# Tight Regret Bounds for Fixed-Price Bilateral Trade

**Houshuang Chen**
Shanghai Jiao Tong University, China

**Yaonan Jin**
Huawei's Taylor Lab, Beijing, China

**Pinyan Lu**
Laboratory of Interdisciplinary Research of Computation and Economics (SUFE), Shanghai University of Finance and Economics, China
Huawei's Taylor Lab, Beijing, China

**Chihao Zhang**
Shanghai Jiao Tong University, China

## Abstract

We examine fixed-price mechanisms in bilateral trade through the lens of regret minimization. Our main results are twofold. (i) For independent values, a near-optimal $\widetilde{\Theta}(T^{2/3})$ tight bound for Global Budget Balance fixed-price mechanisms with two-bit/one-bit feedback. (ii) For correlated/adversarial values, a near-optimal $\Omega(T^{3/4})$ lower bound for Global Budget Balance fixed-price mechanisms with two-bit/one-bit feedback, which improves the best known $\Omega(T^{5/7})$ lower bound obtained in the work [3] and, up to polylogarithmic factors, matches the $\widetilde{\mathcal{O}}(T^{3/4})$ upper bound obtained in the same work. Our work in combination with the previous works [5, 6, 2, 3] (essentially) gives a thorough understanding of regret minimization for fixed-price bilateral trade.

En route, we have developed two technical ingredients that might be of independent interest: (i) A novel algorithmic paradigm, called *fractal elimination*, to address one-bit feedback and independent values. (ii) A new *lower-bound construction* with novel proof techniques, to address the Global Budget Balance constraint and correlated values.

**2012 ACM Subject Classification** Theory of computation → Algorithmic mechanism design; Theory of computation → Computational pricing and auctions; Theory of computation → Query learning; Theory of computation → Regret bounds

**Keywords and phrases** bilateral trade, online learning, regret minimization, budget balance

**Digital Object Identifier** 10.4230/LIPIcs.ICALP.2026.57

**Category** Track A: Algorithms, Complexity and Games

**Related Version** *Full Version*: https://arxiv.org/abs/2504.04349 [7]

**Funding** *Chihao Zhang*: Huawei TC20240329014.

## 1 Introduction

We address a classic problem in Mechanism Design, maximizing *economic efficiency* in repeated bilateral trade: In each round $t \in [T]$, a (new) seller and a (new) buyer seek to trade an indivisible item, which has value $S^t$ to the seller and value $B^t$ to the buyer. There are two standard metrics of economic efficiency:
1. Gains from Trade, defined as $\mathsf{GFT} = \sum_{t\in[T]} \mathsf{GFT}^t = \sum_{t\in[T]}(B^t - S^t) \cdot Z^t$.
2. Social Welfare, defined as $\mathsf{SW} = \sum_{t\in[T]} \mathsf{SW}^t = \sum_{t\in[T]}(B^t \cdot Z^t + S^t \cdot (1 - Z^t))$.

Here, $Z^t = Z^t(S^t, B^t) \in \{0, 1\}$ denotes the trade outcome – either a success or a failure.

As is standard in Mechanism Design, there are three models for generating values $(S^t, B^t)_{t\in[T]}$, listed below from most to least general.

- **Adversarial Values:** An (oblivious) adversary determines an (arbitrary) $2T$-dimensional $[0,1]^{2T}$- supported joint distribution $\mathcal{D}$; then, the values $(S^t, B^t)_{t\in[T]}$ across all rounds are drawn from it.
- **Correlated Values:** An adversary determines an (arbitrary) two-dimensional $[0,1]^2$-supported joint distribution $\mathcal{D}$; then, the values $(S^t, B^t)$ in each round $t \in [T]$ are drawn i.i.d. from it.
- **Independent Values:** This is identical to "correlated values", except that $\mathcal{D}$ is further required to be a product distribution $\mathcal{D}_S \bigotimes \mathcal{D}_B$, making all the $2T$ values $(S^t, B^t)_{t\in[T]}$ mutually independent.

For reference, the "independent values" model is arguably the most canonical model, dating back to seminal works [13, 12], while the other two models are more general and have also received growing attention recently [4, 5, 6, 2, 3, 8].

In this work, we focus on fixed-price mechanisms from the perspective of *(additive) regret minimization*, by which Social Welfare and Gains from Trade are interchangeable since their gap $\mathsf{SW} - \mathsf{GFT} = \sum_{t\in[T]} S^t$ is mechanism-independent. Without loss of generality, we adopt Gains from Trade for our presentation. Such a mechanism posts two possibly randomized prices $(P^t, Q^t)$ in each round $t \in [T]$, one for the seller $P^t$ and one for the buyer $Q^t$. A trade occurs if both agents accept their respective prices: $Z^t = \mathbb{1}[S^t \le P^t \wedge Q^t \le B^t] \in \{0,1\}$. This yields the Gains from Trade $\mathsf{GFT}^t = (B^t - S^t)\cdot Z^t$ and the *profit* $\mathsf{Profit}^t = (Q^t - P^t)\cdot Z^t$. Throughout this work, we use the term *mechanisms* to refer to *fixed-price mechanisms* when no ambiguity arises.

Moreover, economic viability requires the Budget Balance (BB) constraint, which prevents the mechanism from subsidizing the agents. Regarding the BB constraint, prior works have examined two notions, a stricter *Local Budget Balance (LBB)* constraint [12] and a looser *Global Budget Balance (GBB)* constraint [3].

- Local Budget Balance: This can be further differentiated into a stricter *Strong Budget Balance (SBB)* constraint and a looser *Weak Budget Balance (WBB)* constraint.
  - Strong Budget Balance: $\mathsf{Profit}^t = 0$ (or essentially $P^t = Q^t$), $\forall t \in [T]$.
  - Weak Budget Balance: $\mathsf{Profit}^t \ge 0$ (or essentially $P^t \le Q^t$), $\forall t \in [T]$.
- Global Budget Balance: $\sum_{t\in[T]} \mathsf{Profit}^t \ge 0$ ex-post.
  This requires *nonnegative total profit* $\sum_{t\in[T]} \mathsf{Profit}^t \ge 0$, almost surely.

As is usual in Online Optimization, the design and analysis of a mechanism depend on the underlying *feedback model* – the trade information revealed at the end of each round $t \in [T]$. Prior works [5, 6, 2, 3] have studied two feedback models, the more informative *full feedback* and the less informative *partial feedback*.

- **Full Feedback:** This reveals both agents' values $(S^t, B^t) \in [0,1]^2$.
- **Partial Feedback:** This can be further differentiated into the more informative *two-bit feedback* (i.e., both agents' intentions to trade) and the less informative *one-bit feedback* (i.e., the trade's success/failure outcome).
  - **Two-Bit Feedback:** $(X^t, Y^t) = (\mathbb{1}[S^t \le P^t],\ \mathbb{1}[Q^t \le B^t]) \in \{0,1\}^2$.
  - **One-Bit Feedback:** $Z^t = \mathbb{1}[S^t \le P^t \wedge Q^t \le B^t] \equiv X^t \wedge Y^t \in \{0,1\}$.

For full feedback, the picture is essentially complete. Prior work [5, 3] shows $\widetilde{\Theta}(T^{1/2})$ regret for independent/correlated values, and GBB mechanisms also achieve $\widetilde{\Theta}(T^{1/2})$ regret in the adversarial setting, whereas LBB mechanisms suffer linear regret [5, 2]. We therefore

**Table 1** Regret bounds of *GBB partial-feedback* fixed-price mechanisms, up to polylogarithmic factors.

| | Global Budget Balance | |
|---|---|---|
| | Previous Lower/Upper Bounds | Current Tight Bounds |
| Independent | $[T^{1/2}, T^{3/4}]$ [Thm 12] [3] | $T^{2/3}$ [Thms 11 & 1] |
| Correlated | $[T^{5/7}, T^{3/4}]$ [3] | $T^{3/4}$ [Thm 6] [3] |
| Adversarial | | |

focus on partial feedback, where the main gaps remain. Since previous work [5, 2] has shown that linear regret is inevitable under the LBB constraint, we concentrate on the following three settings: GBB partial-feedback mechanisms for independent/correlated/adversarial values.

Together with prior work, our results complete the regret-minimization picture for fixed-price bilateral trade up to polylogarithmic factors. Table 1 summarizes our main results. We also introduce two technical ingredients – fractal elimination and a new GBB lower-bound construction – which may be useful in related online mechanism-design problems.

## 1.1 Technical Overview

To achieve the results presented in Table 1, we have developed several technical ingredients that may be of independent interest. We highlight the two most important contributions:

- **Ingredient 1:** A novel algorithmic paradigm called *fractal elimination* for devising one-bit-feedback mechanisms for "independent values". It is inspired by the recursive structure of fractals; see the full version for the geometric intuition behind the name.
- **Ingredient 2:** A new *lower-bound construction* for "correlated values".

For readability, we omit many minor technical details.

### Ingredient 1: Fractal Elimination for One-Bit Feedback and Independent Values

We begin by explaining why one-bit feedback greatly complicates the design of low-regret mechanisms for "independent values $(S, B) \sim \mathcal{D}_S \bigotimes \mathcal{D}_B$". The algorithmic framework proposed in [3] allows us to first accumulate sufficient profit and then discretize the action space near the diagonal. For an action $(p, q) \in [0, 1]^2$ with $p \approx q$, we can express the expected Gains from Trade as follows [5, Lemma 1]:

$$\mathsf{GFT}(p, q) \;=\; \textstyle\int_0^p \mathcal{D}_S(x)\,\mathrm{d}x \cdot (1 - \mathcal{D}_B(q)) + \mathcal{D}_S(p) \cdot \int_q^1 (1 - \mathcal{D}_B(y))\,\mathrm{d}y + [\text{minor terms}].$$

Here we abuse notation: $\mathcal{D}_S(p) := \mathbb{P}_{S \sim \mathcal{D}_S}[S \le p]$ and $\mathcal{D}_B(q) := 1 - \mathbb{P}_{B \sim \mathcal{D}_B}[B \ge q] = \mathbb{P}_{B \sim \mathcal{D}_B}[B < q]$ denote the corresponding *cumulative distribution functions (CDF's)*.

To adapt a classic *Multi-Armed Bandit (MAB)* algorithm into a low-regret mechanism, we require "good enough" estimates of the $\mathsf{GFT}(p, q)$ values. The query complexity for this task depends critically on the available feedback.

- **Two-Bit Feedback $(X^t, Y^t)$:** As observed in [5], with access to the seller's intention to trade $X^t$, only $\widetilde{\mathcal{O}}(\varepsilon^{-2})$ queries/rounds are needed to estimate the integral $\int_0^p \mathcal{D}_S(x)\,\mathrm{d}x$ within error $\varepsilon > 0$, *pointwise over the entire interval* $p \in [0, 1]$; similarly for $\int_q^1 (1 -$

$\mathcal{D}_B(y))\ \mathrm{d}y$. Thus, estimating $\mathsf{GFT}(p,q)$ reduces to estimating $\mathcal{D}_S(p)$ and $1-\mathcal{D}_B(q)$. Crucially, the two bits $X^t$ and $Y^t$ themselves serve as unbiased estimators for these quantities.
Feeding $X^t$ and $Y^t$ into a standard MAB algorithm directly yields an $\mathcal{O}(T^{2/3})$ regret $\mathsf{GBB}$ mechanism (building on the two-phase meta-mechanism framework from [3, Section 3]).
- **One-Bit Feedback $\boldsymbol{Z^t}$:** In contrast, with access only to the trade outcome $Z^t$, we lose the tailored unbiased estimators for $\mathcal{D}_S(p)$, $1-\mathcal{D}_B(q)$ etc., which constitutes our main technical challenge. Prior work [5, Algorithm 4] demonstrated how to construct unbiased estimators of $\mathsf{GFT}(p,q)$ from the less informative $Z^t$ queries. They showed that $\widetilde{\mathcal{O}}(\varepsilon^{-2})$ queries still suffice for $\varepsilon$-approximations. Although their estimators are optimal in query complexity, the resulting $\widetilde{\mathcal{O}}(T^{3/4})$ regret mechanism is suboptimal in regret, because their estimators inherently incur *constant regret* per query/round, rather than *diminishing regret*.

