected equilibrium price vector: $(0.5, 0.5)$

We overcome problem of static expected equilibrium pricing by increasing prices of over-consumed goods



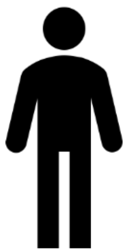


100 units

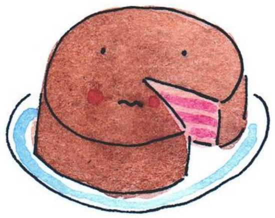
100 units

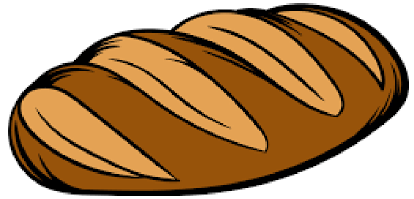
Solve CE($\boldsymbol{d}$) to set price of 0.5 for each good
For 100 users, $d_1 = 1, d_2 = 1$

Type I: (1, 0)



User of Type I and consumes two units of good one





98 units

100 units

1. Update Average remaining Resource Capacities

$$d_1^1 = \frac{98}{99}, d_2^1 = \frac{100}{99}$$

2. Solve CE($\boldsymbol{d}^1$) to set price for next user

Our adaptive expected equilibrium pricing approach achieves constant constraint violation and log regret

Algorithm 1: Adaptive Expected Equilibrium Pricing

Input : Initial Good Capacities $\mathbf{c}$, Number of Users n, Threshold Parameter Vector Δ, Support of Probability Distribution $\{\tilde{w}_k, \tilde{\mathbf{u}}_k\}_{k=1}^K$, Occurrence Probabilities $\{q_k\}_{k=1}^K$

Initialize $\mathbf{c}_1 = \mathbf{c}$ and the average remaining good capacity to $\mathbf{d}_1 = \frac{\mathbf{c}}{n}$;

for $t = 1, 2, ..., n$ **do**

  Phase I: Set Price

  if $\mathbf{d}_{t'} \in [\mathbf{d} - \Delta, \mathbf{d} + \Delta]$ *for all* $t' \leq t$ **then**

    Set price $\mathbf{p}^t$ as the dual variables of the capacity constraints of the certainty equivalent problem $CE(\mathbf{d}_t)$ with capacity $\mathbf{d}_t$;

  else

    Set price $\mathbf{p}^t$ using the dual variables of the capacity constraints of the certainty equivalent problem $CE(\mathbf{d})$ with capacity $\mathbf{d} = \mathbf{d}_1$;

  end

  Phase II: Observed User Consumption and Update Available Good Capacities

  User purchases optimal bundle of goods $\mathbf{x}_t$ given price $\mathbf{p}^t$;

  Update the available good capacities $\mathbf{c}_{t+1} = \mathbf{c}_t - \mathbf{x}_t$;

  Compute the average remaining good capacities $\mathbf{d}_{t+1} = \frac{\mathbf{c}_{t+1}}{n-t}$;

end

Set price based on dual variable of capacity constraints of certainty equivalent problem

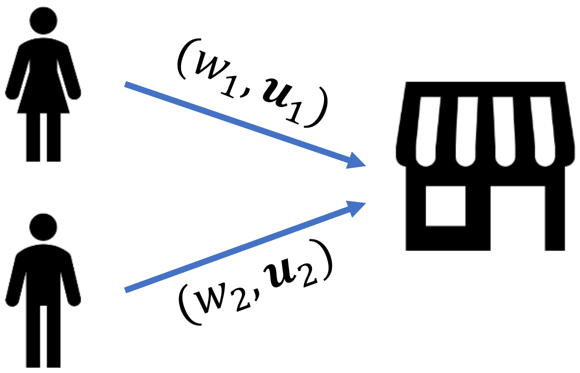
Users consume optimal bundle of goods

Update average remaining resource capacities

Theorem: Under i.i.d. budget and utility parameters with a discrete probability distribution and when good capacities are $O(n)$, Algorithm 1 achieves an expected regret of $R_n(\boldsymbol{\pi}) \leq O(\log(n))$ and expected constraint violation of $V_n(\boldsymbol{\pi}) \leq O(1)$

Primal algorithms are often computationally expensive and do not preserve user privacy

User parameters $(w, \boldsymbol{u})$ are revealed



Such algorithms require information on user parameters, which may not be known in practice

