BILATERAL TRADE WITH INDEPENDENT SELLERS AND BUYERS: TIGHT REGRET UNDER ONE-BIT FEEDBACK

Anonymous authors
Paper under double-blind review

ABSTRACT

We study regret minimization in repeated bilateral trade under one-bit feedback: over T rounds, a learner posts a price to a seller and a price to a buyer to maximize gains-from-trade, but observes only whether trade occurs. When the two valuations are independent and drawn i.i.d. from an unknown distribution with bounded density, the tight regret was still open in two settings: weak budget balance against the best fixed price, and global budget balance against the best feasible distribution over price pairs. In both settings, the best known upper bound was $\widetilde{\mathcal{O}}(T^{3/4})$, against a lower bound of $\Omega(T^{2/3})$.

We close both gaps by designing two learning algorithms, one for each setting, each achieving regret $\widetilde{\mathcal{O}}(T^{2/3})$. Combined with the known $\Omega(T^{3/4})$ lower bound for correlated valuations, our results reveal that independence makes learning strictly easier in both settings. Our key technical ingredient is a recursion that relies on valuation independence, enabling our algorithms to learn gains-from-trade from trade bits alone while probing only inside the price interval being explored.

1 INTRODUCTION

Bilateral trade is a fundamental problem in economics, in which a broker intermediates between two rational agents: a seller who owns an indivisible item and a buyer who wants it. Each agent holds a private valuation for the item, and in the classic setting the two valuations are *independent* (Myerson & Satterthwaite, 1983). The broker aims to design a mechanism that is efficient: it maximizes social welfare by making trade occur whenever the buyer values the item more than the seller. However, the landmark result of Myerson & Satterthwaite (1983) shows that no mechanism can achieve this while still ensuring that both agents report truthfully, that both are willing to participate, and that the broker never loses money.

To bypass the fundamental limitations of this single-shot setting, a recent line of research (Blum et al., 2006; Cesa-Bianchi et al., 2024b;a; Azar et al., 2024; Bernasconi et al., 2024; Lunghi et al., 2025; Chen et al., 2026) studies *repeated* bilateral trade with fixed-price mechanisms, in which each agent simply accepts or rejects a posted price, so both agents are automatically truthful and willing to participate (Hagerty & Rogerson, 1987). In this framework, the broker faces a sequence of seller–buyer pairs over T rounds. In each round t, a new seller and a new buyer arrive with private valuations s_t and b_t, and the broker, without observing the valuations, posts a price p_t to the seller and a price q_t to the buyer. Trade happens exactly when both agents accept, that is, when $s_t \le p_t$ and $b_t \ge q_t$. The broker then collects q_t from the buyer and pays p_t to the seller, keeping $q_t - p_t$ as profit, while the trade generates gains-from-trade (GFT) $b_t - s_t$, which is the increase in social welfare the trade produces.

The broker can learn from past prices and trade outcomes, and typically wants to maximize the cumulative GFT subject to a constraint on the profit. The constraints typically include, in decreasing order of strictness, strong budget balance (SBB), which requires $p_t = q_t$ in every round, and weak budget balance (WBB), which requires $p_t \le q_t$ in every round. In the *adversarial* setting, where the valuations may be chosen by an adversary, no SBB or WBB mechanism achieves sublinear regret against the best fixed price, even with full feedback, in which the broker observes both valuations after each round (Cesa-Bianchi et al., 2024b; Azar et al., 2024). This impossibility motivates the *stochastic* setting, where the valuations are drawn i.i.d. from a fixed distribution; it turns out that this relaxation, which is reasonable in practice, makes learning easier. Under WBB, sublinear regret against the best fixed price becomes possible once the distribution has a bounded density: Cesa-Bianchi et al. (2024b) obtain $\widetilde{\mathcal{O}}(T^{3/4})$ regret with one-bit feedback. This result holds even for correlated valuations, where the seller's and buyer's valuations may depend on each other; thus, it also holds for independent

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Table 1: Regret bounds for independent valuations with bounded densities under one-bit feedback.