This leads to the core challenge for designing a low-regret $\mathsf{GBB}$ one-bit-feedback mechanism:

*How can we **regret-optimally** estimate the* $\mathsf{GFT}(p,q)$ *values?*

Our main contribution is an *elimination-based multi-stage* algorithm, FractalElimination, that addresses this challenge and yields an $\widetilde{\mathcal{O}}(T^{2/3})$ regret mechanism. The key idea is that when estimating $\mathcal{D}_S(p)$, we decompose it into a product of ratios and estimate these ratios recursively. Although the initial actions incur constant regret, they converge exponentially fast to the diagonal $\{(p,q) \mid 0 \le p = q \le 1\}$ as the recursion depth increases (i.e., as the algorithm proceeds stage by stage), resulting in diminishing regret. Overall, the total regret is at most $\widetilde{\mathcal{O}}(T^{2/3})$.

**Elimination Algorithms for the Standard MAB Problem.** For readability, it is helpful to review the standard MAB problem and the elimination algorithms: there are $K$ arms whose (independent) random rewards follow the Bernoulli distributions $\mathtt{Bern}(\rho_i)$ with *mean rewards* $\rho_i \in [0,1]$, $\forall i \in [K]$.

An elimination algorithm operates in $L+1 = \log\big(\frac{1}{\varepsilon}\big)+1$ stages, for some parameter $\varepsilon = \varepsilon(T,K) > 0$ to be determined. In stage $\ell \in [0:L]$, we pull each arm $2^\ell/\varepsilon$ times; by standard concentration inequalities, with high probability we can identify and eliminate those $\Omega(\sqrt{\varepsilon/2^\ell})$-suboptimal mean rewards. After all stages, the rest of the entire time horizon $[T]$ (if existential) can arbitrarily exploit the non-eliminated mean rewards–every such reward is $\mathcal{O}(\sqrt{\varepsilon/2^L}) = \mathcal{O}(\varepsilon)$-approximately optimal.

This elimination algorithm has two key features in each stage $\ell \in [0:L]$, which we call the *accuracy guarantee* (**MAB-AG**) and the *regret guarantee* (**MAB-RG**).
- (**MAB-AG**) After stage $\ell$, every non-eliminated mean reward is $\mathcal{O}(\sqrt{\varepsilon/2^\ell})$-approximately optimal.
- (**MAB-RG**) Stage $\ell$ accumulates at most $\mathcal{O}(\sqrt{\varepsilon/2^{\ell-1}}) \cdot 2^\ell/\varepsilon \cdot K = \mathcal{O}(K/\varepsilon)$ regret.[1]

In combination, the total regret of this elimination algorithm is at most

$$\textstyle\sum_{\ell\in[0:L]} \mathcal{O}(K/\varepsilon) + T\cdot\mathcal{O}(\sqrt{\varepsilon/2^L}) \;=\; \widetilde{\mathcal{O}}(K/\varepsilon + T\cdot\varepsilon).$$

By choosing $\varepsilon = \widetilde{\Theta}(\sqrt{K/T})$, we achieve (nearly) optimal $\widetilde{\mathcal{O}}(\sqrt{KT})$ regret.

[1] More rigorously, the initial stage $\ell = 0$ requires a different deduction to the same regret bound $1\cdot 2^0/\varepsilon\cdot K = \mathcal{O}(K/\varepsilon)$.

**Elimination Algorithms for the Bilateral Trade Model.** In the bilateral trade model, it turns out that by considering the $\frac{1}{K}$-*net* of the entire action space $\{a_{i,j} := (\frac{i}{K}, \frac{j-1}{K})\}_{i,j\in[K]} \subseteq [0,1]^2$, at least one of the near-diagonal actions $\{a_{k,k}\}_{k\in[K]}$ is a $\frac{1}{K}$-*approximation* to the benchmark action $(p^*, q^*)$; we thus designate $\{a_{k,k}\}_{k\in[K]}$ as the *(initial) candidates* and index them by $\mathcal{C}_0 := [1 : K]$.

If we could mechanically implement the above elimination algorithm for the standard MAB problem, then the total regret would be at most

$$\widetilde{\mathcal{O}}(K/\varepsilon + \varepsilon T) \;=\; \widetilde{\mathcal{O}}(T^{2/3}),$$

by choosing $K = \widetilde{\Theta}(\varepsilon^{-1})$ and $\varepsilon = \widetilde{\Theta}(T^{-1/3})$, which gives the desired regret bound.

**Fractal elimination.** The standard elimination template suggests maintaining near-diagonal candidates $a_{k,k}$ and eliminating clearly suboptimal ones stage by stage. The obstacle is that, under one-bit feedback, playing a candidate itself does not provide an unbiased estimator for its GFT. A direct estimator queries separate polylines for different candidates and therefore incurs too much regret.

Our key idea is to reuse information across candidates. At each stage, the surviving candidates are enclosed by short diagonal segments. Around each segment, we query shared horizontal and vertical actions. By independence of the seller's and buyer's values, the desired GFT values can be reconstructed through recursive ratio decompositions. As the stages progress, the queried actions move exponentially closer to the diagonal, so their per-query regret diminishes. This simultaneously gives the accuracy invariant and the regret invariant needed for the $\widetilde{O}(T^{2/3})$ bound. The full recursive implementation and concentration analysis appear in the full version.

## Ingredient 2: A New Lower-Bound Construction, for Correlated Values

Our second main contribution is an $\Omega(T^{3/4})$ lower bound for GBB two-bit-feedback mechanisms in the "correlated values" setting, which improves the $\Omega(T^{5/7})$ lower bound obtained in [3, Theorem 5.5] for the same context and matches the $\widetilde{\mathcal{O}}(T^{3/4})$ upper bound obtained in [3, Theorem 5.4] even for more general contexts.

At a high level, the technical challenge is due in large part to the GBB constraint – it introduces relevance among different rounds – and the crux of our proof is a new remedy for it. For readability, let us first omit the GBB constraint and sketch a general lower-bound approach. Then, we will carefully compare the previous GBB remedy [3] and our new GBB remedy.

**A General Lower-Bound Approach.** Basically, a regret lower bound requires constructing a family of *hard-to-distinguish* instances: When facing some instance from this family, a mechanism must determine its identity in the online learning process, namely "finding a needle in a haystack".

To address our problem using this approach, we shall construct one *base instance* $\mathcal{D}_0$ and $K \geq 1$ *hard instances* $\{\mathcal{D}_k\}_{k\in[K]}$ – recall that an instance is a $[0,1]^2$-supported joint distribution. Each hard instance $\mathcal{D}_k$ shall differ from the base instance $\mathcal{D}_0$ by some $\delta > 0$ in the total variation distance. As such, $\mathcal{D}_k$ can simply perturb $\mathcal{D}_0$ by total probability mass of $\Theta(\delta)$, distributed across constant number of actions, which forms the "needle". The construction shall follow two criteria:

- Information-Regret Dilemma: Each hard instance $\mathcal{D}_k$ has a set of *informative actions*. Only those actions can provide information (on playing) that helps distinguish this hard instance $\mathcal{D}_k$, but each of them will incur *constant regret* $\Omega(1)$.
- Disjointness: All hard instances $\{\mathcal{D}_k\}_{k\in[K]}$ shall have *disjoint* sets of informative actions, thus no information sharing on individual plays of informative actions of different $\mathcal{D}_k$'s.

Given such a construction (if possible), we can informally reason about the total regret of a mechanism as follows:

- If all individual hard instances $\{\mathcal{D}_k\}_{k\in[K]}$ are distinguishable from the base instance $\mathcal{D}_0$, given the total variation distances of $\delta$, this necessitates $\Omega(\delta^{-2})$ number of plays of a single $\mathcal{D}_k$'s informative actions and thus $\Omega(K\delta^{-2})$ number of such plays altogether (Disjointness). However, then, we will suffer from $\Omega(K\delta^{-2})$ total regret (Information-Regret Dilemma).
- Otherwise, some hard instances $\mathcal{D}_k$ are indistinguishable, and (roughly speaking) we will suffer from $\Omega(\delta T)$ total regret when facing one of them.

By choosing $\delta = K^{1/3}T^{-1/3}$, any mechanism will incur total regret

$$\min\left\{\Omega(K\delta^{-2}),\ \Omega(\delta T)\right\} \ = \ \Omega(K^{1/3}T^{2/3}). \tag{1}$$

Consequently, proving an optimal lower bound reduces to the task of seeking a construction that has the largest possible $K \geq 1$ and, simultaneously, retains the above criteria.

It remains to determine the bottleneck of $K \geq 1$. Since each hard instance $\mathcal{D}_k$ for $k \in [K]$ needs to perturb the base instance $\mathcal{D}_0$ by total probability mass of $\Theta(\delta)$ (planting a "mass-$\delta$ needle"), and these perturbations are "disjoint", we naturally require $K = \mathcal{O}(\delta^{-1})$. So if a construction can satisfy $K = \Theta(\delta^{-1})$, plugging this into Equation (1) directly gives an $\Omega(T^{3/4})$ lower bound, as desired.

To implement the above general approach in practice, however, the technical challenge is due in large part to the GBB constraint. As we quote from [3]:

> *"This (the GBB constraint) considerably complicates the construction of the hard instances, as any algorithm could sacrifice some profit temporarily by posting prices with $P^t > Q^t$ to extract a large Gains from Trade."*

The previous work [3] and our work will adopt very different remedies for the GBB constraint.

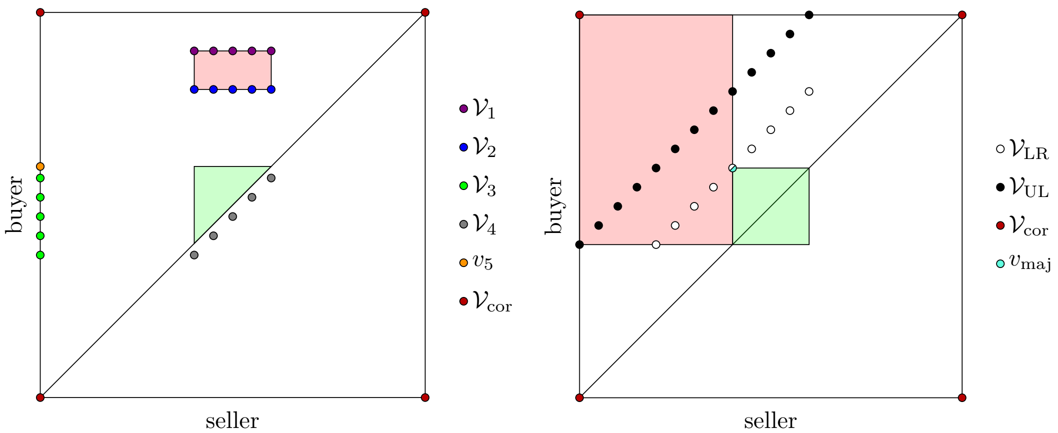
**The Previous GBB Remedy.** The previous work [3] circumvents the GBB constraint in an ingenious manner. Their base instance $\mathcal{D}_0$ (see Figure 1a for a diagram) involves, *just below the diagonal*, value points "so bad" that even a single action play on their *lower right side* will incur intolerable regret.

▷ *These value points refer to grey points $\mathcal{V}^4$ in Figure 1a. Further, "so bad" means the* GFT*-decrease due to a single such value point even dominates the total* GFT*-increase due to all other value points vertically above it.*

This automatically forces a regret-optimal mechanism to satisfy the GBB constraint. In other words, central to their lower-bound construction is such a more "qualitative" principle:

> *Sacrifice of profit* ***cannot*** *produce extra* Gains from Trade.

As it turns out, there are as many "so bad" value points as the hard instances ($K \geq 1$), and they each have probability mass $\Omega(K\delta)$. Then, it can be shown that $K = \Theta(\delta^{-1/2})$. Plugging this in Equation (1) gives an $\Omega(T^{5/7})$ lower bound [3, Theorem 5.5].




**(a)** The previous construction. **(b)** Our new construction.

**Figure 1** Diagrams of the previous $\Omega(T^{5/7})$ lower-bound construction [3, Theorem 5.5] and our new $\Omega(T^{3/4})$ lower-bound construction (Theorem 6).