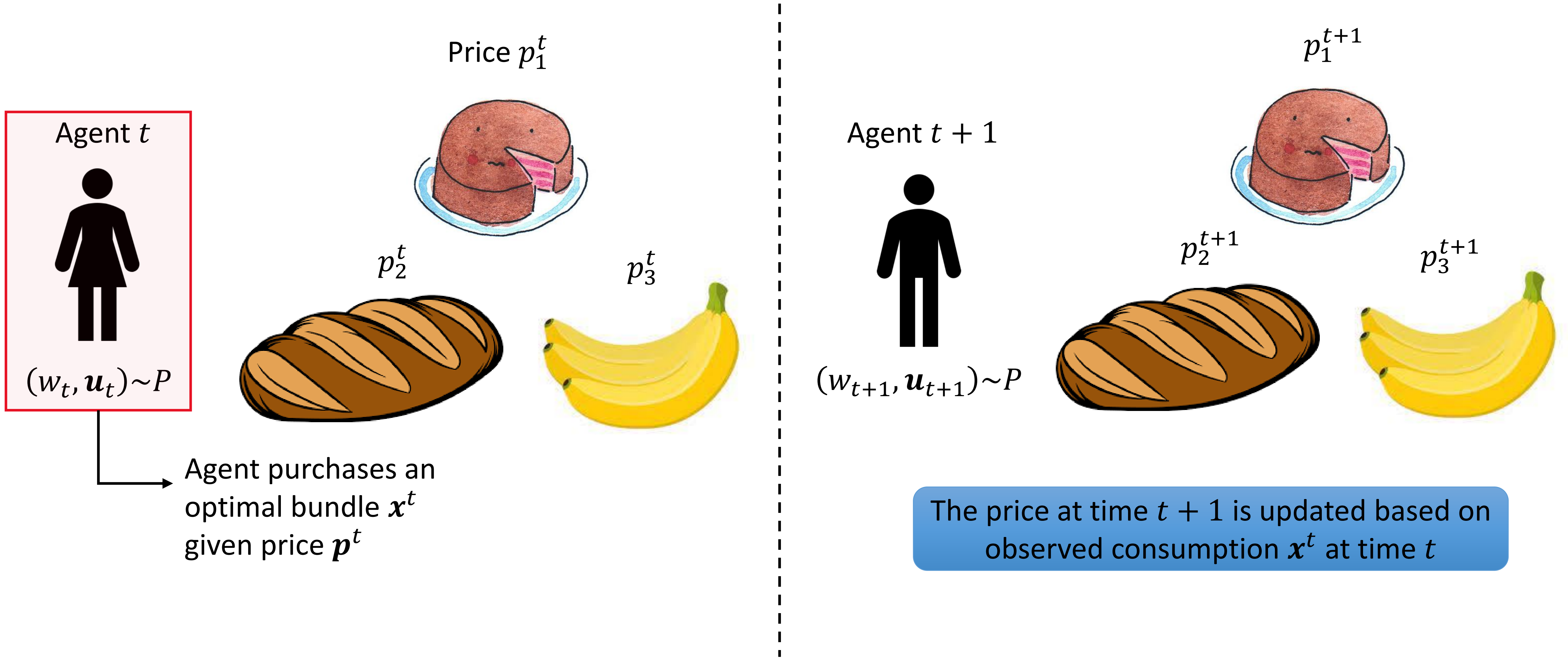
With parameters until user t arrives, we can solve the following primal problem

$$
\begin{aligned}
\max_{\mathbf{x}_i \in \mathbb{R}^m, \forall i \in [t]} \quad & \sum_{i=1}^{t} w_i \log \left(\sum_{j=1}^{m} u_{ij} x_{ij} \right) \\
\text{s.t.} \quad & \sum_{i=1}^{t} x_{ij} \leq \frac{t}{n} c_j, \quad \forall j \in [m] \\
& x_{ij} \geq 0, \quad \forall i \in [t], j \in [m]
\end{aligned}
$$

Prices can be set based on dual of capacity constraints

At each time instance, we solve a larger convex program, which may become computationally expensive in real time

We design a dual based algorithm, wherein users see prices at each time they arrive



Applying gradient descent to the dual of the social optimization problem motivates a natural algorithm