Setting	Previous work		This paper
	Lower bound	Upper bound	
WBB best fixed price	$\Omega\left(T^{2/3}\right)$ Cesa-Bianchi et al., 2024b	$\widetilde{\mathcal{O}}(T^{3/4})$ Cesa-Bianchi et al., 2024b	$\widetilde{\Theta}(T^{2/3})$ [Thm. 1]
GBB best feasible distribution	$\Omega\left(T^{2/3}\right)$ Chen et al., 2026	$\widetilde{\mathcal{O}}(T^{3/4})$ Lunghi et al., 2026b	$\widetilde{\Theta}(T^{2/3})$ [Thm. 2]

Note: Cesa-Bianchi et al. (2024b) state their WBB bounds for the agents' total utility, which equals GFT at common prices; Appendix A explains why both bounds carry over to our benchmark.

valuations. They also show that no fixed-price algorithm can avoid $\Omega(T^{2/3})$ regret for independent valuations. It has remained unknown whether independent valuations can make learning easier by allowing a regret bound better than $\widetilde{\mathcal{O}}(T^{3/4})$.

Instead of restricting the valuations, one can also relax the budget constraint to escape the adversarial impossibility. Bernasconi et al. (2024) introduce global budget balance (GBB), which requires only that the cumulative profit over the T rounds be nonnegative, so that a subsidy in one round can be repaid by profit in another, and show that it allows sublinear regret against the best fixed price even for adversarial valuations. Since a GBB mechanism may also post a pair with $p_t > q_t$, they further introduce a stronger benchmark, the best distribution over price pairs with nonnegative expected profit, and show that no GBB mechanism achieves sublinear regret against it in the adversarial setting. In the stochastic setting with a bounded density, Lunghi et al. (2026b) obtain $\widetilde{\mathcal{O}}(T^{3/4})$ regret against this benchmark with one-bit feedback, again even for correlated valuations. For independent valuations, the best known lower bound is only $\Omega(T^{2/3})$ (Chen et al., 2026), leaving the same gap between $T^{2/3}$ and $T^{3/4}$ as under WBB. Motivated by these two gaps, we ask:

> *Can independent valuations beat the $\widetilde{\mathcal{O}}(T^{3/4})$ regret under WBB and GBB? If so, what is the optimal regret bound?*

1.1 Our Contribution and Techniques

We answer the above questions firmly in the affirmative. In particular, we study GFT maximization in repeated bilateral trade with one-bit feedback, where the valuations are i.i.d. draws from a product distribution with bounded density. We consider two settings: WBB against the best fixed price, and GBB against the best feasible distribution over price pairs. For each setting, we propose an algorithm with near-optimal regret $\widetilde{\mathcal{O}}(T^{2/3})$, improving the previous $\widetilde{\mathcal{O}}(T^{3/4})$ bounds of Cesa-Bianchi et al. (2024b) and Lunghi et al. (2026b), respectively, and matching the corresponding $\Omega(T^{2/3})$ lower bounds of Cesa-Bianchi et al. (2024b) and Chen et al. (2026) up to polylogarithmic factors. Table 1 compares our results with theirs.

The main challenge in algorithm design is to exploit the structure of independent valuations. The known algorithms (Cesa-Bianchi et al., 2024b; Lunghi et al., 2026b) assume only a bounded density and thus cannot use this structure: they pair the posted price with an auxiliary uniform price, which turns the trade bit into an unbiased sample of a joint probability, and then learn the value of every price on a grid from samples of that kind, which keeps their regret at $T^{3/4}$. Independence is what lifts that limitation. Under independence, the trade probability is the product of the probability that the seller accepts the seller's price and the probability that the buyer accepts the buyer's price. Holding one price fixed and varying the other therefore reveals how the latter side's acceptance probability changes, up to a fixed multiplicative factor. An algorithm can thus spend its rounds resolving two one-dimensional functions rather than a two-dimensional table of values. The freedom to post two prices is what makes this possible, and it is not a technicality, since with a single common price the trade bit reports only the chance that the price falls between the two valuations, which does not pin down the optimum even under independence and bounded density (Cesa-Bianchi et al., 2024b). Our WBB and GBB algorithms both exploit this product form through the recursive estimation method described next, but each in a different way, because the benchmarks differ: the WBB benchmark is a single common price, whereas the GBB benchmark is a distribution over price pairs whose expected profit need only be nonnegative overall, and thus individual pairs in its support may lose money.

