An α-regret Analysis of Adversarial Bilateral Trade

Yossi Azar[1] Amos Fiat[1] **Federico Fusco**[2]

NeurIPS'22 - New Orleans

[1]Tel Aviv University [2]Sapienza University of Rome

Bilateral Trade

We study the *bilateral trade* problem in the framework of online learning

Learning Protocol of Sequential Bilateral Trade

```
for time t = 1, 2, ... do
    a new seller/buyer pair arrives with (hidden) valuations (s_t, b_t) ∈ [0,1]^2
    the learner posts prices p_t, q_t ∈ [0,1] with p_t ≤ q_t
    the learner receives a (hidden) reward GFT_t(p_t, q_t)
    a feedback z_t is revealed
```

Gain from trade

$$\mathrm{GFT}_t(p_t, q_t) = \Big(\underbrace{(b_t - q_t)}_{\text{buyer's gain}} + \underbrace{(q_t - p_t)}_{\text{platform's gain}} + \underbrace{(p_t - s_t)}_{\text{seller's gain}} \Big) \underbrace{\mathbb{I}\{s_t \le p_t \le q_t \le b_t\}}_{\text{trade happens}}$$

$$= (b_t - s_t)\mathbb{I}\{s_t \le p_t \le q_t \le b_t\}$$

Bilateral Trade

We study the *bilateral trade* problem in the framework of online learning

Learning Protocol of Sequential Bilateral Trade

for time $t = 1, 2, \ldots$ **do**
    a new seller/buyer pair arrives with (hidden) valuations $(s_t, b_t) \in [0,1]^2$
    the learner posts prices $p_t, q_t \in [0,1]$ with $p_t \le q_t$
    the learner receives a (hidden) reward $\mathrm{GFT}_t(p_t, q_t)$
    a feedback z_t is revealed

Gain from trade

$$\mathrm{GFT}_t(p_t, q_t) = \Big(\underbrace{(b_t - q_t)}_{\text{buyer's gain}} + \underbrace{(q_t - p_t)}_{\text{platform's gain}} + \underbrace{(p_t - s_t)}_{\text{seller's gain}} \Big) \underbrace{\mathbb{I}\{s_t \le p_t \le q_t \le b_t\}}_{\text{trade happens}}$$

$$= (b_t - s_t)\mathbb{I}\{s_t \le p_t \le q_t \le b_t\}$$

Bilateral Trade

We study the *bilateral trade* problem in the framework of online learning

Learning Protocol of Sequential Bilateral Trade

1: **for** time $t = 1, 2, \ldots$ **do**
2: a new seller/buyer pair arrives with (hidden) valuations $(s_t, b_t) \in [0,1]^2$
3: the learner posts prices $p_t, q_t \in [0,1]$ with $p_t \leq q_t$
4: the learner receives a (hidden) reward $\mathrm{GFT}_t(p_t, q_t)$
5: a feedback z_t is revealed

Gain from trade

$$\mathrm{GFT}_t(p_t,q_t) = \Big(\underbrace{(b_t - q_t)}_{\text{buyer's gain}} + \underbrace{(q_t - p_t)}_{\text{platform's gain}} + \underbrace{(p_t - s_t)}_{\text{seller's gain}}\Big)\underbrace{\mathbb{I}\{s_t \leq p_t \leq q_t \leq b_t\}}_{\text{trade happens}}$$
$$= (b_t - s_t)\mathbb{I}\{s_t \leq p_t \leq q_t \leq b_t\}$$

Feedback models and regret

Feedback models

- Full feedback (direct revelation): $z_t = (s_t, b_t)$
- Two-bit feedback (posted-price): $z_t = \big(\mathbb{I}\{s_t \le p_t\}, \mathbb{I}\{q_t \le b_t\}\big)$
- One-bit feedback (minimal): $z_t = \mathbb{I}\{s_t \le p_t \le q_t \le b_t\}$

Prices

- Single Price (strong budget balance): $p_t = q_t$
- Two Prices (budget balance): $p_t \le q_t$

Adversaries

- Stochasthic setting: *i.i.d.* valuations ✓ [Cesa-Bianchi et al. 2021]
- Adversarial setting: *any (oblivious) sequence*