Dual of social optimization problem with Lagrange multiplier of the capacity constraints p_j

$$\min_{\mathbf{p}} \quad \sum_{t=1}^{n} w_t \log(w_t) - \sum_{t=1}^{n} w_t \log\left(\min_{j\in[m]} \frac{p_j}{u_{tj}}\right) + \sum_{j=1}^{m} p_j c_j - \sum_{t=1}^{n} w_t$$

Equivalent Sample Average Approximation (SAA) of Dual Problem

$$\min_{\mathbf{p}} \quad D_n(\mathbf{p}) = \sum_{j=1}^{m} p_j \frac{c_j}{n} + \frac{1}{n}\sum_{t=1}^{n}\left(w_t \log(w_t) - w_t \log(\min_{j\in[m]} \frac{p_j}{u_{tj}}) - w_t\right)$$

(Sub)-gradient descent of dual problem for each agent: $O(m)$ complexity of price update

$$\partial_{\mathbf{p}}\left(\sum_{j\in[m]} p_j \frac{c_j}{n} + w \log(w) - w \log\left(\min_{j\in[m]} \frac{p_j}{u_j}\right) - w\right)\Bigg|_{\mathbf{p}=\mathbf{p}^t} = \frac{1}{n}\mathbf{c} - \mathbf{x}_t$$

Difference between market share of each agent and goods purchased

We develop a revealed preference algorithm with sub-linear regret and constraint violation guarantees

Algorithm 2: Revealed Preference Algorithm for Online Fisher Markets

Input : Number of users n, Vector of good capacities per user $\mathbf{d} = \frac{\mathbf{c}}{n}$

Initialize $\mathbf{p}^1 > \mathbf{0}$;

for $t = 1, 2, ..., n$ **do**

  Phase I: ;

  User purchases an optimal bundle of goods $\mathbf{x}_t$ given the price $\mathbf{p}^t$;

  Phase II (Price Update): ;

  $\mathbf{p}^{t+1} \leftarrow \mathbf{p}^t - \gamma_t (\mathbf{d} - \mathbf{x}_t)$;

end

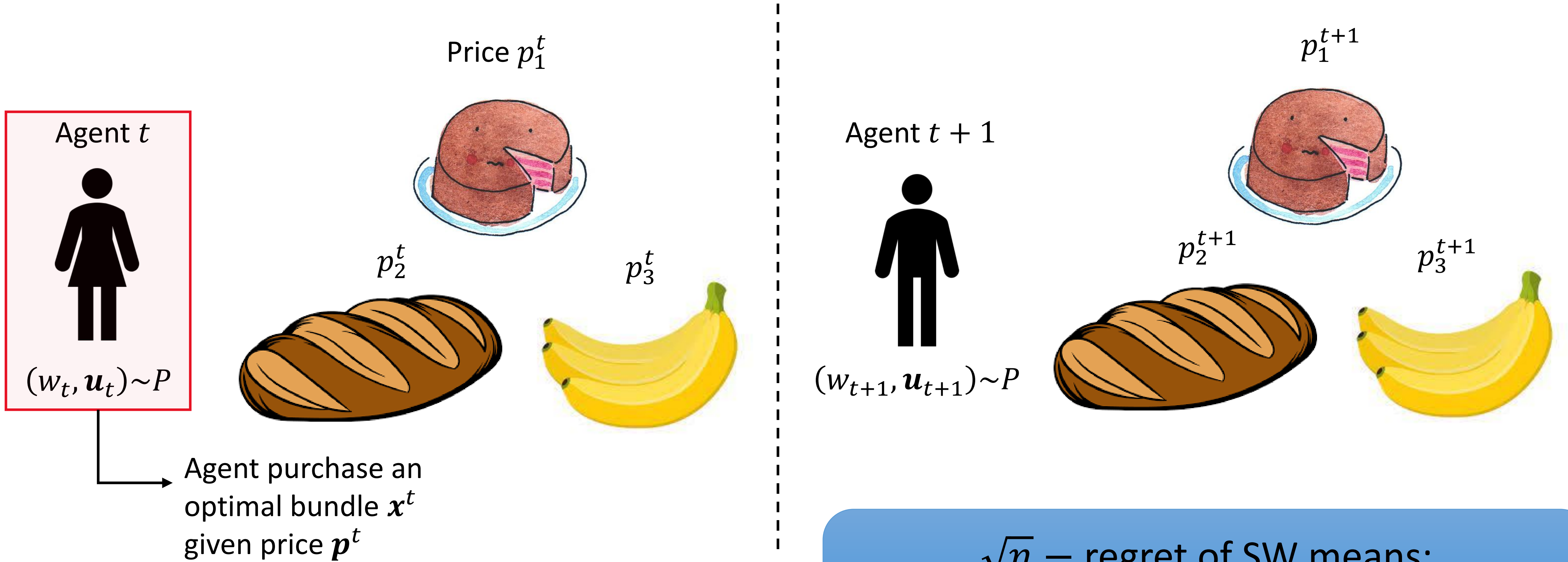
Difference between market share of each agent and goods purchased