A recursive estimation method. At a high level, we split the expected GFT of a price pair into the conditional utilities of the two sides, weighted by the trade probability, and the profit (see eq. (1)). The trade probability and the profit are estimated directly from the bit at the posted pair, so what remains is to observe the two utilities, which one bit does not reveal. We resolve this with a recursion (Lemma 1 for WBB and Lemma 2 for GBB) that, on any interval, expresses the utility at one endpoint through the utility at the other and trade probabilities alone, one of them taken at the same auxiliary uniform price, so that every term is estimable from the bit. The utilities are known exactly at the boundary: the seller's at price 0 and the buyer's at price 1. Starting from these values, the recursion lets the algorithm estimate the utilities at any price in $[0, 1]$. This propagation is the key idea of our algorithms for both settings. Since the recursion uses independence only through the product form of the trade probability, it may also be useful beyond bilateral trade.

The WBB algorithm. For the WBB constraint, the recursion lets the algorithm refine a dyadic tree of prices while probing only inside the interval being refined. Pruning keeps an interval only while its midpoint loses at most a constant times its width against the optimal GFT, so a round spent at depth j costs $\mathcal{O}(2^{-j})$ regret rather than a constant, which cuts exploration regret from the $\widetilde{\mathcal{O}}(K^3)$ of a grid at resolution $1/K$ (Cesa-Bianchi et al., 2024b) to $\widetilde{\mathcal{O}}(K^2)$ and the rate from $T^{3/4}$ to $T^{2/3}$.

The GBB algorithm. For the GBB constraint, the algorithm learns the valuation law itself instead of the profit of each grid pair, which removes the K^2-armed profit bandit behind the previous $T^{3/4}$ bound (Lunghi et al., 2026b). What remains is to learn the law cheaply from a single bit. Chen et al. (2026) also learn from one bit under independence, but their fractal elimination discards candidate prices near the diagonal, where the best fixed price lies and the profit is negligible, by estimated GFT alone. A distribution may instead mix off-diagonal pairs that break even only jointly, so the profit must be learned alongside the GFT. For every pair (p, q) in the support of the optimistic distribution the algorithm plays, the recursion lays out two dyadic paths of local conditional probabilities, and their uncertainties control both the GFT and the profit of the pair. These statistics are shared by every pair whose paths pass through a node, so the algorithm refines nodes rather than learning pairs one at a time. The width argument of the WBB tree then carries over: probing a node I on a path of (p, q) moves one price of (p, q) by less than $2\,|I|$ and costs $\mathcal{O}(|I|)$ regret per round. Each node is refined only logarithmically often and the nodes at each depth have total width $\mathcal{O}(1)$, so refinement costs $\widetilde{\mathcal{O}}(K^2)$. Balancing this against the $\widetilde{\mathcal{O}}(T/K)$ grid error yields $\widetilde{\mathcal{O}}(T^{2/3})$.

1.2 Related Work

Regret minimization for fixed-price bilateral trade was initiated by Cesa-Bianchi et al. (2024b), who showed that posted prices achieve regret of order $T^{2/3}$ under two-bit feedback when the two valuations are independent with bounded densities. Bacchiocchi et al. (2025) show that asynchronous feedback, which lies between one-bit and two-bit feedback, suffices for the same rate. Subsequent work sought to relax the distributional assumptions: Azar et al. (2024) study α-regret against adversarial valuations, and Cesa-Bianchi et al. (2024a) consider a smoothed adversary. Global budget balance, introduced by Bernasconi et al. (2024), enables sublinear regret against the best fixed price even for adversarial valuations; Chen et al. (2026) and Jin (2026) then determine the tight rates under one- or two-bit feedback and under semi feedback, respectively, and Lunghi et al. (2026a) characterize the trade-off between regret and budget violation. Lunghi et al. (2026b) obtain sublinear regret against the stronger benchmark of Bernasconi et al. (2024), the best feasible distribution over price pairs, when the joint density is bounded, and Di Gregorio et al. (2025) turn to profit maximization. Together with our results, these works give an essentially complete picture across budget constraints, feedback models, and valuation classes, which we summarize in Appendix A.