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

Feedback models and regret

Feedback models

- Full feedback (direct revelation): $z_t = (s_t, b_t)$
- Two-bit feedback (posted-price): $z_t = \big(\mathbb{I}\{s_t \le p_t\}, \mathbb{I}\{q_t \le b_t\}\big)$
- One-bit feedback (minimal): $z_t = \mathbb{I}\{s_t \le p_t \le q_t \le b_t\}$

Prices

- Single Price (strong budget balance): $p_t = q_t$
- Two Prices (budget balance): $p_t \le q_t$

Adversaries

- Stochasthic setting: *i.i.d.* valuations ✓ [Cesa-Bianchi et al. 2021]
- Adversarial setting: *any* (oblivious) sequence

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

Feedback models and regret

Feedback models

- Full feedback (direct revelation): $z_t = (s_t, b_t)$
- Two-bit feedback (posted-price): $z_t = \big(\mathbb{I}\{s_t \le p_t\}, \mathbb{I}\{q_t \le b_t\}\big)$
- One-bit feedback (minimal): $z_t = \mathbb{I}\{s_t \le p_t \le q_t \le b_t\}$

Prices

- Single Price (strong budget balance): $p_t = q_t$
- Two Prices (budget balance): $p_t \le q_t$

Adversaries

- Stochasthic setting: *i.i.d.* valuations ✓ [Cesa-Bianchi et al. 2021]
- Adversarial setting: *any* (oblivious) sequence

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

Feedback models and regret

Feedback models

- Full feedback (direct revelation): $z_t = (s_t, b_t)$
- Two-bit feedback (posted-price): $z_t = \big(\mathbb{I}\{s_t \le p_t\}, \mathbb{I}\{q_t \le b_t\}\big)$
- One-bit feedback (minimal): $z_t = \mathbb{I}\{s_t \le p_t \le q_t \le b_t\}$

Prices

- Single Price (strong budget balance): $p_t = q_t$
- Two Prices (budget balance): $p_t \le q_t$

Adversaries

- Stochasthic setting: *i.i.d.* valuations ✓ [Cesa-Bianchi et al. 2021]
- Adversarial setting: *any* (oblivious) sequence ✗ [Cesa-Bianchi et al. 2021]

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

Feedback models and regret

Feedback models

- Full feedback (direct revelation): $z_t = (s_t, b_t)$
- Two-bit feedback (posted-price): $z_t = \big(\mathbb{I}\{s_t \le p_t\}, \mathbb{I}\{q_t \le b_t\}\big)$
- One-bit feedback (minimal): $z_t = \mathbb{I}\{s_t \le p_t \le q_t \le b_t\}$

Prices

- Single Price (strong budget balance): $p_t = q_t$
- Two Prices (budget balance): $p_t \le q_t$

Adversaries

- Stochasthic setting: *i.i.d.* valuations ✓ [Cesa-Bianchi et al. 2021]
- Adversarial setting: *any* (oblivious) sequence ✓ [This work]

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

The soft spot: $\alpha = 2$

Theorem

For any $\alpha \in [1, 2)$, there is a linear lower bound on the α-regret achievable, even with full feedback and posting two different prices.

	Full Feedback	Two-bit feedback	One-bit feedback
Single price	$O(\sqrt{T})$	$\Omega(T)$	
Two prices	$\Omega(\sqrt{T})$	$\Omega(T^{2/3})$	$O(T^{3/4})$

Table: Summary of 2-regret results in various settings.

- Upper bound with 1-bit feedback: "Magic" Estimator
- Lower bounds: Partial monitoring and feedback structure

The soft spot: $\alpha = 2$

Theorem

For any $\alpha \in [1, 2)$, there is a linear lower bound on the α-regret achievable, even with full feedback and posting two different prices.

	Full Feedback	Two-bit feedback	One-bit feedback
Single price	$O(\sqrt{T})$	$\Omega(T)$	
Two prices	$\Omega(\sqrt{T})$	$\Omega(T^{2/3})$	$O(T^{3/4})$

Table: Summary of 2-regret results in various settings.

- Upper bound with 1-bit feedback: "Magic" Estimator
- Lower bounds: Partial monitoring and feedback structure

Thank you!

See you in New Orleans