**Our New GBB Remedy.** Here, our contribution is a more careful treatment of the GBB constraint instead of simply circumventing it like [3]. Specifically, instead of incorporating "so bad" value points into the base instance $\mathcal{D}_0$, our lower-bound construction derives from a more "quantitative" principle:

> *Sacrifice of profit ("investment")* ***can*** *produce extra Gains from Trade ("return").*
> *Just, the "return on investment" is* ***not worthy enough*** *for regret minimization.*

This means a regret-optimal mechanism must abandon "investment" and, thus, restrict its action plays to the *upper left side* of the diagonal. In this manner, our more "quantitative" principle reaches the same goal – the GBB constraint – as the more "qualitative" principle by [3]. Technically, our principle is also more flexible for lower-bound construction, which enables us to construct $K = \Theta(\delta^{-1})$ many hard instances (while retaining the mentioned criteria) and thus show an $\Omega(T^{3/4})$ lower bound.

We have developed new techniques to implement the above discussions into a formal proof. In particular, we introduce a new constraint called *Global Price Balance*, given by

$$\mathbb{E}\left[\sum_{t\in[T]}(Q^t - P^t)\right] \;\geq\; 0.$$

Under our lower-bound construction, this new constraint turns out to be a consequence/relaxation of the original GBB constraint – every GBB mechanism must satisfy it – and is relatively easier to manipulate. Indeed, our $\Omega(T^{3/4})$ lower bound holds "more generally" for any mechanism that satisfies this new constraint; see Section 4 for details.

## 1.2 Concurrent Works [11, 10, 9]

Concurrent and independent works studied closely related questions. In particular, [11] independently introduced a feedback model equivalent to our semi feedback, while [10] independently proved a matching $\Omega(T^{3/4})$ lower bound for GBB partial-feedback mechanisms in the correlated/adversarial settings using a different construction. The recent work [9] studies the related profit-maximization objective.

## 2 Notations and Preliminaries

Given two nonnegative integers $m \geq n \geq 0$, define the sets $[n\colon m] := \{n, n+1, \dots, m-1, m\}$ and $[n] := [1\colon n] = \{1, 2, \dots, n\}$. Given a (possibly random) event $\mathcal{E}$, let $\mathbb{1}[\mathcal{E}] \in \{0, 1\}$ be the indicator function. Also, given a real number $x \in \mathbb{R}$, let $[x]_+ := \max\{x, 0\}$.

**Repeated Bilateral Trade.** Our model involves a $T$-round[2] repeated game against an (oblivious) adversary: In each round $t \in [T]$, a (new) seller and a (new) buyer seek to trade an indivisible item, which has value $S^t$ to the seller and value $B^t$ to the buyer. We aim to maximize the *economic efficiency* by designing a (repeated) mechanism that, ideally, enables trade whenever $S^t \leq B^t$. However, the adversary controls the value generation $(S^t, B^t)_{t \in [T]}$; there are three classic models, listed below from the most to the least general.

- **Adversarial Values:** The adversary determines an (arbitrary) $2T$-dimensional $[0, 1]^{2T}$-supported joint distribution $\mathcal{D}$; then, the values $(S^t, B^t)_{t \in [T]}$ across all rounds are drawn from it.
- **Correlated Values:** The adversary determines an (arbitrary) two-dimensional $[0, 1]^2$-supported joint distribution $\mathcal{D}$; then, the values $(S^t, B^t)$ in each round $t \in [T]$ are drawn i.i.d. from it.
- **Independent Values:** This is identical to "correlated values", except that $\mathcal{D}$ is further required to be a product distribution $\mathcal{D}_S \bigotimes \mathcal{D}_B$, making all the $2T$ values $(S^t, B^t)_{t \in [T]}$ mutually independent.

Some of our lower bounds hold even under the *density-boundedness* assumption, which was introduced to the repeated bilateral trade model by [5]. Clearly, this can only strengthen our hardness results.

▶ **Assumption 1** (Density Boundedness [5]). *Parameterized by $M \geq 1$, a joint distribution $\mathcal{D}$ satisfies the density-boundedness assumption when its joint density function is upper-bounded by $M$.*

**Fixed-Price Mechanisms.** A fixed-price mechanism $\mathcal{M} = (P^t, Q^t)_{t \in [T]}$ posts two possibly randomized prices in each round $t \in [T]$, one for the seller $P^t$ and one for the buyer $Q^t$. A trade occurs when both agents accept their respective prices. This yields the *Gains from Trade* $\mathsf{GFT}(S^t, B^t, P^t, Q^t)$ and the *profit* $\mathsf{Profit}(S^t, B^t, P^t, Q^t)$, $\forall t \in [T]$:

$$\begin{aligned} \mathsf{GFT}(S^t, B^t, P^t, Q^t) &:= (B^t - S^t) \cdot \mathbb{1}[S^t \leq P^t] \cdot \mathbb{1}[Q^t \leq B^t], \\ \mathsf{Profit}(S^t, B^t, P^t, Q^t) &:= (Q^t - P^t) \cdot \mathbb{1}[S^t \leq P^t] \cdot \mathbb{1}[Q^t \leq B^t]. \end{aligned}$$

For notational brevity, we often simply write $\mathsf{GFT}^t = \mathsf{GFT}(S^t, B^t, P^t, Q^t)$ when the values $(S^t, B^t)$ and the prices $(P^t, Q^t)$ are clear from the context; such conventions extend to other notations.

When no ambiguity arises, we use the term *mechanisms* to refer to *fixed-price mechanisms*. A mechanism $\mathcal{M}$ is initially ignorant of the underlying joint distribution $\mathcal{D}$ and the values $(S^t, B^t)_{t \in [T]} \sim \mathcal{D}$, but receives certain feedback at the end of each round $t \in [T]$; given the past prices $(P^r, Q^r)_{r \in [t]}$ and the past feedback, it proceeds to the next round $t + 1$ and computes the prices $(P^{t+1}, Q^{t+1})$. There are three natural feedback models, listed below from the most to the least informative.

[2] Throughout this work, we fix $T \gg 1$ as a sufficiently large integer.

- **Full Feedback:** $(S^t, B^t) \in [0,1]^2$ reveals both agents' values.
- **Semi Feedback:** This consists of four specific types:
  - $(S^t, Y^t) \in [0,1] \times \{0,1\}$ reveals the seller's value $S^t$ and the buyer's intention to trade $Y^t = Y(B^t, Q^t) := \mathbb{1}[Q^t \leq B^t]$.
  - $(X^t, B^t) \in \{0,1\} \times [0,1]$ reveals the seller's intention to trade $X^t = X(S^t, P^t) := \mathbb{1}[S^t \leq P^t]$ and the buyer's value $B^t$.
  - $(S^t, Z^t) \in [0,1] \times \{0,1\}$ reveals the seller's value $S^t$ and the trade outcome $Z^t = Z(S^t, B^t, P^t, Q^t) := \mathbb{1}[S^t \leq P^t \wedge Q^t \leq B^t] \equiv X^t \wedge Y^t$.
  - $(Z^t, B^t) \in \{0,1\} \times [0,1]$ reveals the trade outcome $Z^t \equiv X^t \wedge Y^t$ and the buyer's value $B^t$.
- **Partial Feedback:** This consists of two specific types:
  - **Two-Bit Feedback:** $(X^t, Y^t) \in \{0,1\}^2$ reveals both agents' intentions to trade.
  - **One-Bit Feedback:** $Z^t \equiv X^t \wedge Y^t \in \{0,1\}$ reveals the trade outcome.

Previous works [5, 2, 3] have already acquired a thorough understanding of "full feedback", so we concentrate on "semi feedback" and "partial feedback" in this work.

For the sake of economic viability, we further impose the *Budget Balance (BB)* constraint. There are two notions, a stricter *Local Budget Balance (LBB)* constraint [12] and a looser *Global Budget Balance (GBB)* constraint [3].

- Local Budget Balance: This can be further differentiated into a stricter *Strong Budget Balance (SBB)* constraint and a looser *Weak Budget Balance (WBB)* constraint.
  - Strong Budget Balance: $\mathsf{Profit}^t = 0$ (or essentially $P^t = Q^t$), $\forall t \in [T]$.
  - Weak Budget Balance: $\mathsf{Profit}^t \geq 0$ (or essentially $P^t \leq Q^t$), $\forall t \in [T]$.
- Global Budget Balance: $\sum_{t \in [T]} \mathsf{Profit}^t \geq 0$.

To summarize the scope of this work, there are $3 \times 3 \times 2 = 18$ specific contexts. After omitting the well-understood "adversarial values" and "full feedback", $2 \times 2 \times 2 = 8$ contexts remain. In this proceedings version, we focus on the contexts needed for our main positive and negative results; the remaining auxiliary lower-bound statements are deferred to the full version.

**Regret Minimization.** We evaluate the economic efficiency of a fixed-price mechanism $\mathcal{M}$ within the *regret minimization* framework.

Over the entire repeated game, this mechanism $\mathcal{M}$ induces *Total Gains from Trade* $\mathsf{GFT}_{\mathcal{D}}^{\mathcal{M}}$ in expectation (over all possible randomness $(S^t, B^t)_{t \in [T]} \sim \mathcal{D}$ and $(P^t, Q^t)_{t \in [T]} \leftarrow \mathcal{M}$). We compare this to *Bayesian-Optimal Total Gains from Trade* $\mathsf{GFT}_{\mathcal{D}}^*$, which refers to the optimal-in-expectation fixed prices $(p^*, q^*)$ with respect to the underlying joint distribution $\mathcal{D}$.

$$
\begin{aligned}
\mathsf{GFT}_{\mathcal{D}}^{\mathcal{M}} &:= \mathbb{E}_{(S^t, B^t)_{t \in [T]} \sim \mathcal{D},\ (P^t, Q^t)_{t \in [T]} \leftarrow \mathcal{M}} \left[ \sum_{t \in [T]} \mathsf{GFT}(S^t, B^t, P^t, Q^t) \right], \\
\mathsf{GFT}_{\mathcal{D}}^{*} &:= \max_{0 \leq p^* = q^* \leq 1} \mathbb{E}_{(S^t, B^t)_{t \in [T]} \sim \mathcal{D}} \left[ \sum_{t \in [T]} \mathsf{GFT}(S^t, B^t, p^*, q^*) \right].
\end{aligned}
$$

Without ambiguity, we often write $\mathbb{E}_{\mathcal{D}}[\cdot] = \mathbb{E}_{(S^t, B^t)_{t \in [T]} \sim \mathcal{D}}[\cdot]$ etc for notational brevity.

For a mechanism $\mathcal{M}$, we can define its *(worst-case) regret* $\mathsf{Regret}^{\mathcal{M}}$ by taking into account all possible joint distributions $\mathcal{D}$. In this regard, we aim to find the *minimax regret* $\mathsf{Regret}^*$ by designing a regret-optimal mechanism.

**Algorithm 1** GBB-OneBit.

1: Run the subroutine ProfitMax$(K, \beta)$. ▷ *Cf. Proposition 3 and Corollary 4.*
2: Run the subroutine FractalElimination$(0, [1 : K], 0, 0)$. ▷ *Cf. Algorithm 2.*
3: Take actions $\{a_{k,k}\}_{k \in \mathcal{C}_{L+1}}$ survived in Line 2, in an arbitrary manner, for the remaining rounds.

$$\begin{aligned} \mathsf{Regret}^{\mathcal{M}} &:= \max_{\mathcal{D}} \left(\mathsf{GFT}^*_{\mathcal{D}} - \mathsf{GFT}^{\mathcal{M}}_{\mathcal{D}}\right), \\ \mathsf{Regret}^* &:= \min_{\mathcal{M}} \mathsf{Regret}^{\mathcal{M}}. \end{aligned}$$

We will also use $\mathsf{Regret}^{\mathcal{M}}_{\mathcal{D}} := \mathsf{GFT}^*_{\mathcal{D}} - \mathsf{GFT}^{\mathcal{M}}_{\mathcal{D}}$ to denote the regret of $\mathcal{M}$ on a specific distribution $\mathcal{D}$. When $\mathcal{M}$ is clear from the context, we may drop it from the superscript.