Step-size: $O\left(\frac{1}{\sqrt{n}}\right)$

Only requires knowledge of user consumption (and not their budgets or utilities) to update prices

Theorem: Under i.i.d. budget and utility parameters with strictly positive support and when good capacities are $O(n)$, Algorithm 2 achieves an expected regret of $R_n(\boldsymbol{\pi}) \leq O(\sqrt{n})$ and expected constraint violation of $V_n(\boldsymbol{\pi}) \leq O(\sqrt{n})$, where n is the number of arriving users.

Again, the price of a good is increased if the arriving user purchase more than its market share of the good



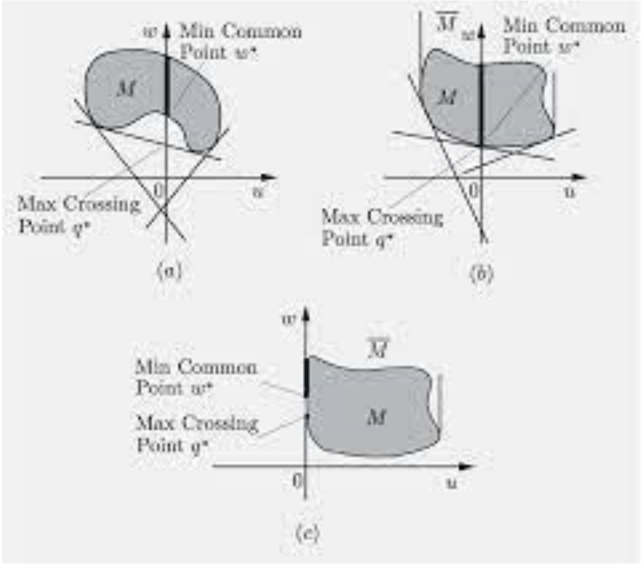
Increase Prices: $p_j^{t+1} > p_j^t$ if $x_j^{t+1} > \frac{c_j}{n}$

Decrease Prices: $p_j^{t+1} < p_j^t$ if $x_j^{t+1} < \frac{c_j}{n}$

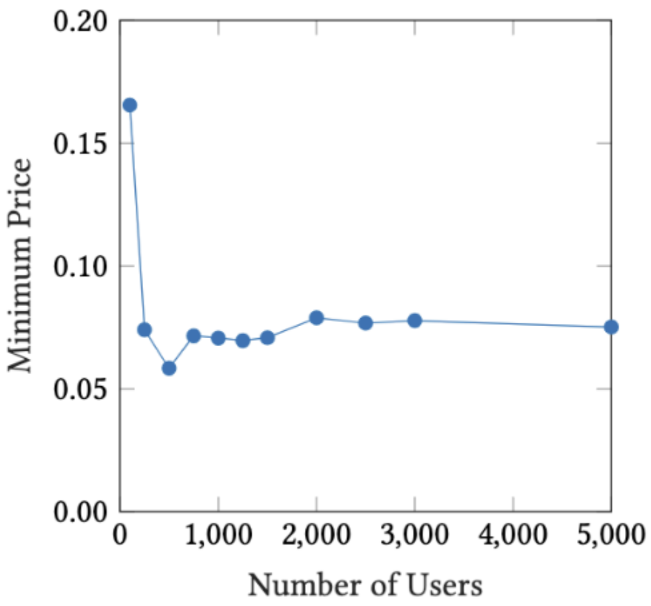
$\sqrt{n}$ – regret of SW means:

$$\frac{\text{SW optimal geometric mean}}{\text{SW geometric mean of online algorithm}} \leq e^{\frac{1}{\sqrt{n}}}$$

The regret and constraint violation guarantees follow from duality and a novel potential function argument



Use convex programming duality to establish the regret and constraint violation guarantees if the prices are strictly positive and bounded



