Learning Pricing Mechanisms in Two-sided Markets

Stefano Leonardi (Sapienza University of Rome)

From Matchings to Markets, CIRM, 11-15 September 2023

Based on joint work with

- Paul Dütting, Federico Fusco, Philip Lazos, Rebecca Reiffenhäuser [2021]
- Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni and Federico Fusco [2021, 2023]

Bilateral Trade and Two-sided Markets

One vs Two sided markets

One sided market

- Buyers want to buy
- Mechanism and sellers *coincide*
- VCG-like optimal mechanisms

Two sided market

- Buyers want to buy
- Sellers want to sell
- The mechanism intermediates between the two parties
- Both buyers and sellers are strategic agents

One vs Two sided markets

One sided market

- Buyers want to buy
- Mechanism and sellers *coincide*
- VCG-like optimal mechanisms

Two sided market

- Buyers want to buy
- Sellers want to sell
- The mechanism intermediates between the two parties
- Both buyers and sellers are strategic agents

Seller s has an item, which buyer b wants.

Seller has value v_s for keeping it, buyer has value v_b for buying it.

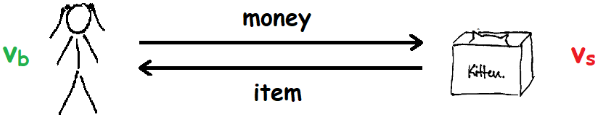
v_b and v_s are drawn from independent distributions D_b, D_s.

Social Welfare

Maximize social welfare!

$$SW = (v_b - v_s)\mathbb{I}_{trade} + v_s;$$

i.e., the value of the player holding the item in the end.



Seller s has an item, which buyer b wants.

Seller has value v_s for keeping it, buyer has value v_b for buying it.

v_b and v_s are drawn from independent distributions D_b, D_s.

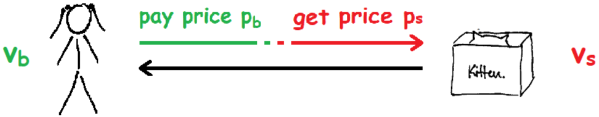
Social Welfare

Maximize social welfare!

$$SW = (v_b - v_s)\mathbb{1}_{trade} + v_s;$$

i.e., the value of the player holding the item in the end.

Warm up: Bilateral Trade



Seller s has an item, which buyer b wants.

Seller has value v_s for keeping it, buyer has value v_b for buying it.

v_b and v_s are drawn from independent distributions D_b, D_s.

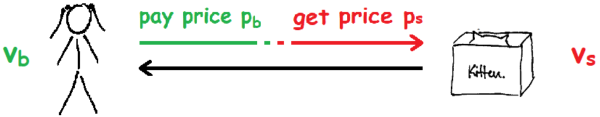
Social Welfare

Maximize social welfare!

$$SW = (v_b - v_s)\mathbb{I}_{trade} + v_s;$$

i.e., the value of the player holding the item in the end.

Warm up: Bilateral Trade



Seller s has an item, which buyer b wants.

Seller has value v_s for keeping it, buyer has value v_b for buying it.

v_b and v_s are drawn from independent distributions D_b, D_s.

Social Welfare

Maximize social welfare!

$$SW = (v_b - v_s)\mathbb{I}_{trade} + v_s;$$

i.e., the value of the player holding the item in the end.

$$u_s = \begin{cases} p_s - v_s & \text{if } s \text{ sells the item} \\ v_s & \text{otherwise} \end{cases} \qquad u_b = \begin{cases} v_b - p_b & \text{if } b \text{ buys the item} \\ 0 & \text{otherwise} \end{cases}$$

Individual Rationality (IR)

Participation does not have negative utility for the agents.

Incentive Compatibility / Truthfulness (IC)

Each agent maximizes his utility when reporting his true valuation, given the other agents' reports.

Budget Balance (BB)

The mechanism does not pay more money to the agents than it collects from them.

Wanted List: Mechanism Design

$$u_s = \begin{cases} p_s - v_s & \text{if } s \text{ sells the item} \\ v_s & \text{otherwise} \end{cases} \qquad u_b = \begin{cases} v_b - p_b & \text{if } b \text{ buys the item} \\ 0 & \text{otherwise} \end{cases}$$

Individual Rationality (IR)

Participation does not have negative utility for the agents.

Incentive Compatibility / Truthfulness (IC)

Each agent maximizes his utility when reporting his true valuation, given the other agents' reports.

Budget Balance (BB)

The mechanism does not pay more money to the agents than it collects from them.

$$u_s = \begin{cases} p_s - v_s & \text{if } s \text{ sells the item} \\ v_s & \text{otherwise} \end{cases} \qquad u_b = \begin{cases} v_b - p_b & \text{if } b \text{ buys the item} \\ 0 & \text{otherwise} \end{cases}$$

Individual Rationality (IR)

Participation does not have negative utility for the agents.

Incentive Compatibility / Truthfulness (IC)

Each agent maximizes his utility when reporting his true valuation, given the other agents' reports.

Budget Balance (BB)

The mechanism does not pay more money to the agents than it collects from them.

Wanted List: Mechanism Design

$$u_s = \begin{cases} p_s - v_s & \text{if } s \text{ sells the item} \\ v_s & \text{otherwise} \end{cases} \qquad u_b = \begin{cases} v_b - p_b & \text{if } b \text{ buys the item} \\ 0 & \text{otherwise} \end{cases}$$

Individual Rationality (IR)

Participation does not have negative utility for the agents.

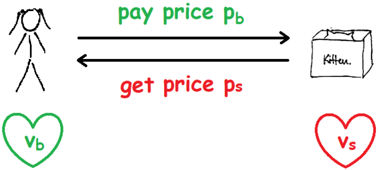
Incentive Compatibility / Truthfulness (IC)

Each agent maximizes his utility when reporting his true valuation, given the other agents' reports.

Budget Balance (BB)

The mechanism does not pay more money to the agents than it collects from them.

Bilateral Trade: The Problem with Truthfulness



Easy Algorithm: propose price $v_s \quad p \quad v_b$, trade if both accept.

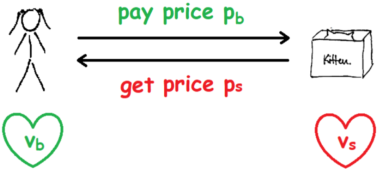
IC Problem: price of an agent cannot depend on his reported value

Obvious solution: set $p_s = v_b$ and $p_b = v_s$.

VCG is not budget-balanced

- Trade occurs only when $v_s \quad v_b$.
- Taking the item away from the seller costs v_s in social welfare: b pays v_s.
- Introducing s and his item yields social welfare v_b: s is paid v_b.

Bilateral Trade: The Problem with Truthfulness



Easy Algorithm: propose price $v_s \quad p \quad v_b$, trade if both accept.

IC Problem: price of an agent cannot depend on his reported value

Obvious solution: set $p_s = v_b$ and $p_b = v_s$.

VCG is not budget-balanced

- Trade occurs only when $v_s \quad v_b$.
- Taking the item away from the seller costs v_s in social welfare: b pays v_s.
- Introducing s and his item yields social welfare v_b: s is paid v_b.

Easy Algorithm: propose price $v_s \quad p \quad v_b$, trade if both accept.

IC Problem: price of an agent cannot depend on his reported value

Obvious solution: set $p_s = v_b$ and $p_b = v_s$.

VCG is not budget-balanced

- Trade occurs only when $v_s \quad v_b$.
- Taking the item away from the seller costs v_s in social welfare: b pays v_s.
- Introducing s and his item yields social welfare v_b: s is paid v_b.

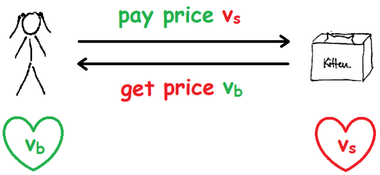
Easy Algorithm: propose price v_s p v_b, trade if both accept.

IC Problem: price of an agent cannot depend on his reported value

Obvious solution: set $p_s = v_b$ and $p_b = v_s$.