# 3 $\widetilde{\mathcal{O}}(T^{2/3})$ Partial-Feedback Upper Bound for Independent Values

In this section, we examine the power of "fixed-price mechanisms with the Global Budget Balance (GBB) constraint and one-bit feedback" in the "independent values" setting. Specifically, we will establish (Theorem 1) the following algorithmic result.

▶ **Theorem 1** (GBB One-Bit-Feedback Upper Bound for Independent Values)**.** *In the "independent values" setting, there is a "GBB one-bit-feedback fixed-price mechanism" achieving $\widetilde{\mathcal{O}}(T^{2/3})$ regret.*

Later in Appendix A, up to polylogarithmic factors, we will establish (Theorem 11) a matching lower bound (even if one-bit feedback is replaced by the more informative *semi feedback*).

In the literature, only a trivial $\widetilde{\mathcal{O}}(T^{3/4})$ upper bound and a trivial $\Omega(T^{1/2})$ lower bound were known – the upper bound is an implication from [3, Theorem 5.4] for "one-bit feedback, adversarial values", and the lower bound is an implication from Theorem 12 for "full feedback, independent values". Nonetheless, our algorithmic result here and our hardness result later in Theorem 11 together close the gap.

## 3.1 Mechanism Design

Our fixed-price mechanism, called GBB-OneBit and presented in Algorithm 1, is built on the mechanism design framework proposed by [3, Section 3]. This fixed-price mechanism has three phases:

Phase 1 invokes a subroutine, called ProfitMax and depicted in Proposition 3, which takes actions only from the *upper-left* action halfspace $\{(p, q) \in [0, 1]^2 \mid p \le q\}$, thus *nonnegative* profit $\ge 0$ per round. The main purpose of this phase is to accumulate sufficient profit, say $\widetilde{\Omega}(T^{2/3})$, while just incurring tolerable regret, say $\widetilde{\mathcal{O}}(T^{2/3})$. Regarding the GBB constraint, this cumulative profit makes mechanism design in subsequent phases more flexible.

Phase 2 invokes a subroutine, called FractalElimination and shown in Algorithm 2, which takes actions only from the *lower-right* action halfspace $\{(p, q) \in [0, 1]^2 \mid p > q\}$, thus *nonpositive* profit $\le 0$ per round. Concretely, this subroutine begins with a set of $K = \widetilde{\Theta}(T^{1/3})$ many *candidate nearly* GFT*-optimal actions* (or *candidates* in short) indexed by $\mathcal{C}_0 = [1 : K]$; the GFT-optimal candidate is ensured to be a good enough approximation to the GFT-optimal action $(p^*, q^*) \in [0, 1]^2$ in the whole action space. The subroutine works in $L + 1 \approx \log(K)$

many stages; a single stage $\ell \in [0 : L]$ leverages *one-bit feedback* to distinguish the survival candidates $\mathcal{C}_\ell$ hitherto (by taking actions from not only $\mathcal{C}_\ell$ themselves, but also other actions in the lower-right action halfspace ), obtaining a more accurate location $\mathcal{C}_{\ell+1} \subseteq \mathcal{C}_\ell$ of the optimal candidate. After all the $L + 1 \approx \log(K)$ many stages, the ultimate candidates $\mathcal{C}_{L+1}$ all will be good enough, compared even with the optimal action $(p^*, q^*) \in [0, 1]^2$ in the whole action space.

Phase 3 simply exploits the ultimate candidates $\mathcal{C}_{L+1}$, in an arbitrary manner.

Remarkably, as it turns out, Phases 2 and 3 never exhaust the profit accumulated in Phase 1, so the whole fixed-price mechanism GBB-OneBit satisfies the GBB constraint.

**Preliminaries.** Our fixed-price mechanism GBB-OneBit will use the following parameters. Specifically, both subroutines ProfitMax and FractalElimination will use the *discretization parameter* $K$, only the former will use the *profit threshold* $\beta$, and only the latter will use the others $L$, $\delta$, and $\gamma_\ell$'s.

$$\begin{aligned}
K &:= \tfrac{1}{8} T^{1/3} \log^{-2/3}(T), && \text{(the discretization parameter)}\\
\beta &:= 9T^{2/3} \log^{2/3}(T), && \text{(the profit threshold)}\\
L &:= \tfrac{1}{3} \log(T), && \text{(the number of stages)}\\
\mathcal{C}_0 &:= [1 : K], && \text{(for initialization)}\\
\delta &:= T^{-4/3} \log^{-1/3}(T), && \text{(for confidence levels)}\\
\gamma_\ell &:= 2^{-\ell/2} K^{-1/2} + (6\ell - 1)K^{-1}, \qquad \forall \ell \in [0 : L + 1]. && \text{(for confidence intervals)}
\end{aligned}$$

In expectation over the randomness of values $(S, B) \sim \mathcal{D}_S \bigotimes \mathcal{D}_B$, an action $(p, q) \in [0, 1]^2$ induces (expected) Gains from Trade $\mathsf{GFT}(p, q)$, (expected) profit $\mathsf{Profit}(p, q)$, and (expected) regret $\mathsf{Regret}(p, q)$. It is easy to see that these formulae are $[0, 1]$-valued.

$$\begin{aligned}
\mathsf{GFT}(p, q) &:= \mathbb{E}_{(S,B)\sim\mathcal{D}_S \bigotimes \mathcal{D}_B}[\mathsf{GFT}(S, B, p, q)], && \forall (p, q) \in [0, 1]^2,\\
\mathsf{Regret}(p, q) &:= \big(\max_{0\le p'\le q'\le 1} \mathsf{GFT}(p, q)\big) - \mathsf{GFT}(p, q), && \forall (p, q) \in [0, 1]^2,\\
\mathsf{Profit}(p, q) &:= \mathbb{E}_{(S,B)\sim\mathcal{D}_S \bigotimes \mathcal{D}_B}[\mathsf{Profit}(S, B, p, q)]\\
&= (q - p) \cdot \mathcal{D}_S(p) \cdot (1 - \mathcal{D}_B(q)), && \forall (p, q) \in [0, 1]^2.
\end{aligned}$$

Here we abuse notation: $\mathcal{D}_S(p) := \mathbb{P}_{S\sim\mathcal{D}_S}[S \le p]$ and $\mathcal{D}_B(q) := 1 - \mathbb{P}_{B\sim\mathcal{D}_B}[B \ge q] = \mathbb{P}_{B\sim\mathcal{D}_B}[B < q]$ denote the corresponding *cumulative distribution functions (CDF's)*.

The following Lemma 2 shows a useful decomposition of the formula $\mathsf{GFT}(p, q)$.

▶ **Lemma 2** (Gains from Trade for Independent Values). *In the "independent values" settings,*

$$\mathsf{GFT}(p, q) \;=\; \mathsf{H}(p, q) + \mathsf{V}(p, q) + \mathsf{Profit}(p, q), \qquad \forall (p, q) \in [0, 1]^2.$$

*Here the terms* $\mathsf{H}(p, q) := \int_0^p \mathcal{D}_S(x)\mathrm{d}x \cdot (1 - \mathcal{D}_B(q))$ *and* $\mathsf{V}(p, q) := \mathcal{D}_S(p) \cdot \int_q^1 (1 - \mathcal{D}_B(y))\mathrm{d}y$.

In addition, we recall that *one-bit feedback* $Z^t = Z(S^t, B^t, P^t, Q^t) \in \{0, 1\}$ reveals whether or not the trade succeeded in a single round $t \in [T]$.

$$Z(S^t, B^t, P^t, Q^t) \;=\; \mathbb{1}[S^t \le P^t] \cdot \mathbb{1}[Q^t \le B^t].$$

**The Subroutine** ProfitMax. To understand the performance guarantees of our fixed-price mechanism GBB-OneBit, all we need to know about (Phase 1 of GBB-OneBit) the subroutine ProfitMax can be summarized into the following Proposition 3, which is quoted (or, indeed, slightly rephrased) from [3, Lemma 5.1].[3] As mentioned, the purpose of ProfitMax is to accumulate sufficient profit (at the cost of tolerable regret), making mechanism design in subsequent phases more flexible. For detailed implementation of ProfitMax, the interested reader can reference [3, Sections 3 and 5].

▶ **Proposition 3** ([3, Lemma 5.1]). *There exists a fixed-price mechanism* ProfitMax$(K', \beta')$ *with one-bit feedback, on input a discretization parameter $K' \geq 1$ and a profit threshold $\beta' > 0$, such that:*

1. *It takes actions $\{(P^t, Q^t)\}_{t=1,2,\ldots}$ only from a size-$|\mathcal{F}_{K'}| = 2K'(\log(T)+1)$ discrete subset $\mathcal{F}_{K'} \subseteq \{(p,q) \in [0,1]^2 \mid p \leq q\}$ of the upper-left action halfspace.*
   *Thus, the per-round profit is nonnegative $\mathsf{Profit}(S^t, B^t, P^t, Q^t) \geq 0$, $\forall t = 1, 2, \ldots$, almost surely.*
2. *It terminates at the end of some round $T' \in [T]$, which has two possibilities:*
   *(i) $T' \in [T]$ is the first round such that $\sum_{t\in[T']} \mathsf{Profit}(S^t, B^t, P^t, Q^t) \geq \beta'$, if existential.*
   *(ii) $T' = T$, if $\sum_{t\in[T]} \mathsf{Profit}(S^t, B^t, P^t, Q^t) < \beta'$.*
   *In either case, with probability $1 - T^{-1}$, the cumulative regret $\sum_{t\in[T']} \mathsf{Regret}(P^t, Q^t)$ satisfies that*

$$\textstyle\sum_{t\in[T']} \mathsf{Regret}(P^t, Q^t) \;\leq\; (8\beta' + 8)\log(T) + \frac{5T}{K'} + 256\sqrt{T|\mathcal{F}_{K'}|\log(T|\mathcal{F}_{K'}|)} \cdot \log(T).$$

▶ **Corollary 4** (ProfitMax; Instantiation). *In the context of Proposition 3, set $K' \leftarrow K$ and $\beta' \leftarrow \beta$. Then in either case, with probability $1 - T^{-1}$, the cumulative regret $\sum_{t\in[T']} \mathsf{Regret}(P^t, Q^t) \leq 220T^{2/3}\log^{5/3}(T)$.*

In the rest of Section 3, we would call the subroutine ProfitMax *"successful"* if its cumulative regret satisfies the bound $\sum_{t\in[T']} \mathsf{Regret}(P^t, Q^t) \leq 220T^{2/3}\log^{5/3}(T)$ given in Corollary 4, or *"failed"* otherwise.

**The Subroutine** FractalElimination. Based on the profit accumulated above ($\geq \beta = \widetilde{\Theta}(T^{2/3})$, say), our fixed-price mechanism GBB-OneBit (Phase 2 thereof) then invokes the subroutine FractalElimination to distinguish the optimal action $a^* = (p^*, q^*) \in [0,1]^2$, or rather, its good enough approximations, while preserving the GBB constraint.

Now let us elaborate on this subroutine FractalElimination; see Algorithm 2 for its implementation and Figure 2 for a diagram. Before all else, we use our discretization parameter $K = \widetilde{\Theta}(T^{2/3})$ to construct the following $\frac{1}{K}$*-net* $\{a_{i,j}\}_{1\leq i,j\leq K}$ of the whole action space $[0,1]^2$ (yet FractalElimination only takes actions from the *lower-right* half $\{a_{i,j}\}_{1\leq j\leq i\leq K}$). Among these discrete actions, we designate $\{a_{k,k}\}_{k\in[1:K]}$ as *candidates* of "good enough approximations to the optimal action $a^*$"; indeed, the optimal candidate $a_{\mu,\mu}$ (say) is a good enough $\frac{1}{K}$-approximation to the optimal action $a^*$.

$$a_{i,j} \;:=\; (\tfrac{i}{K}, \tfrac{j-1}{K}), \qquad\qquad \forall 1 \leq i,j \leq K.$$

[3] This fixed-price mechanism ProfitMax is built on the EXP3.P learning algorithm by [1]; its performance guarantees given in Proposition 3 were shown by [3, Lemma 5.1] for "adversarial values", which accommodates "independent values".