Establish the positivity and boundedness of prices during the operation of Algorithm 2

1. Establish that the positivity of prices implies their boundedness

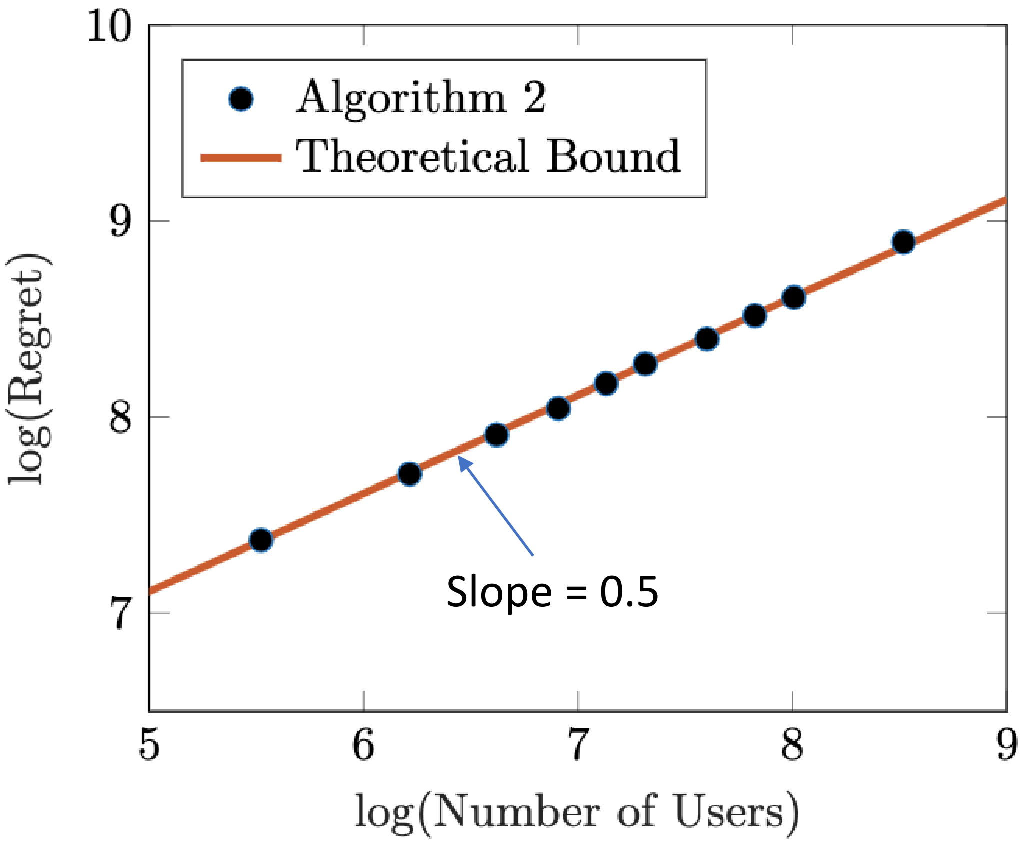
2. Use a potential function argument to show the positivity of prices