Beyond a single seller and a single buyer, Babaioff et al. (2024) and Lunghi et al. (2025) study markets with several buyers. Other extensions consider contextual information (Gaucher et al., 2025; Feng et al., 2026), fairness objectives (Bachoc et al., 2024), and brokerage, where the traders' roles are not fixed in advance (Bolić et al., 2024; Bachoc et al., 2025a;b). A large but more distantly related literature studies bilateral trade in the Bayesian setting: following the impossibility result of Myerson & Satterthwaite (1983), it asks how well simple mechanisms, fixed prices in particular, approximate the optimal GFT (Blumrosen & Mizrahi, 2016; Colini-Baldeschi et al., 2017; Brustle et al., 2017; Dütting et al., 2021; Kang et al., 2022; Deng et al., 2022).

2 Preliminaries

We study online learning in *repeated bilateral trade* over a horizon of T rounds, in which a learner faces a sequence of seller–buyer pairs seeking to trade an indivisible item. In each round $t \in [T]$,

a new seller and a new buyer arrive with private valuations $s_t \in [0,1]$ and $b_t \in [0,1]$ for the item, respectively. We consider the *stochastic setting* with independent valuations, where the valuation pairs are i.i.d. samples from a product distribution $\mu_S \otimes \mu_B$. We further assume *bounded density*: μ_S and μ_B have densities bounded above by a constant $M \geq 1$.

Notation. For a positive integer k, let $[k] \triangleq \{1, \ldots, k\}$. For a set A, let $\Delta(A)$ denote the set of probability distributions on A, and $\mathrm{Unif}(A)$ the uniform distribution on A. The $\mathcal{O}(\cdot)$ notation hides the dependence on M, and $\widetilde{\mathcal{O}}(\cdot)$ further hides polylogarithmic factors in T.

Learning protocol. At each round $t \in [T]$, the learner posts two prices: $p_t \in [0,1]$ to the seller and $q_t \in [0,1]$ to the buyer. Then, the seller accepts if and only if $s_t \leq p_t$, the buyer accepts if and only if $b_t \geq q_t$, and trade occurs if and only if both accept. The learner observes only the trade outcome $Z_t \triangleq \mathbb{1}\{s_t \leq p_t,\ b_t \geq q_t\}$, known as *one-bit feedback*. The learner knows M and T but not μ_S or μ_B, and chooses the price pair (p_t, q_t) based on the history $\mathcal{H}_{t-1} \triangleq (p_r, q_r, Z_r)_{r\in[t-1]}$ of past prices and feedback.

We consider the most fundamental economic objective for the learner: maximizing the *gains-from-trade* (GFT), which is the difference between the buyer's and the seller's valuations if trade occurs, and 0 otherwise, i.e., $(b_t - s_t) \cdot Z_t$. For a price pair $(p,q) \in [0,1]^2$, we define the expected GFT with respect to the distribution of valuations as

$$\mathsf{GFT}(p,q) \triangleq \mathbb{E}_{(s,b)\sim\mu_S\otimes\mu_B}\left[(b-s)\cdot\mathbb{1}\{s\leq p,\ b\geq q\}\right].$$

The GFT represents the increase in social welfare generated by the trade.

Budget balance. A key requirement in bilateral trade is budget balance, which guarantees that the learner does *not* subsidize the market. When trade occurs, the seller receives p_t from the learner and the buyer pays q_t to the learner, so the learner obtains a *profit* of $q_t - p_t$, which can be negative. Formally, the profit in round t is $(q_t - p_t)\cdot Z_t$, and the cumulative profit after round t is $\Pi_t \triangleq \sum_{r\in[t]}(q_r - p_r)\cdot Z_r$, with $\Pi_0 = 0$. We consider two budget balance constraints:

- *weak budget balance* (WBB): for every round $t \in [T]$, $p_t \leq q_t$;
- *global budget balance* (GBB): the cumulative profit over T rounds $\Pi_T \geq 0$.