VCG is not budget-balanced

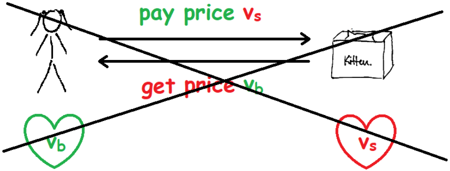
- Trade occurs only when v_s v_b.
- Taking the item away from the seller costs v_s in social welfare: b pays v_s.
- Introducing s and his item yields social welfare v_b: s is paid v_b.



Myerson-Satterthwaite Theorem

No mechanism for bilateral trade is IR, IC, BB and at the same time maximizes the social welfare even in the Bayesian setting.

[Myerson, Satterthwaite, 1983]



Myerson-Satterthwaite Theorem

No mechanism for bilateral trade is IR, IC, BB and at the same time maximizes the social welfare even in the Bayesian setting.

[Myerson, Satterthwaite, 1983]

Known Results with Access to Prior Information

Mechanisms with Full Prior Information

Bilateral Trade:
$0.72 = 0.71$ approximation.

[Cai,Zhu,STOC 2023], [Liu,Ren, Wang,STOC 2023]

Prior-Information Tradeoff:

- Knowing full distributions: usually unrealistic
- Knowing nothing: bad for mechanism's performance

Mechanism Design with limited information

- How good are mechanisms that use limited or no a-priori information?
- Two main possibilities to learn from the environment
    1. **Sampling:** Near optimal mechanisms that use a single sample from each prior distribution, and this is the minimum amount of information needed!
    2. **Strategic interaction:** Online learning of regret minimizing mechanisms for repeated bilateral trade

Myerson-Satterthwaite-Theorem

No mechanism for bilateral trade is IR, IC, BB and at the same time maximizes the social welfare.

[Myerson, Satterthwaite , 1983]

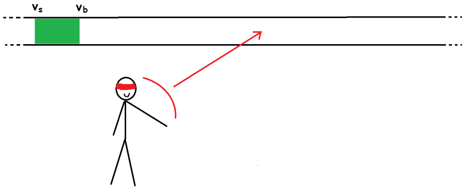
Impossibility Theorem

No IC, BB, IR mechanism without knowledge about the underlying distributions $D_b; D_s$ can achieve an -approximation to the optimal social welfare, for any $2\ \mathrm{R}_{>0}$.

Proof Glimpse: How to price?

Prices: between v_s and v_b, but

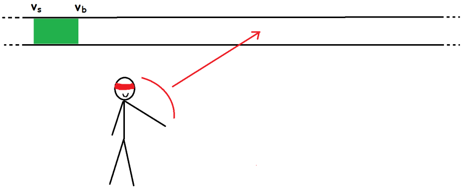
- p_s p_b - BB!
- can't use v_s for p_s or v_b for p_b - not IC!
- can't use v_s for p_b and v_b for p_s - not BB!



Find the *feasible* interval in an infinite range of numbers...

Prices: between v_s and v_b, but

- p_s $\quad$ p_b - BB!
- can't use v_s for p_s or v_b for p_b - not IC!
- can't use v_s for p_b and v_b for p_s - not BB!



Find the *feasible* interval in an infinite range of numbers...

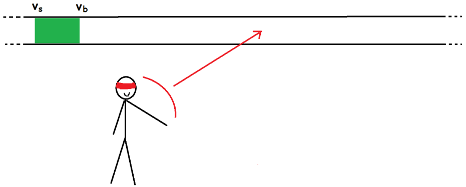
Imagine to be b, with v_b v_s

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all v_b^0 v_s
- To be individually rational and budget balanced p_b 2 $[v_s; \quad v_s]$

Imagine to be s, with v_s $\frac{v_b}{}$

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all v_s^0 $\frac{v_b}{}$
- To be individually rational and budget balanced p_s 2 $[\frac{v_b}{}; v_b]$

If $[\frac{v_b}{}; v_b]$ \ $[v_s; \quad v_s]$ = ; there is no hope to post the right price.



Proof Glimpse: How to price? (Deterministic Mechanism)

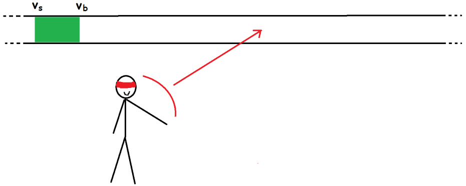
Imagine to be b, with v_b $\quad$ v_s

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all v_b^0 $\quad$ v_s
- To be individually rational and budget balanced p_b 2 $[v_s; \quad v_s]$

Imagine to be s, with v_s $\quad$ $\underline{v_b}$

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all v_s^0 $\quad$ $\underline{v_b}$
- To be individually rational and budget balanced p_s 2 $[\underline{v_b}; v_b]$

If $[\underline{v_b}; v_b] \setminus [v_s; \quad v_s] =$; there is no hope to post the right price.



Proof Glimpse: How to price? (Deterministic Mechanism)

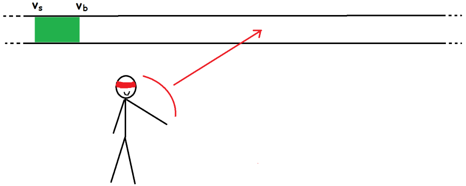
Imagine to be b, with v_b v_s

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all v_b^0 v_s
- To be individually rational and budget balanced p_b 2 $[v_s;$ $v_s]$

Imagine to be s, with v_s $\frac{v_b}{}$

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all v_s^0 $\frac{v_b}{}$
- To be individually rational and budget balanced p_s 2 $[\frac{v_b}{}; v_b]$

If $[\frac{v_b}{}; v_b]$ \ $[v_s;$ $v_s]$ = ; there is no hope to post the right price.



Proof Glimpse: How to price? (Deterministic Mechanism)

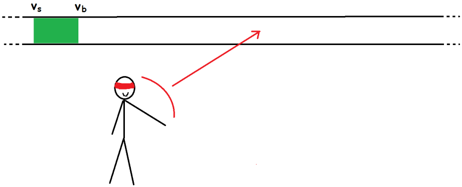
Imagine to be *b*, with $v_b \quad\quad v_s$

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all $v_b^0 \quad\quad v_s$
- To be individually rational and budget balanced p_b 2 $[v_s; \quad v_s]$

Imagine to be *s*, with $v_s \quad \frac{v_b}{}$

- To maintain the approximation, the mechanism has to trade
- To be incentive compatible, the price has to be the same for all $v_s^0 \quad \frac{v_b}{}$
- To be individually rational and budget balanced p_s 2 $[\frac{v_b}{}; v_b]$

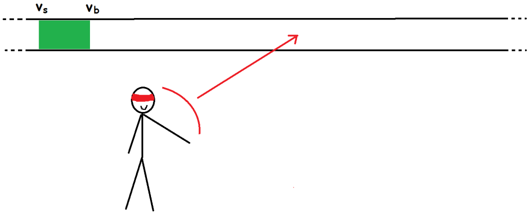
If $[\frac{v_b}{}; v_b] \setminus [v_s; \quad v_s] = ;$ there is no hope to post the right price.



Single Sample Mechanisms

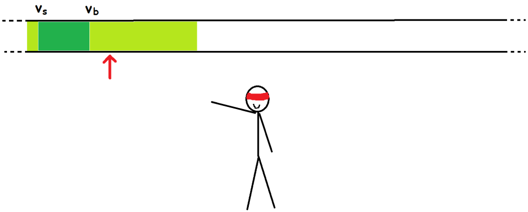
Paul Dütting, Federico Fusco, Philip Lazos, Stefano Leonardi, Rebecca Reiffenhäuser (2021,2022)

Single Sample from the Sellers Heals Impossibility



Idea: sample could give us a hint towards the feasible price interval!

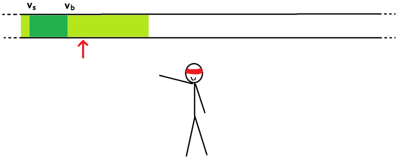
Single Sample from the Sellers Heals Impossibility



Idea: sample could give us a hint towards the feasible price interval!

Proposition: Sample Access Enables Constant Approximation

Posting v_s^0 from D_s as price is a 2 approximation for the bilateral trade problem



Solution: Draw a **sample** v_s^0 from D_s!

- v_s v_s^0 with probability 1=2
- $E[v_s^0] = E[v_s]$, and $E[v_s^0 j v_s$ $v_s^0]$ $2E[v_s]$.