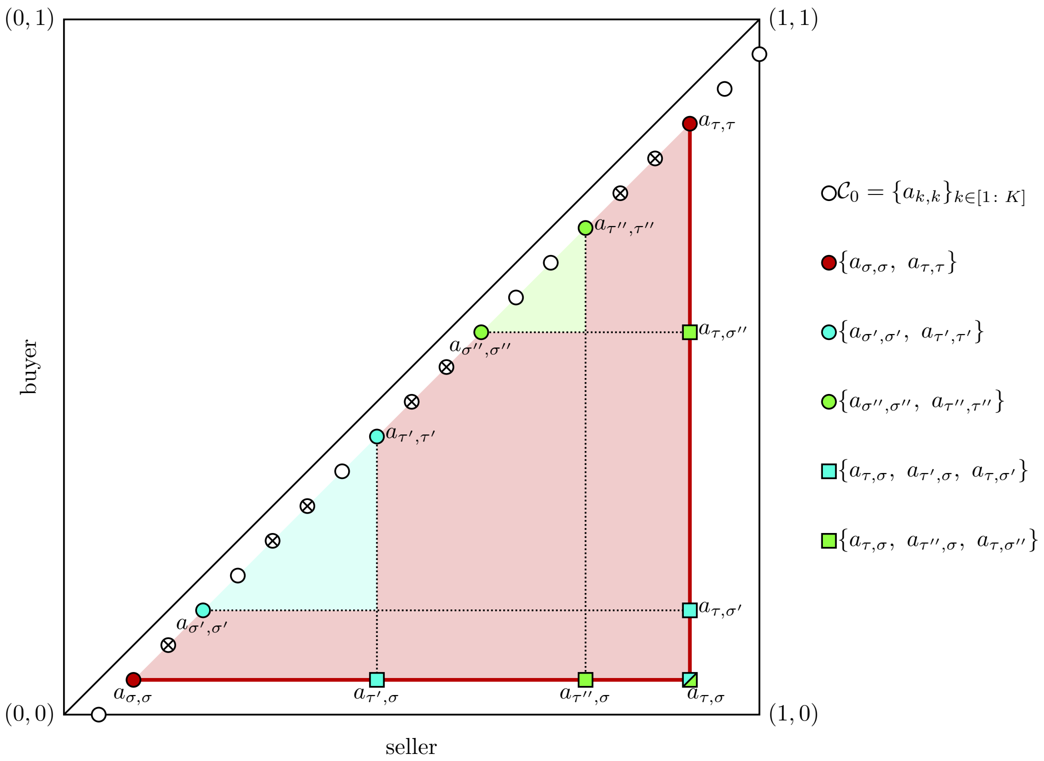



**Figure 2** Diagram of a specific stage $\ell \in [0 : L]$ of the subroutine FractalElimination (Algorithm 2).
Here, ○'s in general refer to candidates $\{a_{k,k}\}_{k\in[1:K]}$, and ⊗'s in particular refer to candidates $\{a_{k,k}\}_{k\in[1:K]}$ eliminated in the current stage $\ell \in [0 : L]$.
Also, the *red* horizontal/vertical lines $a_{\sigma,\sigma} - a_{\tau,\sigma}$ and $a_{\tau,\sigma} - a_{\tau,\tau}$ refer to actions taken in Line 2, and the six *blue/green* □'s (with $a_{\tau,\sigma}$ counted twice) refer to actions taken in Lines 11 and 15.
When FractalElimination proceeds from the current stage $\ell \in [0 : L]$ to the next stage $\ell + 1 \in [1 : L + 1]$, the considered segment $[\sigma : \tau]$ (and its associated *red* triangle) shrinks to two smaller segments $[\sigma' : \tau']$ and $[\sigma'' : \tau'']$ (and their associated *blue/green* triangles).

In regard to discrete actions $\{a_{i,j}\}_{1\le i,j\le K}$ and one-bit feedback $Z^t = \mathbb{1}[S^t \le P^t] \cdot \mathbb{1}[Q^t \le B^t] \in \{0, 1\}$, we define the *trade rates* $\{\mathsf{Z}_{i,j}\}_{1\le i,j\le K}$ as follows. It is easy to see that $\mathsf{Z}_{i,j} \in [0, 1]$ and the monotonicity $\mathsf{Z}_{1,j} \le \cdots \le \mathsf{Z}_{i,j} \le \cdots \le \mathsf{Z}_{K,j}$ and $\mathsf{Z}_{i,1} \ge \cdots \ge \mathsf{Z}_{i,j} \ge \cdots \ge \mathsf{Z}_{i,K}$.

$$\begin{aligned}\mathsf{Z}_{i,j} &:= \mathbb{E}_{(S,B)\sim\mathcal{D}_S \bigotimes \mathcal{D}_B}[Z(S, B, a_{i,j})] \\ &= \mathcal{D}_S(\tfrac{i}{K}) \cdot (1 - \mathcal{D}_B(\tfrac{j-1}{K})), \qquad \forall 1 \le i, j \le K. \end{aligned} \tag{2}$$

In regard to (Lemma 2) the decomposition $\mathsf{GFT}(p, q) = \mathsf{H}(p, q) + \mathsf{V}(p, q) + \mathsf{Profit}(p, q)$, $\forall (p, q) \in [0, 1]^2$, we define the following terms $\widetilde{\mathsf{H}}([\sigma : \tau], j)$ and $\widetilde{\mathsf{V}}(i, [\sigma : \tau])$.

$$\widetilde{\mathsf{H}}([\sigma : \tau], j) := \tfrac{1}{K} \textstyle\sum_{i\in[\sigma:\tau]} \mathsf{Z}_{i,j}, \qquad \forall[\sigma : \tau] \subseteq [1 : K],\ \forall j \in [1 : K], \tag{3}$$

$$\widetilde{\mathsf{V}}(i, [\sigma : \tau]) := \tfrac{1}{K} \textstyle\sum_{j\in[\sigma:\tau]} \mathsf{Z}_{i,j}, \qquad \forall[\sigma : \tau] \subseteq [1 : K],\ \forall i \in [1 : K]. \tag{4}$$

It is easy to check that all these terms are $[0, 1]$-bounded and, specifically, that $\widetilde{\mathsf{H}}([1 : i], j) = \mathsf{H}(a_{i,j}) \pm K^{-1}$ and $\widetilde{\mathsf{V}}(i, [j : K]) = \mathsf{V}(a_{i,j}) \pm K^{-1}$, $\forall 1 \le i, j \le K$. Moreover, every candidate $k \in [1 : K]$ induces negligible profit $\mathsf{Profit}(a_{k,k}) = \pm K^{-1}$.

**Algorithm 2** FractalElimination$(\ell, [\sigma : \tau], \mathsf{h}, \mathsf{v})$.

**Input:** $\ell \in [0 : L]$ – the current stage.
$[\sigma : \tau] \subseteq [1 : K]$ – the considered segment.
$\mathsf{h} \in [0, 1]$ – an estimate of $\widetilde{\mathsf{H}}([1 : \sigma - 1], \sigma)$.
$\mathsf{v} \in [0, 1]$ – an estimate of $\widetilde{\mathsf{V}}(\tau, [\tau + 1 : K])$.

1: **if** $\ell > L$ **then** Quit.
2: Take actions $\{a_{i,\sigma}\}_{i\in[\sigma:\tau]} \cup \{a_{\tau,j}\}_{j\in[\sigma:\tau]}$ each for $2^{\ell+2}K \ln\big(\frac{2}{\delta}\big)$ rounds, thus one-bit feedback $Z^t$'s.
3: $\{\widehat{\mathsf{Z}}_{i,\sigma}\}_{i\in[\sigma:\tau]} \cup \{\widehat{\mathsf{Z}}_{\tau,j}\}_{j\in[\sigma:\tau]} \leftarrow$ "empirical means of (index-wise) one-bit feedback $Z^t$'s by Line 2".

4: **for** every candidate $k \in \mathcal{C}_\ell \bigcap [\sigma : \tau]$ in the considered segment **do**
5: $\quad \widehat{\mathsf{H}}([\sigma : k], \sigma) \leftarrow \underbrace{\frac{1}{K} \sum_{i\in[\sigma:k]} \widehat{\mathsf{Z}}_{i,\sigma}}_{\text{Line 3}}$ and $\widehat{\mathsf{V}}(\tau, [k : \tau]) \leftarrow \underbrace{\frac{1}{K} \sum_{j\in[k:\tau]} \widehat{\mathsf{Z}}_{\tau,j}}_{\text{Line 3}}$.
6: $\quad \widehat{\mathsf{GFT}}_\ell[k] \leftarrow (\mathsf{h} + \underbrace{\widehat{\mathsf{H}}([\sigma : k], \sigma)}_{\text{Line 5}}) \cdot \underbrace{[\widehat{\mathsf{Z}}_{\tau,k}/\widehat{\mathsf{Z}}_{\tau,\sigma}]_{\downarrow 1}}_{\text{Line 3}} + (\mathsf{v} + \underbrace{\widehat{\mathsf{V}}(\tau, [k : \tau])}_{\text{Line 5}}) \cdot \underbrace{[\widehat{\mathsf{Z}}_{k,\sigma}/\widehat{\mathsf{Z}}_{\tau,\sigma}]_{\downarrow 1}}_{\text{Line 3}}$.

7: $\mathcal{C}_{\ell+1} \leftarrow \{k \in \mathcal{C}_\ell \mid \widehat{\mathsf{GFT}}_\ell[k] \geq \max_{c\in\mathcal{C}_\ell} \widehat{\mathsf{GFT}}_\ell[c] - 2\gamma_{\ell+1}\}$. ▷ $\mathcal{C}_0 = [1 : K]$.
8: $\sigma' \leftarrow \min \mathcal{C}_{\ell+1} \bigcap [\sigma : \frac{\sigma+\tau}{2}]$ and $\tau' \leftarrow \max \mathcal{C}_{\ell+1} \bigcap [\sigma : \frac{\sigma+\tau}{2}]$. ▷ *The "lower-left" half-segment.*
9: $\sigma'' \leftarrow \min \mathcal{C}_{\ell+1} \bigcap [\frac{\sigma+\tau}{2} + 1 : \tau]$ and $\tau'' \leftarrow \max \mathcal{C}_{\ell+1} \bigcap [\frac{\sigma+\tau}{2} + 1 : \tau]$. ▷ *The "upper-right" half-segment.*
10: Skip Lines 11 to 14 (resp. Lines 15 to 18) when $\mathcal{C}_{\ell+1} \bigcap [\sigma : \frac{\sigma+\tau}{2}] = \varnothing$ (resp. $\mathcal{C}_{\ell+1} \bigcap [\frac{\sigma+\tau}{2} + 1 : \tau] = \varnothing$).

11: Take actions $\{a_{\tau,\sigma},\ a_{\tau',\sigma},\ a_{\tau,\sigma'}\}$ each for $\frac{1}{2}K^2 \ln\big(\frac{2}{\delta}\big)$ rounds.
12: $\{\widehat{\mathsf{Z}}_{\tau,\sigma},\ \widehat{\mathsf{Z}}_{\tau',\sigma},\ \widehat{\mathsf{Z}}_{\tau,\sigma'}\} \leftarrow$ "empirical means of (index-wise) one-bit feedback $Z^t$'s by Line 11".
13: $\mathsf{h}' \leftarrow (\mathsf{h} + \underbrace{\widehat{\mathsf{H}}([\sigma : \sigma' - 1], \sigma)}_{\text{Line 5}}) \cdot \underbrace{[\widehat{\mathsf{Z}}_{\tau,\sigma'}/\widehat{\mathsf{Z}}_{\tau,\sigma}]_{\downarrow 1}}_{\text{Line 12}}$ and $\mathsf{v}' \leftarrow (\mathsf{v} + \underbrace{\widehat{\mathsf{V}}(\tau, [\tau' + 1 : \tau])}_{\text{Line 5}}) \cdot \underbrace{[\widehat{\mathsf{Z}}_{\tau',\sigma}/\widehat{\mathsf{Z}}_{\tau,\sigma}]_{\downarrow 1}}_{\text{Line 12}}$.
14: FractalElimination$(\ell + 1, [\sigma' : \tau'], \mathsf{h}', \mathsf{v}')$.