WBB enforces that the learner does not lose money in any round, while GBB only enforces this in total, so $p_t > q_t$ is allowed in some rounds as long as the loss is recovered later. For a price pair $(p,q) \in [0,1]^2$, we also define the expected profit with respect to the distribution of valuations as

$$\mathsf{PRO}(p,q) \triangleq (q-p)\cdot\Pr_{(s,b)\sim\mu_S\otimes\mu_B}[s\leq p,\ b\geq q].$$

Benchmarks and regret. We aim to design learning algorithms that post prices to maximize the GFT under the WBB or GBB constraint. We measure the performance of an algorithm by its regret against the best fixed strategy that satisfies the same budget balance constraint. For WBB, the benchmark is the best fixed price pair (p,q) with $p \leq q$, and it suffices to consider common prices: for $p \leq z \leq q$, every trade at (p,q) also occurs at (z,z) and generates the same GFT. Hence, the value of the best fixed price benchmark is $\mathsf{GFT}^\star_{\mathrm{W}} \triangleq \sup_{z\in[0,1]}\mathsf{GFT}(z,z)$, and the regret of an algorithm satisfying WBB is

$$\mathsf{REG}_{\mathrm{W}} \triangleq T\cdot\mathsf{GFT}^\star_{\mathrm{W}} - \mathbb{E}\left[\sum_{t\in[T]}\mathsf{GFT}(p_t,q_t)\right],$$

where the expectation is over the valuations and the randomness of the algorithm.

For GBB, following Lunghi et al. (2026b), the benchmark is the best distribution over price pairs with nonnegative expected profit, whose value is

$$\mathsf{GFT}^\star_{\mathrm{G}} \triangleq \sup_{\gamma\in\Delta([0,1]^2)}\mathbb{E}_{(p,q)\sim\gamma}[\mathsf{GFT}(p,q)] \quad \text{s.t.} \quad \mathbb{E}_{(p,q)\sim\gamma}[\mathsf{PRO}(p,q)] \geq 0.$$

Every common price has zero profit, so the best feasible distribution benchmark is stronger than the best fixed price benchmark, and we have $\mathsf{GFT}^\star_{\mathrm{G}} \geq \mathsf{GFT}^\star_{\mathrm{W}}$. The regret of an algorithm satisfying GBB is defined in the same way against this benchmark: $\mathsf{REG}_{\mathrm{G}} \triangleq T\cdot\mathsf{GFT}^\star_{\mathrm{G}} - \mathbb{E}\left[\sum_{t\in[T]}\mathsf{GFT}(p_t,q_t)\right]$.

GFT decomposition. We conclude this section with a decomposition of the expected GFT, which is useful in the design and analysis of our algorithms for WBB and GBB. For a valuation pair $(s,b) \sim \mu_S \otimes \mu_B$ and a price pair (p,q), we define the *conditional utilities* of the seller and the buyer

$$U_S(p) \triangleq \mathbb{E}_{s\sim\mu_S}[p - s \mid s \leq p] \in [0,p], \qquad U_B(q) \triangleq \mathbb{E}_{b\sim\mu_B}[b - q \mid b \geq q] \in [0, 1-q],$$