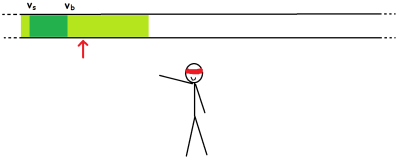
Intuition: For price $p = v_s^0$, s will accept w.pr. 1=2 and if b rejects, that's ok since the seller value (in exp.) is also good!

Note: Instead of a sample, any percentile works fine, too. E.g., same approximation for having the *median*. Blumrosen and Dobzinski [2014, 2016]

Minimum Prior Knowledge: Single Sample

Proposition: Sample Access Enables Constant Approximation

Posting v_s^0 from D_s as price is a 2 approximation for the bilateral trade problem



Solution: Draw a **sample** v_s^0 from D_s!

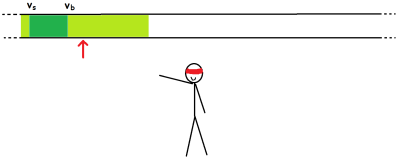
- $v_s \quad v_s^0$ with probability $\quad 1{=}2$
- $\mathrm{E}[v_s^0] = \mathrm{E}[v_s]$, and $\mathrm{E}[v_s^0 j v_s \quad v_s^0] \quad 2\mathrm{E}[v_s]$.

Intuition: For price $p = v_s^0$, s will accept w.pr. $\quad 1{=}2$ and if b rejects, that's ok since the seller value (in exp.) is also good!

Note: Instead of a sample, any percentile works fine, too. E.g., same approximation for having the *median*. Blumrosen and Dobzinski [2014, 2016]

Proposition: Sample Access Enables Constant Approximation

Posting v_s^0 from D_s as price is a 2 approximation for the bilateral trade problem



Solution: Draw a **sample** v_s^0 from D_s!

- v_s v_s^0 with probability 1=2
- $E[v_s^0] = E[v_s]$, and $E[v_s^0 j v_s$ $v_s^0]$ $2E[v_s]$.

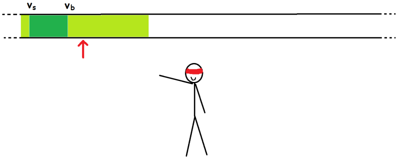
Intuition: For price $p = v_s^0$, s will accept w.pr. 1=2 and if b rejects, that's ok since the seller value (in exp.) is also good!

Note: Instead of a sample, any percentile works fine, too. E.g., same approximation for having the *median*. Blumrosen and Dobzinski [2014, 2016]

Minimum Prior Knowledge: Single Sample

Proposition: Sample Access Enables Constant Approximation

Posting v_s^0 from D_s as price is a 2 approximation for the bilateral trade problem



Solution: Draw a **sample** v_s^0 from D_s!

- v_s v_s^0 with probability 1=2
- $E[v_s^0] = E[v_s]$, and $E[v_s^0 j v_s$ $v_s^0]$ $2E[v_s]$.

Intuition: For price $p = v_s^0$, s will accept w.pr. 1=2 and if b rejects, that's ok since the seller value (in exp.) is also good!

Note: Instead of a sample, any percentile works fine, too. E.g., same approximation for having the *median*. Blumrosen and Dobzinski [2014, 2016]

Results with Limited Prior Information

Lower Bound of 2 **when only** $\{D_s\}_{s \in S}$ **known:**
no deterministic mechanism for bilateral trade achieves approximation better than 2 when only information about the seller distributions is used.

[**Blumrosen, Dobzinski, 2016**]

Limited Information seems to work quite well!

Theorem (Single-Sample 2-Lower Bound)

There exists no (deterministic), IR IC and BB mechanism that approximates social welfare better than to a factor of 2 and uses only a single sample from the seller distribution.

- Here, mechanisms have access to randomness (via the available sample). Therefore, this generalizes the existing deterministic 2 lower bound.
  [Dütting, Fusco, Lazos, Leonardi, Reiffenhäser, STOC 2021]
- This lower bound has recently been extended to randomized fixed price mechanisms that use one single sample.
  [Liu, Ren, Wang, STOC 2023]

Theorem (Black Box I)

Denote by the approximation guarantee of any one-sided IR, IC offline/online mechanism for maximizing social welfare for XOS valuations.

There exists a two-sided mechanism for XOS buyers and unit-supply sellers that is IR, IC, BB, uses a single sample from each seller and provides a

max f2 ; 3g

approximation to the optimal social welfare.

The two-sided mechanism inherits the offline/online properties of the one-sided mechanism on the buyer side and is offline on the seller side.

Theorem (Black Box I)

Denote by the approximation guarantee of any one-sided IR, IC offline/online mechanism for maximizing social welfare for XOS valuations.

There exists a two-sided mechanism for XOS buyers and unit-supply sellers that is IR, IC, BB, uses a single sample from each seller and provides a

$$\max\{2\ ;3\}$$

approximation to the optimal social welfare.

The two-sided mechanism inherits the offline/online properties of the one-sided mechanism on the buyer side and is offline on the seller side.

Theorem (Black Box I)

Denote by the approximation guarantee of any one-sided IR, IC offline/online mechanism for maximizing social welfare for XOS valuations.

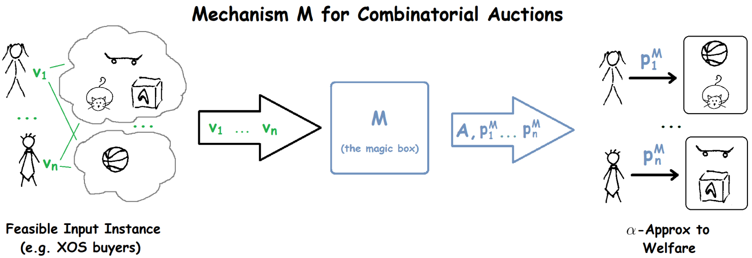
There exists a two-sided mechanism for XOS buyers and unit-supply sellers that is IR, IC, BB, uses a single sample from each seller and provides a

$$\max\{2\ ; 3\}$$

approximation to the optimal social welfare.

The two-sided mechanism inherits the offline/online properties of the one-sided mechanism on the buyer side and is offline on the seller side.

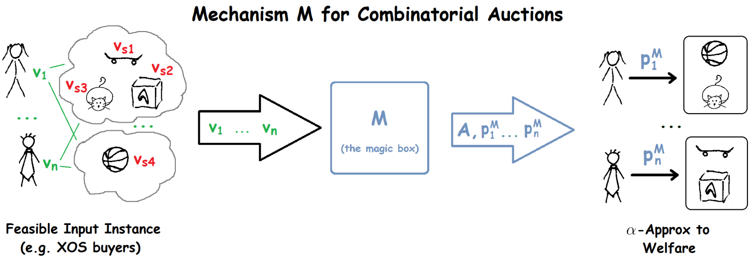
Blackbox Construction



- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \overline{T}} (a_{b,\overline{T}}(s) - v_s^0); \quad \text{where } \overline{T} = \text{argmax}_{T^* \subseteq T \cap \hat{S}} \left\{ v_b(T^*) - \sum_{s \in T^*} v_s^0 \right\}$$

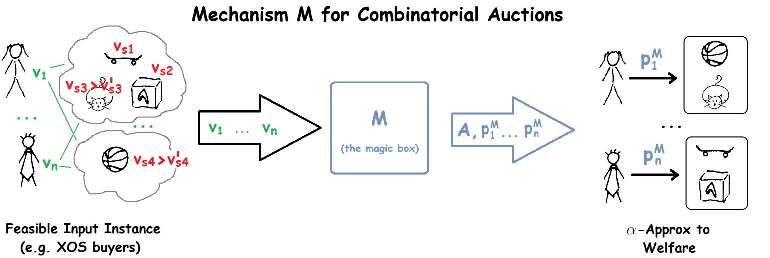
- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).



- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \bar{T}} (a_{b,\bar{T}}(s) - v_s^0); \quad \text{where } \bar{T} = \text{argmax}_{T^\star \subseteq T \cap \hat{S}} \left\{ v_b(T^\star) - \sum_{s \in T^\star} v_s^0 \right\}$$

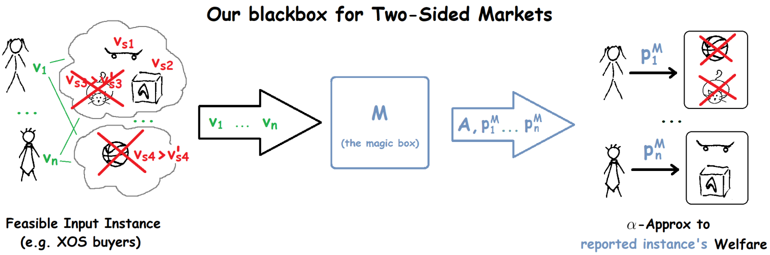
- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).



- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \bar{T}} (a_{b,\bar{T}}(s) - v_s^0); \quad \text{where } \bar{T} = \operatorname{argmax}_{T^\star \subseteq T \cap \hat{S}} \left\{ v_b(T^\star) - \sum_{s \in T^\star} v_s^0 \right\}$$

- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).

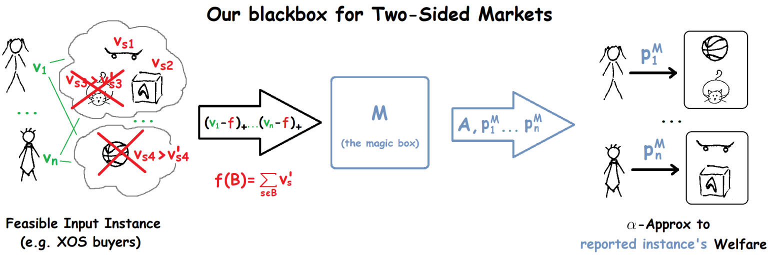


- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \overline{T}} (a_{b,\overline{T}}(s) - v_s^0); \quad \text{where } \overline{T} = \text{argmax}_{T^* \subseteq T \cap \hat{S}} \left\{ v_b(T^*) - \sum_{s \in T^*} v_s^0 \right\}$$

- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).

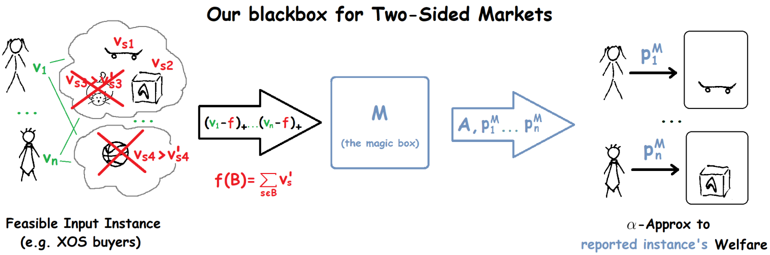
Blackbox Construction



- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \overline{T}} (a_{b;\overline{T}}(s) - v_s^0); \quad \text{where } \overline{T} = \text{argmax}_{T^\star \subseteq T \cap \hat{S}} \left\{ v_b(T^\star) - \sum_{s \in T^\star} v_s^0 \right\}$$

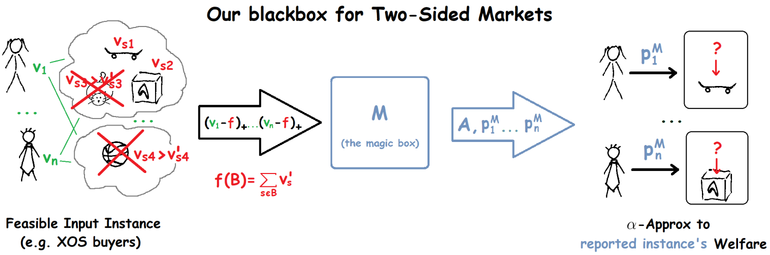
- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).



- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \bar{T}} (a_{b;\bar{T}}(s) - v_s^0); \quad \text{where } \bar{T} = \operatorname{argmax}_{T' \subseteq T \cap \hat{S}} \left\{ v_b(T') - \sum_{s \in T'} v_s^0 \right\}$$

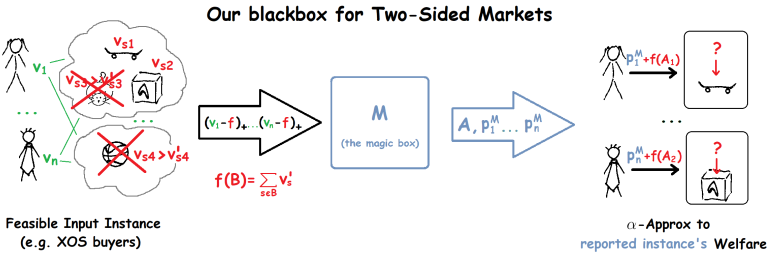
- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).



- Promise each seller s to get v^0_s for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \overline{T}} (a_{b;\overline{T}}(s) - v^0_s); \quad \text{where } \overline{T} = \text{argmax}_{T' \subseteq T \setminus \hat{S}} \left\{ v_b(T') - \sum_{s \in T'} v^0_s \right\}$$

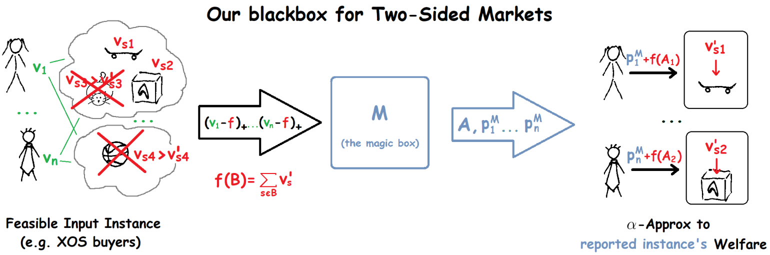
- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v^0_s$ for his bundle.
- Pay sellers the promised amount (if item was allocated).



- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s \in \overline{T}} (a_{b;\overline{T}}(s) - v_s^0); \quad \text{where } \overline{T} = \text{argmax}_{T' \subseteq T \cup \hat{S}} \left\{ v_b(T') - \sum_{s \in T'} v_s^0 \right\}$$

- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s \in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).



- Promise each seller s to get v_s^0 for his item, if accepts, add s to $\hat{S}$
- Report to M the modified buyers' valuations:

$$\hat{v}_b(T) = \sum_{s\in\overline{T}} (a_{b;\overline{T}}(s) - v_s^0); \quad \text{where } \overline{T} = \text{argmax}_{T' \subseteq T \setminus \hat{S}} \left\{ v_b(T \cup T') - \sum_{s\in T'} v_s^0 \right\}$$

- In addition to the mechanism price, charge each buyer b the sum of seller samples $\sum_{s\in A_b} v_s^0$ for his bundle.
- Pay sellers the promised amount (if item was allocated).

Results from Blackbox Application

For example, we obtain the following single-sample IR IC BB mechanisms:

- $2e$ **for max-weight matching**
  Online random order on the buyer side, offline sellers.
  Using: [**Reiffenhäuser, 2019.**]
- $O((\log\log m)^2)$ **for general XOS buyers**
  Offline on the buyer and seller side.
  Using: [**Assadi, Kesselheim, Singla, 2021**]

Regret Analysis of Bilateral Trade

Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, Stefano Leonardi (2021,2023)

Bilateral Trade

A *seller* and a *buyer* join an online platform to trade a good or a service

- The seller wants to sell at a price greater than some value S
- The buyer wants to buy at a price smaller than some value B
- The values are *private* information

Goal: Design an *efficient* mechanism to intermediate between the agents that is robust to *strategic behaviour*

A *seller* and a *buyer* join an online platform to trade a good or a service

- The seller wants to sell at a price greater than some value S
- The buyer wants to buy at a price smaller than some value B
- The values are *private* information

Goal: Design an *efficient* mechanism to intermediate between the agents that is robust to *strategic behaviour*

1. Individually rational (seller and buyer should not *lose* money in the trade)
2. Dominant strategy incentive compatibile (Agents incentivized to be *truthful*)
3. Strongly budget balanced (no *subsidy*, no *money draining*)
4. Economically efficient (no *trading opportunity* is lost)

Theorem (Myerson and Satterthwaite, 1981)

No mechanism satisfying 1 4 exists even in the Bayesian setting.