15: Take actions $\{a_{\tau,\sigma},\ a_{\tau'',\sigma},\ a_{\tau,\sigma''}\}$ each for $\frac{1}{2}K^2 \ln\big(\frac{2}{\delta}\big)$ rounds.
16: $\{\widehat{\mathsf{Z}}_{\tau,\sigma},\ \widehat{\mathsf{Z}}_{\tau'',\sigma},\ \widehat{\mathsf{Z}}_{\tau,\sigma''}\} \leftarrow$ "empirical means of (index-wise) one-bit feedback $Z^t$'s by Line 15".
17: $\mathsf{h}'' \leftarrow (\mathsf{h} + \underbrace{\widehat{\mathsf{H}}([\sigma : \sigma'' - 1], \sigma)}_{\text{Line 5}}) \cdot \underbrace{[\widehat{\mathsf{Z}}_{\tau,\sigma''}/\widehat{\mathsf{Z}}_{\tau,\sigma}]_{\downarrow 1}}_{\text{Line 16}}$ and $\mathsf{v}'' \leftarrow (\mathsf{v} + \underbrace{\widehat{\mathsf{V}}(\tau, [\tau'' + 1 : \tau])}_{\text{Line 5}}) \cdot \underbrace{[\widehat{\mathsf{Z}}_{\tau'',\sigma}/\widehat{\mathsf{Z}}_{\tau,\sigma}]_{\downarrow 1}}_{\text{Line 16}}$.
18: FractalElimination$(\ell + 1, [\sigma'' : \tau''], \mathsf{h}'', \mathsf{v}'')$.

FractalElimination follows a *divide-and-conquer* principle and, over $L + 1 \approx \log(K)$ stages, locates the optimal candidate $a_{\mu,\mu}$ more and more accurately. In more details:

*Induction Hypothesis.* Before a specific stage $\ell \in [0 : L]$, we have already located $a_{\mu,\mu}$ in a *candidate set* $\mathcal{C}_\ell \subseteq [1 : K]$, and we have up to $2^\ell$ many disjoint *segments* $[\sigma : \tau] \subseteq [1 : K]$ whose union covers $\mathcal{C}_\ell$. For every considered segment $[\sigma : \tau]$, estimates $\mathsf{h} \approx \widetilde{\mathsf{H}}([1 : \sigma - 1], \sigma)$ and $\mathsf{v} \approx \widetilde{\mathsf{V}}(\tau, [\tau + 1 : K])$ are good enough.

*Base Case.* Before the initial stage $\ell = 0$, we just consider a "universal" candidate set $\mathcal{C}_0 := [1 : K]$ and a single "universal" segment $[\sigma : \tau] = [1 : K]$. Therefore, (Phase 2 of GBB-OneBit) estimates $\mathsf{h} = 0 = \widetilde{\mathsf{H}}(\varnothing, \sigma) = \widetilde{\mathsf{H}}([1 : \sigma - 1], \sigma)$ and $\mathsf{v} = 0 = \widetilde{\mathsf{V}}(\tau, \varnothing) = \widetilde{\mathsf{V}}(\tau, [\tau + 1 : K])$ are perfect.

*Induction Step.* In a specific stage $\ell \in [0 : L]$, we aim at locating $a_{\mu,\mu}$ more accurately $\mathcal{C}_{\ell+1} \subseteq \mathcal{C}_\ell$, by leveraging one-bit feedback, and retain Induction Hypothesis for the next stage $\ell + 1 \in [1 : L + 1]$.

For every considered segment $[\sigma : \tau]$ (cf. the two *red* ○'s in Figure 2), we *can* obtain (Lines 2 and 3) the following good enough estimates for trade rates $\{\mathsf{Z}_{i,\sigma}\}_{i\in[\sigma:\tau]} \cup \{\mathsf{Z}_{\tau,j}\}_{j\in[\sigma:\tau]}$ (cf. the *red* horizontal/vertical lines in Figure 2).

$$\begin{aligned} \widehat{\mathsf{Z}}_{i,\sigma} &\approx \mathsf{Z}_{i,\sigma}, && \forall i \in [\sigma : \tau], \\ \widehat{\mathsf{Z}}_{\tau,j} &\approx \mathsf{Z}_{\tau,j}, && \forall j \in [\sigma : \tau]. \end{aligned}$$

Also, for these candidates $[\sigma : \tau]$, especially the survival ones $k \in \mathcal{C}_\ell \bigcap [\sigma : \tau]$, we *can* obtain (Lines 4 and 5) the following good enough estimates for terms $\widetilde{\mathsf{H}}([\sigma : k], \sigma)$ and $\widetilde{\mathsf{V}}(\tau, [k : \tau])$.

$$\begin{aligned} \widehat{\mathsf{H}}([\sigma : k], \sigma) &= \tfrac{1}{K} \textstyle\sum_{i\in[\sigma:k]} \widehat{\mathsf{Z}}_{i,\sigma} \approx \widetilde{\mathsf{H}}([\sigma : k], \sigma), && \forall k \in \mathcal{C}_\ell \bigcap [\sigma : \tau], \\ \widehat{\mathsf{V}}(\tau, [k : \tau]) &= \tfrac{1}{K} \textstyle\sum_{j\in[k:\tau]} \widehat{\mathsf{Z}}_{\tau,j} \approx \widetilde{\mathsf{V}}(\tau, [k : \tau]), && \forall k \in \mathcal{C}_\ell \bigcap [\sigma : \tau]. \end{aligned}$$

Provided that (Induction Hypothesis) estimates $\mathsf{h} \approx \widetilde{\mathsf{H}}([1 : \sigma - 1], \sigma)$ and $\mathsf{v} \approx \widetilde{\mathsf{V}}(\tau, [\tau + 1 : K])$ are also good enough, we can add either of them to the above ones, thus good enough estimates for terms $\widetilde{\mathsf{H}}([1 : k], \sigma) = \mathsf{H}(a_{k,\sigma}) \pm K^{-1}$ and $\widetilde{\mathsf{V}}(\tau, [k : K]) = \mathsf{V}(a_{\tau,k}) \pm K^{-1}$, $\forall k \in \mathcal{C}_\ell \bigcap [\sigma : \tau]$. In regard to the independence of values $(S, B) \sim \mathcal{D}_S \bigotimes \mathcal{D}_B$ and that a candidate $a_{k,k}$ always induces negligible profit $\mathsf{Profit}(a_{k,k}) = \pm K^{-1}$, we *can* obtain (Lemma 2 and Line 3) the following good enough estimates $\widehat{\mathsf{GFT}}_\ell[k] \approx \mathsf{GFT}(a_{k,k})$ for Gains from Trade from survival candidates $k \in \mathcal{C}_\ell \bigcap [\sigma : \tau]$ in the considered segment.

$$\begin{aligned} \widehat{\mathsf{GFT}}_\ell[k] &= (\mathsf{h} + \widehat{\mathsf{H}}([\sigma : k], \sigma)) \cdot \tfrac{\widehat{\mathsf{Z}}_{\tau,k}}{\widehat{\mathsf{Z}}_{\tau,\sigma}} + (\mathsf{v} + \widehat{\mathsf{V}}(\tau, [k : \tau])) \cdot \tfrac{\widehat{\mathsf{Z}}_{k,\sigma}}{\widehat{\mathsf{Z}}_{\tau,\sigma}} \\ &\approx \widetilde{\mathsf{H}}([1 : K], \sigma) \cdot \tfrac{\mathsf{Z}_{\tau,k}}{\mathsf{Z}_{\tau,\sigma}} + \widetilde{\mathsf{V}}(\tau, [k : K]) \cdot \tfrac{\mathsf{Z}_{k,\sigma}}{\mathsf{Z}_{\tau,\sigma}} \\ &\approx \mathsf{H}(a_{k,k}) + \mathsf{V}(a_{k,k}) \\ &\approx \mathsf{GFT}(a_{k,k}) \end{aligned}$$

Over all of the up to $2^\ell$ many disjoint segments $[\sigma : \tau]$, whose union covers $\mathcal{C}_\ell$ (Induction Hypothesis), we *do* obtain good enough estimates $\widehat{\mathsf{GFT}}_\ell[k] \approx \mathsf{GFT}(a_{k,k})$, for all survival candidates $k \in \mathcal{C}_\ell$ in the current stage $\ell \in [0 : L]$. Then, we *do* locate (Line 7) the optimal candidate $a_{\mu,\mu}$ more accurately $\mathcal{C}_{\ell+1} \subseteq \mathcal{C}_\ell$.

Moreover, we need to retain Induction Hypothesis for the next stage $\ell + 1 \in [1 : L + 1]$. To this end, we simply follow the *divide-and-conquer* principle:
(Lines 8 and 9) For every considered segment $[\sigma : \tau]$, find two disjoint *half-segments* $[\sigma' : \tau']$ and $[\sigma'' : \tau'']$ to cover the new survival candidates $\mathcal{C}_{\ell+1} \bigcap [\sigma : \tau]$ therein (cf. the two *blue* ○'s and the two *green* ○'s in Figure 2). Clearly, there are up to $2^{\ell+1}$ many such half-segments in total, and their union covers the new candidate set $\mathcal{C}_{\ell+1}$.
(Lines 11 to 13 and 15 to 17) For every half-segment $[\sigma' : \tau']$, i.e., a new segment for the next stage $\ell + 1 \in [1 : L + 1]$, obtain new good enough estimates $\mathsf{h}' \approx \widetilde{\mathsf{H}}([1 : \sigma' - 1], \sigma')$ and $\mathsf{v}' \approx \widetilde{\mathsf{V}}(\tau', [\tau' + 1 : K])$, in a similar manner as the above; similarly for every other half-segment $[\sigma'' : \tau'']$.
(Lines 14 and 18) Move on to the next stage $\ell + 1 \in [1 : L + 1]$.

In sum, initially there are $|\mathcal{C}_0| = K = \widetilde{\Theta}(T^{2/3})$ many candidates. After all the $L + 1 \approx \log(K)$ many stages, the ultimate candidates $\mathcal{C}_{L+1}$ all will be good enough, compared with the optimal candidate $a_{\mu,\mu}$ or even the optimal action $a^*$ in the whole action space.

▶ **Proposition 5** (FractalElimination Guarantee, Informal)**.** *For $K = \widetilde{\Theta}(T^{1/3})$, the subroutine* FractalElimination *uses one-bit feedback and, with high probability, returns a candidate $a_{k,k}$ whose expected gains from trade is within $\widetilde{O}(T^{-1/3})$ of the optimal fixed price. Moreover, during the subroutine, the total regret and the total budget deficit are both bounded by $\widetilde{O}(T^{2/3})$.*

The proof proceeds by a stage-wise elimination argument. At each stage, the remaining candidates lie on short diagonal segments. Instead of estimating each candidate separately, the algorithm queries carefully chosen horizontal and vertical segments so that many candidates share the same observations. The independence of $S$ and $B$ allows us to express the required GFT estimates through products of ratios, obtained recursively from one-bit feedback. This yields both the accuracy invariant and the regret invariant. The full induction and concentration bounds are deferred to the full version.

## 3.2 Performance Analysis

**Proof sketch of Theorem 1.** Phase 1 invokes ProfitMax to accumulate $\widetilde{\Omega}(T^{2/3})$ budget surplus while incurring $\widetilde{O}(T^{2/3})$ regret. Phase 2 invokes FractalElimination; by Proposition 5, it identifies an $\widetilde{O}(T^{-1/3})$-optimal diagonal candidate while spending only the surplus accumulated in Phase 1. Phase 3 exploits any remaining candidate. Combining the regret from the three phases gives $\widetilde{O}(T^{2/3})$, and the budget surplus from Phase 1 ensures GBB. ◀

# 4 $\Omega(T^{3/4})$ Partial-Feedback Lower Bound for Correlated Values

In this section, we study the limit of "fixed-price mechanisms with the Global Budget Balance (GBB) constraint and partial feedback" in the "correlated values" setting. Specifically, we will establish (Theorem 6) the following hardness result. By implication, the same lower bound extends to the *more general* "adversarial values" setting.

▶ **Theorem 6** (GBB Partial-Feedback Lower Bound for Correlated Values). *In the "correlated values" setting (with or without the density-boundedness assumption – Assumption 1 with parameter $M = 224$),*
*every "GBB fixed-price mechanism with two-bit feedback" has worst-case regret $\Omega(T^{3/4})$.*

Previously, in both the "correlated values" setting and the "adversarial values" setting, merely an $\widetilde{\mathcal{O}}(T^{3/4})$ upper bound [3, Theorem 5.4] and an unmatching $\Omega(T^{5/7})$ lower bound [3, Theorem 5.5] were known. Nonetheless, our hardness result closes this gap by (up to polylogarithmic factors) establishing a matching $\Omega(T^{3/4})$ lower bound.