Algorithm 1 TREE-SEARCH

1: $\mathcal{A}_0 \leftarrow \{(0,1)\}, \quad \mathcal{A}_1, \ldots, \mathcal{A}_J \leftarrow \emptyset, \quad \widehat{U}_S(0) \leftarrow 0, \quad \widehat{U}_B(1) \leftarrow 0$
/* exploration phase */
2: **for** $j = 0, 1, \ldots, J$ **do**
3: **for all** $(p,q) \in \mathcal{A}_j$ **do**
4: $\widehat{\mathsf{GFT}}\left(\frac{p+q}{2}, \frac{p+q}{2}\right), \widehat{U}_S\left(\frac{p+q}{2}\right), \widehat{U}_B\left(\frac{p+q}{2}\right) \leftarrow$ MIDPOINT-EST $\left(p, q, \widehat{U}_S(p), \widehat{U}_B(q)\right)$
5: **if** $j < J$ **then**
6: $\widehat{\mathsf{GFT}}_{\max} \leftarrow \max\left\{\widehat{\mathsf{GFT}}\left(\frac{p+q}{2}, \frac{p+q}{2}\right) : (p,q) \in \mathcal{A}_j\right\}$
7: **for all** $(p,q) \in \mathcal{A}_j$ **do**
8: **if** $\widehat{\mathsf{GFT}}\left(\frac{p+q}{2}, \frac{p+q}{2}\right) + 2\varepsilon + M2^{-j-1} \geq \widehat{\mathsf{GFT}}_{\max}$ **then**
9: $\mathcal{A}_{j+1} \leftarrow \mathcal{A}_{j+1} \cup \left\{\left(p, \frac{p+q}{2}\right), \left(\frac{p+q}{2}, q\right)\right\}$
/* exploitation phase */
10: $\widehat{z} \leftarrow$ the midpoint of a pair of $\mathcal{A}_J$ maximizing $\widehat{\mathsf{GFT}}$, with fixed tie-breaking
11: post $(\widehat{z}, \widehat{z})$ in every remaining round

set to zero when the conditioning event has probability zero; in particular, $U_S(0) = U_B(1) = 0$. Whenever a trade occurs, the GFT is $b - s = (p - s) + (q - p) + (b - q)$, so by the independence of $s \sim \mu_S$ and $b \sim \mu_B$, we can decompose the expected GFT into the conditional utilities, multiplied by the trade probability, and the expected profit:

$$\mathsf{GFT}(p,q) = \Pr_{(s,b)\sim\mu_S\otimes\mu_B}[s \leq p,\ b \geq q] \cdot (U_S(p) + U_B(q)) + \mathsf{PRO}(p,q). \tag{1}$$

3 WEAK BUDGET BALANCE AGAINST THE FIXED-PRICE BENCHMARK

In this section, we study GFT maximization under WBB, with the best fixed price as the benchmark. We design an online learning algorithm (TREE-SEARCH) that satisfies WBB and achieves $\widetilde{\mathcal{O}}(T^{2/3})$ regret. This improves the $\widetilde{\mathcal{O}}(T^{3/4})$ bound of Cesa-Bianchi et al. (2024b) and matches their $\Omega\left(T^{2/3}\right)$ lower bound up to polylogarithmic factors. Before presenting our algorithm, we introduce the following recursive decomposition as a tool to estimate GFT at a common price.

Lemma 1. *For a valuation pair $(s,b) \sim \mu_S \otimes \mu_B$ and prices p, z, q with $0 \leq p < z < q \leq 1$, let $X \sim \mathrm{Unif}[p,z]$ and $Y \sim \mathrm{Unif}[z,q]$ be independent of (s,b). Then we have*

$$\Pr[s \leq z,\ b \geq z]\, U_S(z) = \Pr[s \leq p,\ b \geq z]\, U_S(p) + (z-p)\Pr[s \leq X,\ b \geq z],$$
$$\Pr[s \leq z,\ b \geq z]\, U_B(z) = \Pr[s \leq z,\ b \geq q]\, U_B(q) + (q-z)\Pr[s \leq z,\ b \geq Y].$$

At a common price z the profit is zero, so eq. (1) becomes $\mathsf{GFT}(z,z) = \Pr[s \leq z,\ b \geq z] \cdot (U_S(z) + U_B(z))$, and estimating GFT at z reduces to estimating the trade probability and the two conditional utilities. The trade probability can be estimated directly from empirical trade frequencies, whereas the utilities cannot be estimated directly from one-bit feedback; Lemma 1 instead lets us estimate the utilities recursively. For $z \in (p,q)$, Lemma 1 expresses the utilities $U_S(z)$ and $U_B(z)$ in terms of the seller's utility $U_S(p)$ at the left endpoint p, the buyer's utility $U_B(q)$ at the right endpoint q, and five trade probabilities, each estimable from trade bits. Since $U_S(0) = U_B(1) = 0$ are known, applying Lemma 1 with $p = 0$ and $q = 1$ would already estimate the utilities at any z, but this would incur large regret, because the auxiliary prices $X \sim \mathrm{Unif}[0,z]$ and $Y \sim \mathrm{Unif}[z,1]$ spread over all of $[0,1]$ and thus miss many trades that occur at (z,z).