1. Individually rational (seller and buyer should not *lose* money in the trade)
2. Dominant strategy incentive compatibile (Agents incentivized to be *truthful*)
3. Strongly budget balanced (no *subsidy*, no *money draining*)
4. Economically efficient (no *trading opportunity* is lost)

Theorem (Myerson and Satterthwaite, 1981)

No mechanism satisfying 1   4 exists even in the Bayesian setting.

Desired Properties

1. Individually rational (seller and buyer should not *lose* money in the trade)
2. Dominant strategy incentive compatibile (Agents incentivized to be *truthful*)
3. Strongly budget balanced (no *subsidy*, no *money draining*)
4. Economically efficient (no *trading opportunity* is lost)

Theorem (Myerson and Satterthwaite, 1981)

No mechanism satisfying 1 4 exists even in the Bayesian setting.

Theorem (Colini-Baldeschi, de Keijzer, Leonardi and Turchetta, 2016)

All mechanism satisfying 1 3 are posted-price mechanism.

Posted price: the mechanism proposes a *price* without consulting the agents

One single price for buyer/seller - Strong Budget Balance

- Seller and buyer arrive with private valuations S and B
- The platform posts a price p, independently
- If S p B, then the trade happens

Theorem (Colini-Baldeschi, de Keijzer, Leonardi and Turchetta, 2016)

All mechanism satisfying 1 3 are posted-price mechanism.

Posted price: the mechanism proposes a *price* without consulting the agents

One single price for buyer/seller - Strong Budget Balance

- Seller and buyer arrive with private valuations S and B
- The platform posts a price p, independently
- If S p B, then the trade happens

For *Regret Minimization*, which is an additive approximation, we can restrict to the net increase in social welfare induced by the mechanism

Gain from trade

$$\mathrm{GFT}(p;S;B) = \Big(\underbrace{(B-p)}_{\text{buyer's net gain}} + \underbrace{(p-S)}_{\text{seller's net gain}}\Big)\underbrace{\mathbb{I}\{S \le p \le B\}}_{\text{trade happens}} = (B-S)\,\mathbb{I}\{S \le p \le B\}$$

Goal: maximize the total gain from trade over a sequence of trades without any *preliminary* information on buyers and sellers valuations

The Gain from Trade

For *Regret Minimization*, which is an additive approximation, we can restrict to the net increase in social welfare induced by the mechanism

Gain from trade

$$\mathrm{GFT}(p; S; B) = \underbrace{(B - p)}_{\text{buyer's net gain}} + \underbrace{(p - S)}_{\text{seller's net gain}} \underbrace{\mathbb{I}\{S \le p \le B\}}_{\text{trade happens}} = (B - S)\mathbb{I}\{S \le p \le B\}$$

Goal: maximize the total gain from trade over a sequence of trades without any *preliminary* information on buyers and sellers valuations

For $t = 1, 2, \dots$

1. Seller and buyer arrive with hidden valuations $S_t, B_t \in [0, 1]$
2. The platform posts a price $p_t \in [0, 1]$
3. The platform observes a *feedback signal* Z_t
4. The seller and the buyer leave

- Full feedback (direct revelation): $Z_t = (S_t, B_t)$
- Two-bits feedback (posted-price): $Z_t = \big(\mathbb{I}\{S_t \le p_t\}, \mathbb{I}\{p_t \le B_t\}\big)$
- One-bit feedback (minimal): $Z_t = \mathbb{I}\{S_t \le p_t \le B_t\}$

The price p_t is determined by $Z_1, \dots, Z_{t-1}$ and possibly by internal randomization

Different generation models of the sequence $(S_1, B_1), (S_2, B_2), \dots$ of valuations

- Adversarial setting: *unknown* deterministic sequence
- Stochasthic setting: *i.i.d.* random variables from an unknown joint distribution

For $t = 1, 2, \ldots$

1. Seller and buyer arrive with hidden valuations $S_t, B_t \in [0,1]$
2. The platform posts a price $p_t \in [0,1]$
3. The platform observes a *feedback signal* Z_t
4. The seller and the buyer leave

- Full feedback (direct revelation): $Z_t = (S_t, B_t)$
- Two-bits feedback (posted-price): $Z_t = \big(\mathbb{I}\{S_t \le p_t\}, \mathbb{I}\{p_t \le B_t\}\big)$
- One-bit feedback (minimal): $Z_t = \mathbb{I}\{S_t \le p_t \le B_t\}$

The price p_t is determined by $Z_1, \ldots, Z_{t-1}$ and possibly by internal randomization

Different generation models of the sequence $(S_1, B_1), (S_2, B_2), \ldots$ of valuations

- Adversarial setting: *unknown* deterministic sequence
- Stochasthic setting: *i.i.d.* random variables from an unknown joint distribution

For $t = 1, 2, \ldots$

1. Seller and buyer arrive with hidden valuations $S_t, B_t \in [0,1]$
2. The platform posts a price $p_t \in [0,1]$
3. The platform observes a *feedback signal* Z_t
4. The seller and the buyer leave

- Full feedback (direct revelation): $Z_t = (S_t, B_t)$
- Two-bits feedback (posted-price): $Z_t = \left(\mathbb{I}\{S_t \le p_t\}, \mathbb{I}\{p_t \le B_t\}\right)$
- One-bit feedback (minimal): $Z_t = \mathbb{I}\{S_t \le p_t \le B_t\}$

The price p_t is determined by $Z_1, \ldots, Z_{t-1}$ and possibly by internal randomization

Different generation models of the sequence $(S_1, B_1), (S_2, B_2), \ldots$ of valuations

- Adversarial setting: *unknown* deterministic sequence
- Stochasthic setting: *i.i.d.* random variables from an unknown joint distribution

For $t = 1, 2, \ldots$

1. Seller and buyer arrive with hidden valuations $S_t, B_t \in [0, 1]$
2. The platform posts a price $p_t \in [0, 1]$
3. The platform observes a *feedback signal* Z_t
4. The seller and the buyer leave

- Full feedback (direct revelation): $Z_t = (S_t, B_t)$
- Two-bits feedback (posted-price): $Z_t = \big(\mathbb{I}\{S_t \le p_t\}, \mathbb{I}\{p_t \le B_t\}\big)$
- One-bit feedback (minimal): $Z_t = \mathbb{I}\{S_t \le p_t \le B_t\}$

The price p_t is determined by $Z_1, \ldots, Z_{t-1}$ and possibly by internal randomization

Different generation models of the sequence $(S_1, B_1), (S_2, B_2), \ldots$ of valuations

- Adversarial setting: *unknown* deterministic sequence
- Stochasthic setting: *i.i.d.* random variables from an unknown joint distribution

Feedback models and generating process

For $t = 1, 2, \ldots$

1. Seller and buyer arrive with hidden valuations $S_t, B_t \in [0, 1]$
2. The platform posts a price $p_t \in [0, 1]$
3. The platform observes a *feedback signal* Z_t
4. The seller and the buyer leave

- Full feedback (direct revelation): $Z_t = (S_t, B_t)$
- Two-bits feedback (posted-price): $Z_t = \big(\mathbb{I}\{S_t \le p_t\}, \mathbb{I}\{p_t \le B_t\}\big)$
- One-bit feedback (minimal): $Z_t = \mathbb{I}\{S_t \le p_t \le B_t\}$

The price p_t is determined by $Z_1, \ldots, Z_{t-1}$ and possibly by internal randomization

Different generation models of the sequence $(S_1, B_1), (S_2, B_2), \ldots$ of valuations

- Adversarial setting: *unknown* deterministic sequence
- Stochasthic setting: *i.i.d.* random variables from an unknown joint distribution

Feedback models and generating process

For $t = 1, 2, \dots$

1. Seller and buyer arrive with hidden valuations $S_t, B_t \in [0, 1]$
2. The platform posts a price $p_t \in [0, 1]$
3. The platform observes a *feedback signal* Z_t
4. The seller and the buyer leave

- Full feedback (direct revelation): $Z_t = (S_t, B_t)$
- Two-bits feedback (posted-price): $Z_t = \big(\mathbb{I}\{S_t \le p_t\}, \mathbb{I}\{p_t \le B_t\}\big)$
- One-bit feedback (minimal): $Z_t = \mathbb{I}\{S_t \le p_t \le B_t\}$