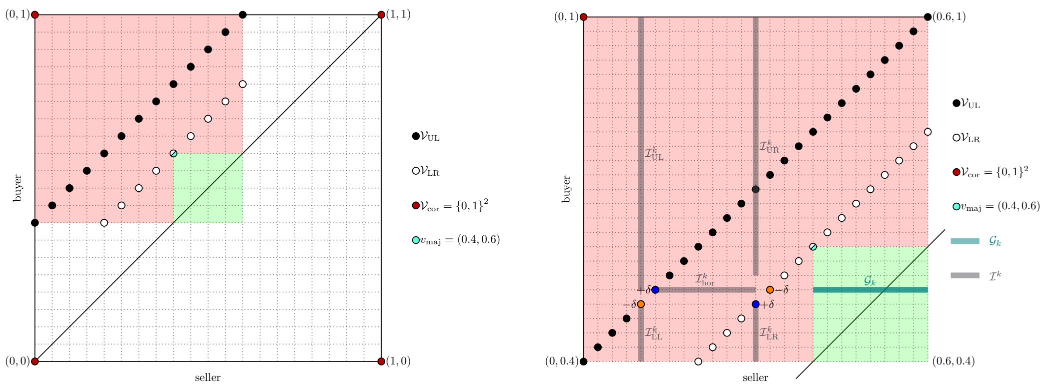
We will only establish the "discrete values" version of Theorem 6 in Sections 4.1 and 4.2; the "density-bounded values" version can be found in the full version.

## 4.1 Lower-Bound Construction for Discrete Values

Our construction utilizes two parameters $K = \Theta(T^{1/4})$ and $\delta = \Theta(T^{-1/4})$, as follows.

$$\begin{aligned} K &:= T^{1/4}, \\ \delta &:= \tfrac{0.1}{5K+2}. \end{aligned}$$

**The Value Support.** We will construct $(K+1)$ base/hard instances $\{\mathcal{D}^k\}_{k\in[0:K]}$, which have a common *discrete* support $\mathcal{V} \subseteq [0,1]^2$, including four types of points that serve different purposes – a size-$(3K+1)$ "upper-left" subset $\mathcal{V}_{\mathrm{UL}}$, a size-$(2K+1)$ "lower-right" subset $\mathcal{V}_{\mathrm{LR}}$, four "corner" points $\mathcal{V}_{\mathrm{cor}} = \{0,1\}^2$, and one "majority" point $v_{\mathrm{maj}} = (0.4, 0.6)$; see Figure 3a for a diagram.




**(a)** The support $\mathcal{V} \subseteq [0,1]^2$.
**(b)** The informative action subsets $\mathcal{I}^k$.

**Figure 3** Diagrams of our new $\Omega(T^{3/4})$ lower-bound construction (Theorem 6).

$$\begin{aligned}
\mathcal{V} &:= \mathcal{V}_{\mathrm{UL}} \cup \mathcal{V}_{\mathrm{LR}} \cup \mathcal{V}_{\mathrm{cor}} \cup \{v_{\mathrm{maj}}\}, \\
\mathcal{V}_{\mathrm{UL}} &:= \{(\tfrac{k}{5K},\ 0.4 + \tfrac{k}{5K})\}_{k \in [0:3K]}, \\
\mathcal{V}_{\mathrm{LR}} &:= \{(0.2 + \tfrac{k}{5K},\ 0.4 + \tfrac{k}{5K})\}_{k \in [0:2K]}, \\
\mathcal{V}_{\mathrm{cor}} &:= \{0,1\}^2, \\
v_{\mathrm{maj}} &:= (0.4, 0.6).
\end{aligned}$$

Note that the majority point $v_{\mathrm{maj}}$ is also the index-$K$ point in the lower-right subset $\mathcal{V}_{\mathrm{LR}}[K] = (0.4, 0.6)$; a base/hard instance $\mathcal{D}^k$ assigns two probability masses to this point, one for $v_{\mathrm{maj}}$ and one for $\mathcal{V}_{\mathrm{LR}}[k]$. However, we can safely treat it as *two* isolated points (or, interchangeably, treat it as *one* point by adding both probability masses together). All other points $v \in \mathcal{V} \setminus \{v_{\mathrm{maj}}\}$ are isolated.

**The Base/Hard Instances.** Among the $K+1$ instances, $\mathcal{D}^0$ is the base instance and is given as follows.

$$\mathbb{P}_{(S,B)\sim\mathcal{D}^0}[(S,B) = v] \ := \begin{cases} \delta, & v \in \mathcal{V}_{\mathrm{UL}} \cup \mathcal{V}_{\mathrm{LR}} \\ 0.1, & v \in \mathcal{V}_{\mathrm{cor}} \\ 0.5, & v = v_{\mathrm{maj}} \end{cases}.$$

This $\mathcal{D}^0$ is a well-defined distribution, given that $\delta \cdot |\mathcal{V}_{\mathrm{UL}} \cup \mathcal{V}_{\mathrm{LR}}| + 0.1 \cdot |\mathcal{V}_{\mathrm{cor}}| + 0.5 \cdot |\{v_{\mathrm{maj}}\}| = 1$.

In contrast, each $\mathcal{D}^k$ for $k \in [K]$ is a hard instance and tweaks the probability masses at two upper-left points $\mathcal{V}_{\mathrm{UL}}[k-1]$, $\mathcal{V}_{\mathrm{UL}}[k]$ and two lower-right points $\mathcal{V}_{\mathrm{LR}}[k-1]$, $\mathcal{V}_{\mathrm{LR}}[k]$ (but otherwise is identical to the base instance $\mathcal{D}^0$); see Figure 3b for a diagram. Again, this $\mathcal{D}^k$ is a well-defined distribution.

$$\mathbb{P}_{(S,B)\sim\mathcal{D}^k}[(S,B) = v] \ := \begin{cases} \mathbb{P}_{(S,B)\sim\mathcal{D}^0}[(S,B) = v] + \delta, & v \in \{\mathcal{V}_{\mathrm{UL}}[k],\ \mathcal{V}_{\mathrm{LR}}[k-1]\} \\ \mathbb{P}_{(S,B)\sim\mathcal{D}^0}[(S,B) = v] - \delta, & v \in \{\mathcal{V}_{\mathrm{UL}}[k-1],\ \mathcal{V}_{\mathrm{LR}}[k]\} \\ \mathbb{P}_{(S,B)\sim\mathcal{D}^0}[(S,B) = v], & \text{otherwise} \end{cases}.$$

## 4.2 Lower-Bound Analysis for Discrete Values

For the base/hard instances $\{\mathcal{D}^k\}_{k\in[0:K]}$ above, we will establish an $\Omega(T^{3/4})$ lower bound in three steps:

- Firstly, we show that a mechanism $\mathcal{M}$'s actions $(P^t, Q^t)_{t\in[T]}$, without loss of generality, can be restricted to a particular discrete set $\mathcal{A}$ (or, more precisely, a $(3K+1)\times(3K+1)$ grid).
- Secondly, we show that the Global Budget Balance constraint can be relaxed to another constraint (which we call Global Price Balance); the $\Omega(T^{3/4})$ lower bound holds even after this relaxation.
- Thirdly, we establish an indistinguishability argument: any algorithm must take sufficiently many informative actions in order to distinguish between hard instances.

## Step 1: Discretization of Actions

It turns out that a *regret-optimal* mechanism $\mathcal{M} = (P^t, Q^t)_{t\in[T]}$ can take actions $(P^t, Q^t)$ only from $\mathcal{A}$, the Cartesian product of the coordinate projections – excluding the "trivial" seller value of 1 and the "trivial" buyer value of 0 – of the support $\mathcal{V} \subseteq [0,1]^2$; cf. the union of the *red/green regions* in Figure 3a.

$$\begin{aligned}\mathcal{A} &:= (\mathrm{proj}_S(\mathcal{V})\setminus\{1\})\times(\mathrm{proj}_B(\mathcal{V})\setminus\{0\})\\ &= \{\tfrac{k}{5K}\}_{k\in[0:3K]}\times\{0.4+\tfrac{k}{5K}\}_{k\in[0:3K]}.\end{aligned}$$

The explicit transformation is provided in the full version.

▶ **Lemma 7** (Discretization of Actions). *A (generic) fixed-price mechanism $\overline{\mathcal{M}} = (\overline{P}^t, \overline{Q}^t)_{t\in[T]}$ can transform into a new fixed-price mechanism $\mathcal{M} = (P^t, Q^t)_{t\in[T]}$ such that, in any possibility $\mathcal{D} = \mathcal{D}^k$ for $k \in [0:K]$:*

- *$(P^t, Q^t)_{t\in[T]} \subseteq \mathcal{A}$, almost surely.*
- *On the same realization $(S^t, B^t)_{t\in[T]} \sim \mathcal{D}^k$, almost surely over the randomness of $\overline{\mathcal{M}}$ and $\mathcal{M}$,*
  *both the Gains from Trade and the profit in every round $t \in [T]$ can only increase:*

$$\begin{aligned}\mathsf{GFT}(S^t, B^t, P^t, Q^t) &\geq \mathsf{GFT}(S^t, B^t, \overline{P}^t, \overline{Q}^t), && \forall t\in[T],\\ \mathsf{Profit}(S^t, B^t, P^t, Q^t) &\geq \mathsf{Profit}(S^t, B^t, \overline{P}^t, \overline{Q}^t), && \forall t\in[T].\end{aligned}$$

In the rest of Section 4, we safely restrict actions to the discrete set $\mathcal{A}$. To emphasize this, we rewrite $\mathcal{M}^{\mathcal{A}}$ for a mechanism $\mathcal{M}$. To proceed, we divide $\mathcal{A}$ into a "good" subset $\mathcal{G}$ (cf. the *green region* in Figure 3a) and a "bad" subset $\mathcal{B}$ (cf. the *red region* in Figure 3a), as follows.

$$\begin{aligned}\mathcal{G} &:= \cup_{k\in[0:K]}\mathcal{G}_k,\\ \mathcal{G}^k &:= \{(0.4+\tfrac{i}{5K},\ 0.4+\tfrac{k}{5K})\}_{i\in[0:K]}, && \forall k\in[0:K],\\ \mathcal{B} &:= \mathcal{A}\setminus\mathcal{G}.\end{aligned}$$

Let us denote by $\mathsf{Regret}_{\mathcal{D}}(p,q)$ the regret incurred by each take of an action $(p,q) \in \mathcal{A} = \mathcal{G}\cup\mathcal{B}$. The following Proposition 8 lower-bounds $\mathsf{Regret}_{\mathcal{D}}(p,q)$ in each possibility $\mathcal{D} = \mathcal{D}^k$ for $k \in [0:K]$.

▶ **Proposition 8** (Gains from Trade). *For the base instance $\mathcal{D}^0$ and the hard instances $\mathcal{D}^k$, $\forall k \in [K]$:*

$$\mathsf{Regret}_{\mathcal{D}^0}(p,q) \;\geq\; \begin{cases} 0.1, & \forall (p,q) \in \mathcal{B} \\ 3\delta K \cdot (q-p), & \forall (p,q) \in \mathcal{G} \end{cases},$$

$$\mathsf{Regret}_{\mathcal{D}^k}(p,q) \;\geq\; \begin{cases} 0.1, & \forall (p,q) \in \mathcal{B} \\ 3\delta K \cdot (q-p) + 0.2\delta \cdot \mathbb{1}[(p,q) \notin \mathcal{G}^k], & \forall (p,q) \in \mathcal{G} \end{cases}.$$

## Step 2: Relaxation of Global Budget Balance

We now show that, under our lower-bound construction, a GBB mechanism $\mathcal{M}^{\mathcal{A}}$ always satisfies the following "Global Price Balance" condition; thus, it serves as a relaxation of the GBB constraint.