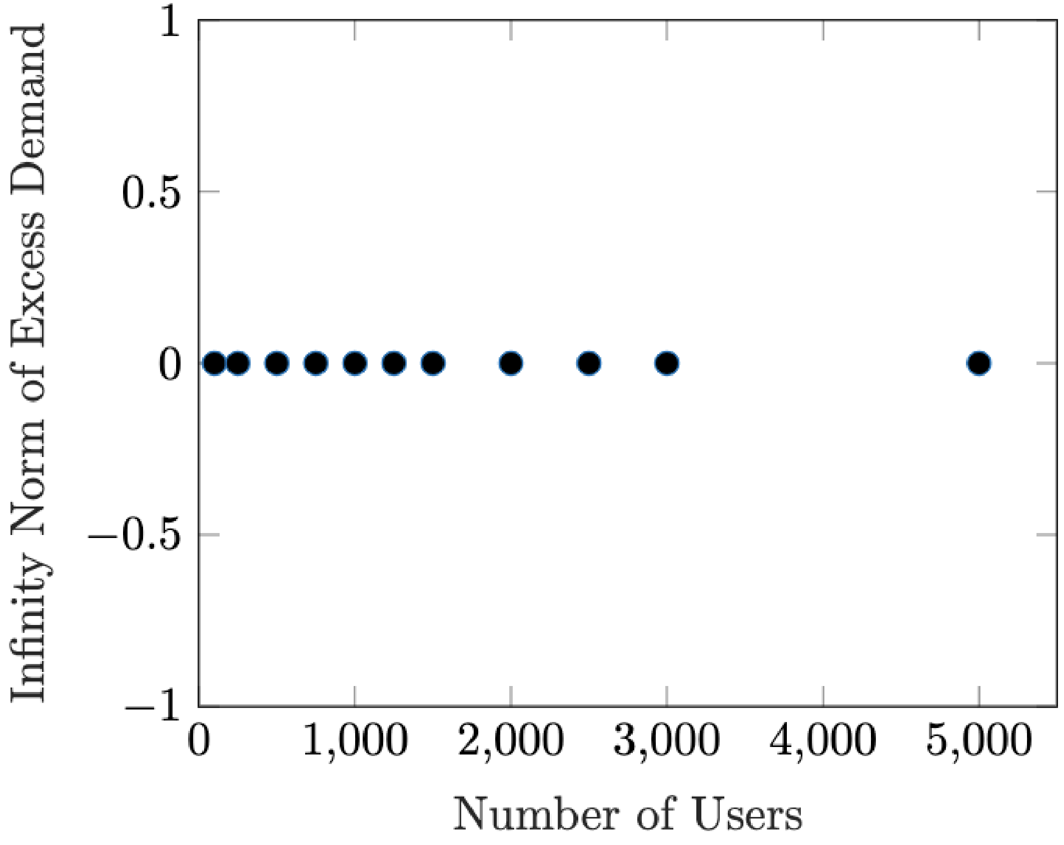
Potential Function

$$V_t = (\boldsymbol{p}^t) \cdot \boldsymbol{d}$$

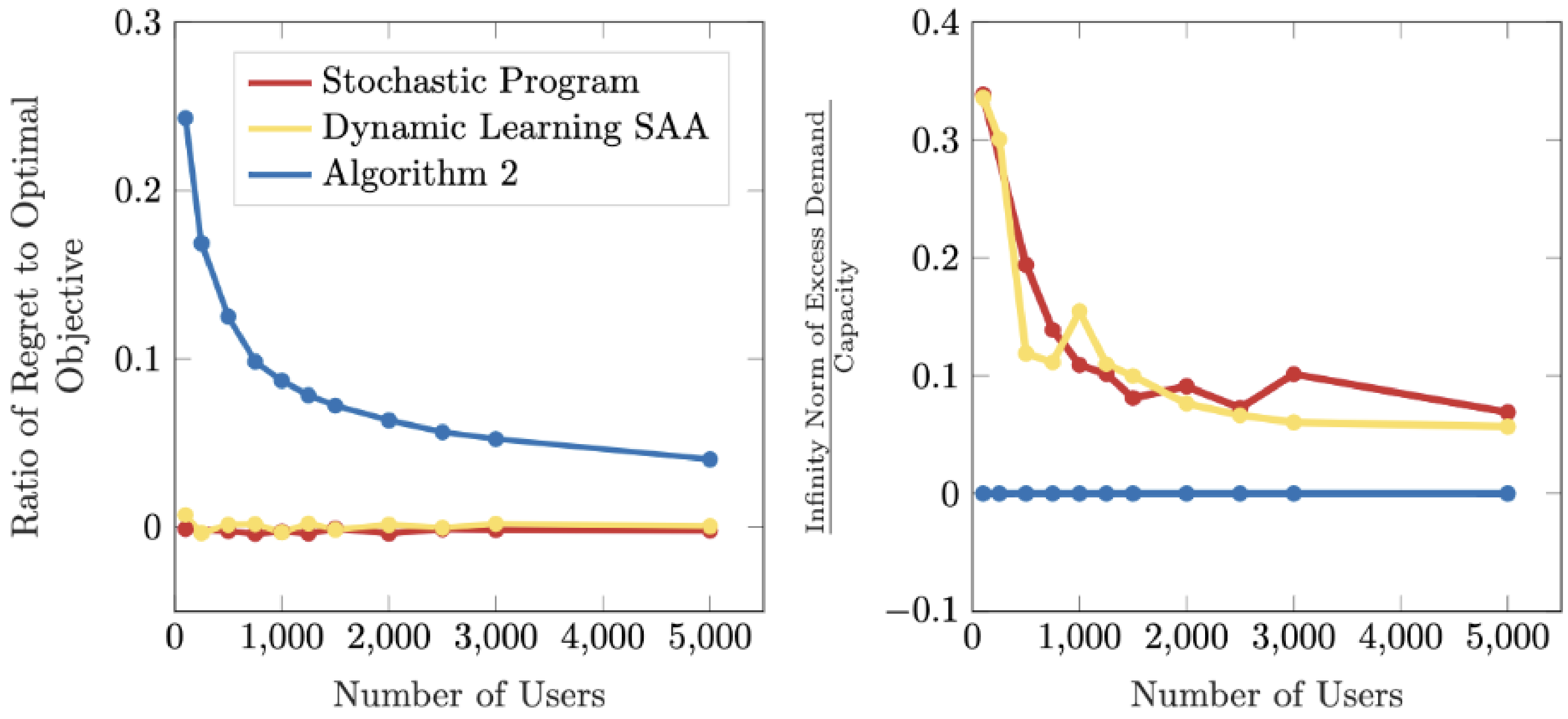
We show that this potential function is non-decreasing when the prices of all goods drops below a threshold, implying that the prices of some goods must increase in the subsequent iteration