The price p_t is determined by $Z_1, \dots, Z_{t-1}$ and possibly by internal randomization

Different generation models of the sequence $(S_1, B_1), (S_2, B_2), \dots$ of valuations

- Adversarial setting: *unknown* deterministic sequence
- Stochasthic setting: *i.i.d.* random variables from an unknown joint distribution

For $t = 1, 2, \ldots$

1. Seller and buyer arrive with hidden valuations $S_t, B_t \in [0, 1]$
2. The platform posts a price $p_t \in [0, 1]$
3. The platform observes a *feedback signal* Z_t
4. The seller and the buyer leave

- Full feedback (direct revelation): $Z_t = (S_t, B_t)$
- Two-bits feedback (posted-price): $Z_t = \big(\mathbb{I}\{S_t \le p_t\}, \mathbb{I}\{p_t \le B_t\}\big)$
- One-bit feedback (minimal): $Z_t = \mathbb{I}\{S_t \le p_t \le B_t\}$

The price p_t is determined by $Z_1, \ldots, Z_{t-1}$ and possibly by internal randomization

Different generation models of the sequence $(S_1, B_1), (S_2, B_2), \ldots$ of valuations

- Adversarial setting: *unknown* deterministic sequence
- Stochasthic setting: *i.i.d.* random variables from an unknown joint distribution

Regret minimization

Compete against best fixed-price strategy

$$R_T = \max_{p \in [0,1]} \mathbb{E}\left[\sum_{t=1}^{T} \mathrm{GFT}(p, S_t, B_t) - \sum_{t=1}^{T} \mathrm{GFT}(p_t, S_t, B_t) \right]$$

Our Contribution: Full characterization of the different *regret regimes* for different combinations of:

- *Feedback Models*
- *Sequence generation Models*

A linear lower bound on the regret

Given any randomized algorithm, we construct a deterministic sequence $(S_1, B_1), (S_2, B_2), \dots$ of valuations such that

1. The probability that $p_t \in [S_t, B_t]$ is at most $\frac{1}{2}$
2. There exists $p^* \in \bigcap_{t \geq 1} [S_t, B_t]$
3. $B_t - S_t \geq \frac{1 - \varepsilon}{2}$

This implies $R_T \geq \frac{1 - \varepsilon}{4} T = \Omega(T)$ even under *full feedback*

The sequence of valuations is inspired from the construction of Cantor's ternary set

A linear lower bound on the regret

Given any randomized algorithm, we construct a deterministic sequence $(S_1, B_1), (S_2, B_2), \ldots$ of valuations such that

1. The probability that $p_t \in [S_t, B_t]$ is at most $\frac{1}{2}$
2. There exists $p^* \in \bigcap_{t \geq 1} [S_t, B_t]$
3. $B_t - S_t \geq \frac{1-\varepsilon}{2}$

This implies $R_T \geq \frac{1-\varepsilon}{4} T = \Omega(T)$ even under *full feedback*

The sequence of valuations is inspired from the construction of Cantor's ternary set

A linear lower bound on the regret

Given any randomized algorithm, we construct a deterministic sequence $(S_1, B_1), (S_2, B_2), \ldots$ of valuations such that

1. The probability that $p_t \in [S_t, B_t]$ is at most $\frac{1}{2}$
2. There exists $p \in \bigcap_{t \geq 1} [S_t, B_t]$
3. $B_t - S_t \geq \frac{1}{2} - \varepsilon$

This implies $R_T \geq \frac{1 - \varepsilon}{4} T = \Omega(T)$ even under *full feedback*

The sequence of valuations is inspired from the construction of Cantor's ternary set

A linear lower bound on the regret

Given any randomized algorithm, we construct a deterministic sequence $(S_1; B_1); (S_2; B_2); : : :$ of valuations such that

1. The probability that $p_t \; 2 \; [S_t; B_t]$ is at most $\frac{1}{2}$
2. There exists $p \quad 2 \quad \bigcap_{t \quad 1} [S_t; B_t]$
3. $B_t \quad S_t \quad \frac{1 \quad ''}{2}$

This implies $R_T \quad \frac{1 \quad ''}{4} \quad T = \quad (T)$ even under *full feedback*

The sequence of valuations is inspired from the construction of Cantor's ternary set

Regret Regimes - Adversarial Setting

	Stochastic	Adversarial
Full		(T)
Two-bits		T

The sequence $(S_1, B_1), (S_2, B_2), \ldots$ is i.i.d. with fixed but unknown distribution

$$p^\star \in \arg\max_{p \in [0,1]} \mathbb{E}\left[\mathrm{GFT}(p)\right]$$

Full feedback

- (S_t, B_t) observed in each round, so algorithm can compute $\mathrm{GFT}(p; S_t, B_t)$, $\forall p$
- We can run *follow the best price* $p_t = \arg\max_{p \in [0,1]} \underbrace{\hat{\mathbb{E}}_t\, \mathrm{GFT}(p)}_{\text{empirical GFT}}$
- For any sequence $(S_1, B_1), \ldots, (S_{t-1}, B_{t-1})$ of valuations
  $\mathbb{E}[\mathrm{GFT}(p^\star)] - \mathbb{E}[\mathrm{GFT}(p_t)] \le 2 \max_{p \in [0,1]} \left| \mathbb{E}[\mathrm{GFT}(p)] - \hat{\mathbb{E}}_t\, \mathrm{GFT}(p) \right|$
- We can use *uniform convergence* over the class
  $\{\mathrm{GFT}(p; \cdot, \cdot) : 0 \le p \le 1\}$ of real-valued functions on $[0,1]^2$
- This gives $R_T = O\left(\sqrt{T \ln T}\right)$

The sequence $(S_1, B_1), (S_2, B_2), \dots$ is i.i.d. with fixed but unknown distribution

$$p^* \in \arg\max_{p \in [0,1]} \mathbb{E}\left[\mathrm{GFT}(p)\right]$$

Full feedback

- (S_t, B_t) observed in each round, so algorithm can compute $\mathrm{GFT}(p, S_t, B_t), \ \forall p$
- We can run *follow the best price* $p_t = \arg\max_{p \in [0,1]} \underbrace{\widehat{\mathbb{E}}_t \, \mathrm{GFT}(p)}_{\text{empirical GFT}}$
- For any sequence $(S_1, B_1), \dots, (S_{t-1}, B_{t-1})$ of valuations $\mathbb{E}[\mathrm{GFT}(p^*)] - \mathbb{E}[\mathrm{GFT}(p_t)] \le 2 \max_{p \in [0,1]} \left| \mathbb{E}[\mathrm{GFT}(p)] - \widehat{\mathbb{E}}_t \, \mathrm{GFT}(p) \right|$
- We can use *uniform convergence* over the class $\{\mathrm{GFT}(p, \cdot, \cdot) : 0 \le p \le 1\}$ of real-valued functions on $[0,1]^2$
- This gives $R_T = O\left(\sqrt{T \ln T}\right)$

The sequence $(S_1, B_1), (S_2, B_2), \dots$ is i.i.d. with fixed but unknown distribution

$$p^* \in \arg\max_{p \in [0,1]} \mathbb{E}[\mathrm{GFT}(p)]$$

Full feedback

- (S_t, B_t) observed in each round, so algorithm can compute $\mathrm{GFT}(p, S_t, B_t),\ \forall p$
- We can run *follow the best price* $p_t = \arg\max_{p \in [0,1]} \underbrace{\hat{\mathbb{E}}_t \, \mathrm{GFT}(p)}_{\text{empirical GFT}}$
- For any sequence $(S_1, B_1), \dots, (S_{t-1}, B_{t-1})$ of valuations $\mathbb{E}[\mathrm{GFT}(p^*)] - \mathbb{E}[\mathrm{GFT}(p_t)] \le 2 \max_{p \in [0,1]} \left| \mathbb{E}[\mathrm{GFT}(p)] - \hat{\mathbb{E}}_t \, \mathrm{GFT}(p) \right|$
- We can use *uniform convergence* over the class $\{\mathrm{GFT}(p, \cdot, \cdot) : 0 \le p \le 1\}$ of real-valued functions on $[0,1]^2$
- This gives $R_T = O\left(\sqrt{T \ln T}\right)$