▶ **Lemma 9** (Relaxation of GBB). *For a GBB fixed-price mechanism $\mathcal{M}^{\mathcal{A}}$, in each possibility $k \in [0 : K]$:*

$$\mathbb{E}^k[\textstyle\sum_{t\in[T]}(Q^t - P^t)] \;\geq\; 0. \tag{Global Price Balance}$$

## Step 3: Indistinguishability Argument

Regarding each hard instance $\mathcal{D}^k$ for $k \in [K]$, we define its *informative action subset* $\mathcal{I}^k \subseteq \mathcal{B}$ – only taking such actions can help in distinguishing this $\mathcal{D}^k$ from the other base/hard instances. The following lemma states that any algorithm behaves similarly on instances $\mathcal{D}^0$ and $\mathcal{D}^k$, if the informative action subset $\mathcal{I}^k$ is explored only a sufficiently small number of times; we refer the reader to the full version for the explicit definition of $\mathcal{I}^k$.

▶ **Lemma 10.** *In each possibility $k \in [K]$:*

$$\mathbb{E}^k[T_{\mathcal{G}^k}] - \mathbb{E}^0[T_{\mathcal{G}^k}] \;\leq\; T \cdot \sqrt{15\delta^2 \mathbb{E}^0[T_{\mathcal{I}^k}]}$$

Finally, we are ready to accomplish Theorem 6.

**Proof of Theorem 6.** Recall that we divide $\mathcal{A}$ into two disjoint subsets $\mathcal{A} = \mathcal{G} \cup \mathcal{B}$. For the base instance $\mathcal{D}^0$, each take of a bad action $(P^t, Q^t) \in \mathcal{B}$ incurs at least 0.1 regret, and each take of a good action $(P^t, Q^t) \in \mathcal{G}$ incurs $3\delta K \cdot (Q^t - P^t)$ regret (Proposition 8), so the total regret is at least

$$\begin{aligned}
\mathsf{Regret}_{\mathcal{D}^0} \;&=\; \mathbb{E}^0\big[\textstyle\sum_{t\in[T]} \big(\mathbb{1}[(P^t,Q^t)\in\mathcal{B}] \cdot \mathsf{Regret}_{\mathcal{D}^0}(P^t,Q^t) \\
&\qquad\qquad + \mathbb{1}[(P^t,Q^t)\in\mathcal{G}] \cdot \mathsf{Regret}_{\mathcal{D}^0}(P^t,Q^t)\big)\big] \\
\triangleright \text{ Proposition 8} \qquad &\geq\; 0.1 \cdot \mathbb{E}^0[T_{\mathcal{B}}] \;+\; 3\delta K \cdot \mathbb{E}^0\big[\textstyle\sum_{t\in[T]}(Q^t-P^t)\cdot \mathbb{1}[(P^t,Q^t)\in\mathcal{G}]\big] \\
\triangleright \text{ Lemma 9 and } \mathcal{B} = \mathcal{A}\setminus\mathcal{G} \qquad &\geq\; 0.1 \cdot \mathbb{E}^0[T_{\mathcal{B}}] \;-\; 3\delta K \cdot \mathbb{E}^0\big[\textstyle\sum_{t\in[T]}(Q^t-P^t)\cdot \mathbb{1}[(P^t,Q^t)\in\mathcal{B}]\big] \\
\triangleright\; Q^t - P^t \leq 1 \qquad &\geq\; 0.1 \cdot \mathbb{E}^0[T_{\mathcal{B}}] \;-\; 3\delta K \cdot \mathbb{E}^0[T_{\mathcal{B}}] \\
\triangleright\; \delta K = \tfrac{0.1K}{5K+2} \leq 0.02 \qquad &\geq\; 0.04 \cdot \mathbb{E}^0[T_{\mathcal{B}}].
\end{aligned}$$

For each hard instance $\mathcal{D}^k$, $\forall k \in [K]$, a moment's reflection will show that the deduction above for $\mathcal{D}^0$ still holds, and each take of an action $(P^t, Q^t) \in \mathcal{G} \setminus \mathcal{G}^k$ will incur $0.2\delta$ more regret (Proposition 8). Thus, we can obtain

$$\begin{aligned}
\mathsf{Regret}_{\mathcal{D}^k} \;&\geq\; 0.04 \cdot \mathbb{E}^k[T_{\mathcal{B}}] \;+\; \mathbb{E}^k\big[\textstyle\sum_{t\in[T]} 0.2\delta \cdot \mathbb{1}[(P^t, Q^t) \in \mathcal{G} \setminus \mathcal{G}^k]\big] \\
&=\; 0.04 \cdot \mathbb{E}^k[T_{\mathcal{B}}] \;+\; 0.2\delta \cdot \big(\mathbb{E}^k[T_{\mathcal{G}}] - \mathbb{E}^k[T_{\mathcal{G}^k}]\big) \\
\triangleright \text{ Lemma 10} \qquad &\geq\; 0.04 \cdot \mathbb{E}^k[T_{\mathcal{B}}] \;+\; 0.2\delta \cdot \big(T - \mathbb{E}^k[T_{\mathcal{B}}] - \mathbb{E}^0[T_{\mathcal{G}^k}] - 4\delta T \cdot \sqrt{\mathbb{E}^0[T_{\mathcal{I}^k}]}\big) \\
\triangleright\, \delta \in [0, \tfrac{1}{20}] \qquad &\geq\; 0.2\delta \cdot \big(T - \mathbb{E}^0[T_{\mathcal{G}^k}] - 4\delta T \cdot \sqrt{\mathbb{E}^0[T_{\mathcal{I}^k}]}\big).
\end{aligned}$$

Taking the average of $\mathsf{Regret}_{\mathcal{D}^k}$'s for $k \in [K]$ gives

$$\begin{aligned}
\tfrac{1}{K} \cdot \textstyle\sum_{k\in[K]} \mathsf{Regret}_{\mathcal{D}^k} \;&\geq\; \tfrac{1}{K} \cdot \textstyle\sum_{k\in[K]} 0.2\delta \cdot \big(T - \mathbb{E}^0[T_{\mathcal{G}^k}] - 4\delta T \cdot \sqrt{\mathbb{E}^0[T_{\mathcal{I}^k}]}\big) \\
\triangleright \textstyle\sum_{k\in[K]} T_{\mathcal{G}^k} \leq T \text{ (a.s.)} \qquad &\geq\; 0.2\delta \cdot \big(T - \tfrac{T}{K} - \tfrac{4\delta T}{K} \cdot \textstyle\sum_{k\in[K]} \sqrt{\mathbb{E}^0[T_{\mathcal{I}^k}]}\big) \\
\triangleright \text{ Cauchy-Schwarz inequality} \qquad &\geq\; 0.2\delta \cdot \big(T - \tfrac{T}{K} - 4\delta T \cdot \sqrt{\tfrac{1}{K} \cdot \textstyle\sum_{k\in[K]} \mathbb{E}^0[T_{\mathcal{I}^k}]}\big) \\
\triangleright \textstyle\sum_{k\in[K]} T_{\mathcal{I}^k} \leq 2T_{\mathcal{B}} \text{ (a.s.)} \qquad &\geq\; 0.2\delta \cdot \big(T - \tfrac{T}{K} - 4\delta T \cdot \sqrt{\tfrac{2}{K} \cdot \mathbb{E}^0[T_{\mathcal{B}}]}\big).
\end{aligned}$$

Here the last step uses $\sum_{k\in[K]} T_{\mathcal{G}^k} \leq T$, as a consequence of that $\mathcal{G}^k$'s for $k \in [K]$ are disjoint. And the last step uses $\sum_{k\in[K]} T_{\mathcal{I}^k} \leq 2T_{\mathcal{B}}$.

Plugging in $K = T^{1/4}$ and $\delta = \frac{0.1}{5K+2}$, it is easy to verify through elementary algebra that

$$\begin{aligned}
&\max\big\{\mathsf{Regret}_{\mathcal{D}^0},\ \tfrac{1}{K} \cdot \textstyle\sum_{k\in[K]} \mathsf{Regret}_{\mathcal{D}^k}\big\} \\
&\geq\; \max\Big\{0.04 \cdot \mathbb{E}^0[T_{\mathcal{B}}],\ 0.004 \cdot T^{3/4} \cdot \big(\tfrac{1-T^{-1/4}}{1+0.4T^{-1/4}} - \tfrac{0.08}{(1+0.4T^{-1/4})^2} \cdot \sqrt{\tfrac{2\cdot\mathbb{E}^0[T_{\mathcal{B}}]}{T^{3/4}}}\big)\Big\} \\
&=\; \Omega(T^{3/4}).
\end{aligned}$$

In sum, the considered mechanism $\mathcal{M}^{\mathcal{A}}$ incurs $\Omega(T^{3/4})$ regret in at least one possibility $k \in [0 : K]$. Then, the arbitrariness of $\mathcal{M}^{\mathcal{A}}$ implies Theorem 6. ◀

## References


1 Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The non-stochastic multiarmed bandit problem. *SIAM J. Comput.*, 32(1):48–77, 2002. `doi:10.1137/S0097539701398375`.

2 Yossi Azar, Amos Fiat, and Federico Fusco. An $\alpha$-regret analysis of adversarial bilateral trade. *Artif. Intell.*, 337:104231, 2024. `doi:10.1016/J.ARTINT.2024.104231`.

3 Martino Bernasconi, Matteo Castiglioni, Andrea Celli, and Federico Fusco. No-regret learning in bilateral trade via global budget balance. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 247–258. ACM, 2024. `doi:10.1145/3618260.3649653`.

4 Avrim Blum, Tuomas Sandholm, and Martin Zinkevich. Online algorithms for market clearing. *J. ACM*, 53(5):845–879, 2006. `doi:10.1145/1183907.1183913`.

5 Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. Bilateral trade: A regret minimization perspective. *Math. Oper. Res.*, 49(1):171–203, 2024. `doi:10.1287/MOOR.2023.1351`.

6 Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. Regret analysis of bilateral trade with a smoothed adversary. *J. Mach. Learn. Res.*, 25:234:1–234:36, 2024. URL: `https://jmlr.org/papers/v25/23-1627.html`.

7 Houshuang Chen, Yaonan Jin, Pinyan Lu, and Chihao Zhang. Tight regret bounds for fixed-price bilateral trade. *arXiv preprint arXiv:2504.04349*, 2025. `doi:10.48550/arXiv.2504.04349`.

**8** Shahar Dobzinski and Ariel Shaulker. Bilateral trade with correlated values. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 237–246. ACM, 2024. doi:10.1145/3618260.3649659.

**9** Simone Di Gregorio, Paul Dütting, Federico Fusco, and Chris Schwiegelshohn. Nearly tight regret bounds for profit maximization in bilateral trade. *arXiv preprint arXiv:2509.22563*, 2025. To appear in FOCS 2025. doi:10.48550/arXiv.2509.22563.

**10** Anna Lunghi, Matteo Castiglioni, and Alberto Marchesi. Better regret rates in bilateral trade via sublinear budget violation. *arXiv preprint arXiv:2507.11419*, 2025. To appear in SODA 2026. doi:10.48550/arXiv.2507.11419.

**11** Anna Lunghi, Matteo Castiglioni, and Alberto Marchesi. Online two-sided markets: Many buyers enhance learning. *arXiv preprint arXiv:2503.01529*, 2025. doi:10.48550/arXiv.2503.01529.

**12** Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983.

**13** William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37, 1961. doi:10.1111/j.1540-6261.1961.tb02789.x.

## A Additional Lower Bounds

▶ **Theorem 11** (GBB Semi-Feedback Lower Bound for Independent Values). *In the "independent values" setting (with or without the density-boundedness assumption – Assumption 1 with parameter $M = 11$), every "GBB semi-feedback fixed-price mechanism" has worst-case regret $\Omega(T^{2/3})$.*

▶ **Theorem 12** (GBB Full-Feedback Lower Bound for Independent Values). *In the "independent values" setting (with or without the density-boundedness assumption – Assumption 1 with parameter $M = 11$), every "GBB full-feedback fixed-price mechanism" has worst-case regret $\Omega(T^{1/2})$.*

▶ **Theorem 13** (WBB Partial-Feedback Lower Bound for Independent Values [5, Theorem 6]). *In the "independent values" setting, every "WBB fixed-price mechanism with two-bit feedback" has worst-case regret $\Omega(T)$.*