Our numerical results verify the obtained theoretical guarantee





Our numerical results demonstrate a tradeoff between regret and constraint violation

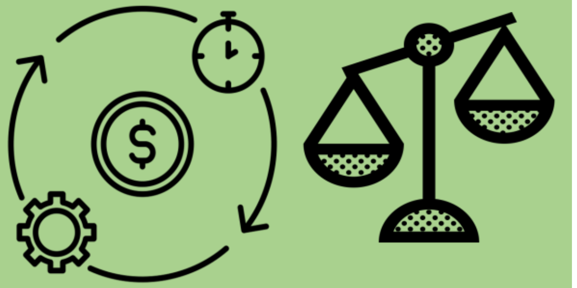


Organization

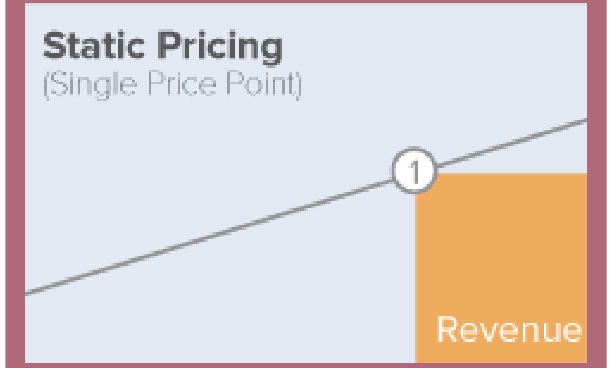
- Online Linear Programming
- OLP Extensions
- Online Equilibrium Pricing for Stochastic Fisher Market
- Conclusion/Takeaway

We study LP and Fisher markets in the online and incomplete information setting and develop algorithms with sub-linear regret guarantees

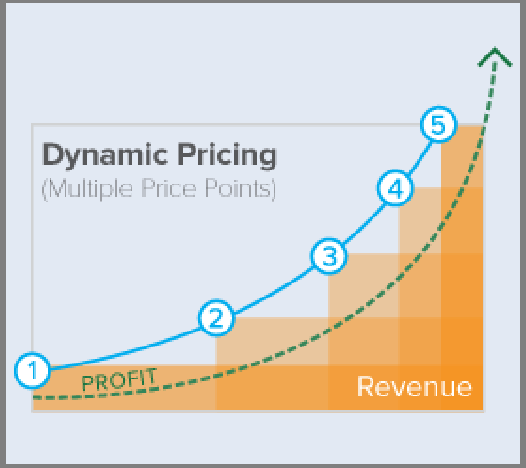
The weighted geometric average objective has both efficiency and fairness properties



Static equilibrium pricing approaches have performance limitations



We develop an adaptive expected equilibrium pricing algorithm with much improved performance



We develop a revealed preference algorithm with sub-linear regret and capacity violation

Jalota, Ye (2023), arXiv link: https://arxiv.org/abs/2205.00825

Overall Takeaways

It is possible to maker online decisions for quantitative decision models with performance guarantees close to that of the offline decision-making with complete information

Many open questions in these areas: nonstationary data, OLP with predictions, Distributionally Robust OLP, strategic players, truthful mechanism, etc…

- THANK YOU