The sequence $(S_1, B_1), (S_2, B_2), \ldots$ is i.i.d. with fixed but unknown distribution

$$p^* \in \arg\max_{p \in [0,1]} \mathbb{E}[\mathrm{GFT}(p)]$$

Full feedback

- (S_t, B_t) observed in each round, so algorithm can compute $\mathrm{GFT}(p, S_t, B_t), \ \forall p$
- We can run *follow the best price* $p_t = \arg\max_{p \in [0,1]} \underbrace{\widehat{\mathbb{E}}_t \, \mathrm{GFT}(p)}_{\text{empirical GFT}}$
- For any sequence $(S_1, B_1), \ldots, (S_{t-1}, B_{t-1})$ of valuations
$$\mathbb{E}[\mathrm{GFT}(p^*)] - \mathbb{E}[\mathrm{GFT}(p_t)] \le 2 \max_{p \in [0,1]} \left| \mathbb{E}[\mathrm{GFT}(p)] - \widehat{\mathbb{E}}_t \, \mathrm{GFT}(p) \right|$$
- We can use *uniform convergence* over the class $\{\mathrm{GFT}(p, \cdot, \cdot) : 0 \le p \le 1\}$ of real-valued functions on $[0,1]^2$
- This gives $R_T = O\left(\sqrt{T \ln T}\right)$

Stochastic valuations and full feedback

The sequence $(S_1, B_1), (S_2, B_2), \dots$ is i.i.d. with fixed but unknown distribution

$$p^* \in \operatorname*{arg\,max}_{p \in [0,1]} \mathbb{E}[\mathrm{GFT}(p)]$$

Full feedback

- (S_t, B_t) observed in each round, so algorithm can compute $\mathrm{GFT}(p, S_t, B_t), \ \forall p$
- We can run *follow the best price* $p_t = \operatorname*{arg\,max}_{p \in [0,1]} \underbrace{\widehat{\mathbb{E}}_t \, \mathrm{GFT}(p)}_{\text{empirical GFT}}$
- For any sequence $(S_1, B_1), \dots, (S_{t-1}, B_{t-1})$ of valuations
  $\mathbb{E}[\mathrm{GFT}(p^*)] - \mathbb{E}[\mathrm{GFT}(p_t)] \le 2 \max_{p \in [0,1]} \left| \mathbb{E}[\mathrm{GFT}(p)] - \widehat{\mathbb{E}}_t \, \mathrm{GFT}(p) \right|$
- We can use *uniform convergence* over the class $\{\mathrm{GFT}(p, \cdot, \cdot) : 0 \le p \le 1\}$ of real-valued functions on $[0,1]^2$
- This gives $R_T = O\left(\sqrt{T \ln T}\right)$

Stochastic valuations and full feedback

The sequence $(S_1, B_1), (S_2, B_2), \ldots$ is i.i.d. with fixed but unknown distribution

$$p^* \in \arg\max_{p \in [0,1]} \mathbb{E}[\mathrm{GFT}(p)]$$

Full feedback

- (S_t, B_t) observed in each round, so algorithm can compute $\mathrm{GFT}(p; S_t, B_t), \ \forall p$
- We can run *follow the best price* $p_t = \arg\max_{p \in [0,1]} \underbrace{\widehat{\mathbb{E}}_t \, \mathrm{GFT}(p)}_{\text{empirical GFT}}$
- For any sequence $(S_1, B_1), \ldots, (S_{t-1}, B_{t-1})$ of valuations $\mathbb{E}[\mathrm{GFT}(p^*)] - \mathbb{E}[\mathrm{GFT}(p_t)] \le 2 \max_{p \in [0,1]} \left| \mathbb{E}[\mathrm{GFT}(p)] - \widehat{\mathbb{E}}_t \, \mathrm{GFT}(p) \right|$
- We can use *uniform convergence* over the class $\{\mathrm{GFT}(p; \cdot, \cdot) : 0 \le p \le 1\}$ of real-valued functions on $[0,1]^2$
- This gives $R_T = O\left(\sqrt{T \ln T}\right)$

Lower bound

$R_T = \sqrt{T}$ even when seller and buyer valuations are *independent* with *bounded densities*.

Proof: reduction from experts

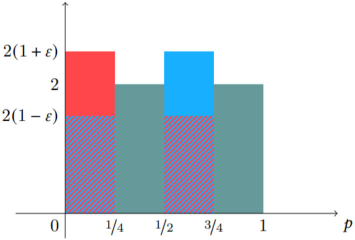


Figure: Seller distribution in green, buyer distribution either red or blue

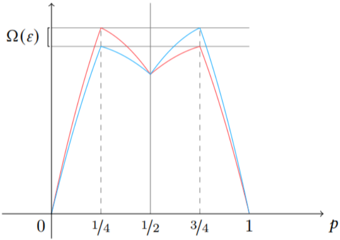


Figure: Gains from trade of the two scenarios

Regret Regimes - Direct Revelation Mechanisms

	Stochastic	Adversarial
Full	$(\frac{p}{T})$	(T)
Two-bits		T

Two-bits feedback - Assumptions needed!

- $\mathbb{I}\{S_t \quad p_t\}$ and $\mathbb{I}\{p_t \quad B_t\}$ observed at each step
- There exist independent valuations $(S; B)$ with unbounded density such that $R_T = \quad (T)$. *Needle in a haystack* phenomenon
- There exist correlated valuations $(S; B)$ with bounded density such that $R_T = \quad (T)$. *Indistinguishable* distributions

Two-bits feedback - Assumptions needed!

- $\mathbb{I}\{S_t \quad p_t\}$ and $\mathbb{I}\{p_t \quad B_t\}$ observed at each step
- There exist independent valuations $(S; B)$ with unbounded density such that $R_T = \quad (T)$. *Needle in a haystack* phenomenon
- There exist correlated valuations $(S; B)$ with bounded density such that $R_T = \quad (T)$. *Indistinguishable* distributions

Regret Regimes - Posted Price Mechanisms

	Stochastic				Adversarial
	iid	+indep.	+bounded	+indep. & bounded	adv
Full	$O(\sqrt{T})$	$\sqrt{T}$	$\sqrt{T}$	$(\sqrt{T})$	(T)
Two-bits	T	(T)	(T)		T

$$
\begin{aligned}
\mathbb{E}[\mathrm{GFT}(p)] &= (B - S)\,\mathbb{I}\{S \le p \le B\} \\
&= \int_0^p \mathbb{P}\left[S \le x;\ p \le B\right] \mathrm{d}x + \int_p^1 \mathbb{P}\left[S \le p;\ x \le B\right] \mathrm{d}x
\end{aligned}
$$

Using *independence*:

$$
\mathbb{E}[\mathrm{GFT}(p)] = \mathbb{P}(p \le B) \int_0^p \mathbb{P}(S \le x)\,\mathrm{d}x + \mathbb{P}(S \le p) \int_p^1 \mathbb{P}(x \le B)\,\mathrm{d}x
$$

If U is uniform over $[0;1]$ and independent on S and B, then

$$
\mathbb{E}[\mathrm{GFT}(p)] = \mathbb{P}(p \le B)\mathbb{P}(S \le U \le p) + \mathbb{P}(S \le p)\mathbb{P}(p \le U \le B)
$$

It is possible to estimate the blue terms via sampling!

$$\mathbb{E}[\mathrm{GFT}(p)] = (B - S)\mathbb{I}\{S \le p \le B\}$$
$$= \int_0^p \mathbb{P}\left(S \le x;\ p \le B\right) dx + \int_p^1 \mathbb{P}\left(S \le p;\ x \le B\right) dx$$

Using *independence*:

$$\mathbb{E}[\mathrm{GFT}(p)] = \mathbb{P}(p \le B)\int_0^p \mathbb{P}(S \le x)\, dx + \mathbb{P}(S \le p)\int_p^1 \mathbb{P}(x \le B)\, dx$$

If U is uniform over $[0; 1]$ and independent on S and B, then

$$\mathbb{E}[\mathrm{GFT}(p)] = \mathbb{P}(p \le B)\mathbb{P}(S \le U \le p) + \mathbb{P}(S \le p)\mathbb{P}(p \le U \le B)$$

It is possible to estimate the blue terms via sampling!

1. ε-grid of prices over $[0, 1]$
2. Post random prices to estimate $\mathbb{P}(S \le U \le p)$ and $\mathbb{P}(p \le U \le B)$ for each p in the ε-grid
3. Bandits: built a uniform grid and play a bandit algorithm on the points of the grid replacing the blue terms with their *approximations*.

Grid *OK* because the *boundedness* assumption on S and B distributions makes $\mathbb{E}[\mathrm{GFT}(p)]$ *Lipschitz*

$$R_T \lesssim \underbrace{\frac{1}{\varepsilon^2}\ln\frac{1}{\varepsilon}}_{\text{random expl.}} + \underbrace{\varepsilon T}_{\text{approx. error}} + \underbrace{\sqrt{T/\varepsilon}}_{\text{bandit regret}} = O\left(T^{2/3}\ln T\right) \text{ for } \varepsilon = T^{-1/3}$$

1. ε-grid of prices over $[0,1]$
2. Post random prices to estimate $\mathbb{P}(S \le U \le p)$ and $\mathbb{P}(p \le U \le B)$ for each p in the ε-grid
3. Bandits: built a uniform grid and play a bandit algorithm on the points of the grid replacing the blue terms with their *approximations*.

Grid *OK* because the *boundedness* assumption on S and B distributions makes $\mathbb{E}[\mathrm{GFT}(p)]$ *Lipschitz*

$$R_T \lesssim \underbrace{\frac{1}{\varepsilon^2}\ln\frac{1}{\varepsilon}}_{\text{random expl.}} + \underbrace{\varepsilon T}_{\text{approx. error}} + \underbrace{\sqrt{T/\varepsilon}}_{\text{bandit regret}} = O\left(T^{2/3}\ln T\right) \text{ for } \varepsilon = T^{-1/3}$$

1. ε-grid of prices over $[0,1]$
2. Post random prices to estimate $\mathbb{P}(S \le U \le p)$ and $\mathbb{P}(p \le U \le B)$ for each p in the ε-grid
3. Bandits: built a uniform grid and play a bandit algorithm on the points of the grid replacing the blue terms with their *approximations*.

Grid *OK* because the *boundedness* assumption on S and B distributions makes $\mathbb{E}[\mathrm{GFT}(p)]$ *Lipschitz*

$$R_T \lesssim \underbrace{\frac{1}{\varepsilon^2}\ln\frac{1}{\varepsilon}}_{\text{random expl.}} + \underbrace{\varepsilon T}_{\text{approx. error}} + \underbrace{\sqrt{T/\varepsilon}}_{\text{bandit regret}} = O\left(T^{2/3}\ln T\right) \text{ for } \varepsilon = T^{-1/3}$$

1. ε-grid of prices over $[0, 1]$
2. Post random prices to estimate $\mathbb{P}(S \le U \le p)$ and $\mathbb{P}(p \le U \le B)$ for each p in the ε-grid
3. Bandits: built a uniform grid and play a bandit algorithm on the points of the grid replacing the blue terms with their *approximations*.

Grid *OK* because the *boundedness* assumption on S and B distributions makes $\mathbb{E}[\mathrm{GFT}(p)]$ *Lipschitz*

$$R_T \lesssim \underbrace{\frac{1}{\varepsilon^2}\ln\frac{1}{\varepsilon}}_{\text{random expl.}} + \underbrace{\varepsilon T}_{\text{approx. error}} + \underbrace{\sqrt{T/\varepsilon}}_{\text{bandit regret}} = O(T^{2/3}\ln T) \quad \text{for } \varepsilon = T^{-1/3}$$

1. ε-grid of prices over $[0,1]$
2. Post random prices to estimate $\mathbb{P}(S \le U \le p)$ and $\mathbb{P}(p \le U \le B)$ for each p in the ε-grid
3. Bandits: built a uniform grid and play a bandit algorithm on the points of the grid replacing the blue terms with their *approximations*.

Grid *OK* because the *boundedness* assumption on S and B distributions makes $\mathbb{E}[\mathrm{GFT}(p)]$ *Lipschitz*

$$R_T \le \underbrace{\frac{1}{\varepsilon^2}\ln\frac{1}{\varepsilon}}_{\text{random expl.}} + \underbrace{\varepsilon T}_{\text{approx. error}} + \underbrace{\sqrt{T/\varepsilon}}_{\text{bandit regret}} = O\left(T^{2/3}\ln T\right) \quad \text{for } \varepsilon = T^{-1/3}$$

1. ε-grid of prices over $[0,1]$
2. Post random prices to estimate $\mathbb{P}(S \le p)$ and $\mathbb{P}(p \le B)$ for each p in the ε-grid
3. Bandits: built a uniform grid and play a bandit algorithm on the points of the grid replacing the blue terms with their *approximations*.

Grid *OK* because the *boundedness* assumption on S and B distributions makes $\mathbb{E}[\mathrm{GFT}(p)]$ *Lipschitz*

$$R_T \lesssim \underbrace{\frac{1}{\varepsilon^2}\ln\frac{1}{\varepsilon}}_{\text{random expl.}} + \underbrace{\varepsilon T}_{\text{approx. error}} + \underbrace{\sqrt{T/\varepsilon}}_{\text{bandit regret}} = O\left(T^{2/3}\ln T\right) \quad \text{for } \varepsilon = T^{-1/3}$$

	Stochastic				Adversarial
	iid	+indep.	+bounded	+indep. & bounded	adv
Full	$O(\sqrt{T})$	$\sqrt{T}$	$\sqrt{T}$	$(\sqrt{T})$	(T)
Two-bits	T	(T)	(T)	$(T^{2=3})$	T

Stochastic model

- Posted price mechanisms require buyer/seller independence and smooth distributions.
- Two-bit feedback is required.

Weaker adversarial models

- Smooth Adversary

Definition (Haghtalab, Roughgarden, 2021)

Let X be a domain supporting a uniform distribution . A measure on X is said to be -smooth if for all measurable subsets A X, we have $(A) \quad \frac{(A)}{\quad}$.

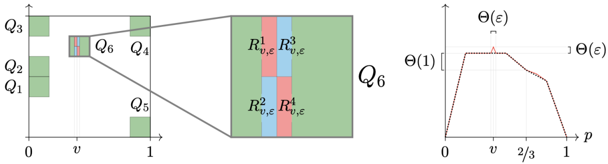
Weak Budget Balance Mechanisms (two prices) are strictly superior to Strong Budget Balance Mechanisms (one price) in the Smooth Adversarial setting

	Full Feedback	Two-bit Feedback	One-bit Feedback
Single Price	$\tilde{O}(\sqrt{T})$	(T)	(T)
Two Prices	$(\sqrt{T})$	$(T^{3/4})$	$\tilde{O}(T^{3/4})$

[Cesa-Bianchi, Cesari, Colomboni, Fusco, Leonardi, COLT 2023]

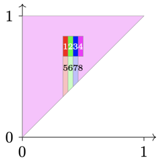
A family of -smooth adversaries:

- The valuations $(S_t; B_t)$ are drawn i.i.d. according to a fixed distribution, obliviously of the actions of the learner.
- We build this family of distributions by suitable perturbations over a base distribution, whose support is given by the union of the six squares $Q_1; : : : ; Q_6$.



The $(T^{3=4})$ Lower Bound

- We consider a perturbation such that the sequence of seller/buyer evaluations $(S; B), (S_1; B_1), (S_2; B_2), :::$ is i.i.d. ands it is -smooth, for all $^{1=9}$.
- Finding the best of K arms requires to pull each arm $(\frac{1}{2})$ time with no guarantee of any reward. Alternatively, the algorithm can exploit a random arm at each step by incurring a regret .
- The lower bound is $\min(\frac{K}{2}; \quad T)$ to obtaining with $K = T^{1=4}$ and $= T^{1=4}$ minimax regret $(T^{3=4})$



Conclusions:

- Limited information, even one single sample, can be sufficient to provide near optimal approximation mechanisms.
- Efficient learning of mechanisms through repeated interaction with the agents is possible in some settings.

Open problems:

- learning in strategic interaction between agents with carry over between rounds.
- Budget balance along the whole time horizon.
- Fair division of the gain from trade between buyers and sellers across time.

Thanks!