Algorithm overview. We instead apply Lemma 1 on short intervals $[p,q]$ whose endpoint utilities have already been estimated on earlier, larger intervals, so the prices posted to estimate GFT at $z \in (p,q)$ all lie in $[p,q]$ and incur little regret when z is near-optimal. Algorithm 1 implements this along a dyadic tree of price intervals, whose depth J is set below. At each depth j, the algorithm evaluates a set $\mathcal{A}_j$ of active intervals: $\mathcal{A}_0 \triangleq \{(0,1)\}$ (line 1), and each later active set is formed at the previous depth. For each active interval $(p,q) \in \mathcal{A}_j$, the algorithm computes the estimated midpoint GFT $\widehat{\mathsf{GFT}}\left(\frac{p+q}{2}, \frac{p+q}{2}\right)$ and the estimated midpoint utilities $\widehat{U}_S\left(\frac{p+q}{2}\right)$ and $\widehat{U}_B\left(\frac{p+q}{2}\right)$ by calling MIDPOINT-EST (Algorithm 4); see line 4. It then takes the largest estimated midpoint GFT $\widehat{\mathsf{GFT}}_{\max}$ (line 6), selects only the intervals $(p,q) \in \mathcal{A}_j$ whose estimated midpoint GFT is close to $\widehat{\mathsf{GFT}}_{\max}$, up to a slack for estimation error and interval width, and adds their two halves $\left(p, \frac{p+q}{2}\right)$ and

$\left(\frac{p+q}{2}, q\right)$ to $\mathcal{A}_{j+1}$ (lines 8 and 9). This focuses exploration on prices that may be optimal: an interval whose estimated midpoint GFT is not close to $\widehat{\mathsf{GFT}}_{\max}$ is pruned, and none of its subintervals is explored at later depths. At depth $j+1$, the midpoint utility estimates obtained at depth j serve as the endpoint estimates of the intervals $\left(p, \frac{p+q}{2}\right)$ and $\left(\frac{p+q}{2}, q\right)$. Finally, the algorithm posts the midpoint of an interval in $\mathcal{A}_J$ with the largest estimated GFT in every remaining round (lines 10 and 11).

For an active interval (p, q), MIDPOINT-EST runs five consecutive blocks of n rounds, each estimating one trade probability in Lemma 1 with $z = (p+q)/2$ by posting pairs with $p_t \le q_t$, e.g., (z, z), (p, z) and (z, q), so TREE-SEARCH satisfies WBB. We set the discretization parameter $K \triangleq T^{1/3}/(\log T)^{4/3}$, the tree depth $J \triangleq \lfloor \log_2 K \rfloor$, and the target accuracy $\varepsilon \triangleq M/K$ for TREE-SEARCH, and set the block length $n \triangleq \lceil \log(20KT)/(2\eta^2) \rceil$ and the concentration tolerance $\eta \triangleq \varepsilon/[16(J+1)]$ for MIDPOINT-EST. Since each call to MIDPOINT-EST uses $5n$ rounds and $|\mathcal{A}_j| \le 2^j$, exploration uses at most $5n \sum_{j=0}^{J} 2^j < 10nK$ rounds, and for all sufficiently large T these choices satisfy $10nK \le T$, so exploration ends within the horizon. Next, we present the guarantee of TREE-SEARCH and outline the analysis of its regret.

Theorem 1. *For GFT maximization in repeated bilateral trade with one-bit feedback, where the valuations are i.i.d. draws from a product distribution with bounded density,* TREE-SEARCH *(Algorithm 1) satisfies WBB and achieves regret $\widetilde{\mathcal{O}}(T^{2/3})$ against the best fixed price.*

Proof sketch. The analysis starts from a standard consequence of bounded densities: the diagonal GFT $z \mapsto \mathsf{GFT}(z, z)$ is M-Lipschitz (Lemma 3; see also Cesa-Bianchi et al., 2024b). This property only needs the density assumption, and it is what allows the algorithm to stop refining at a finite resolution. As is common in bandit analyses, we then define a good event $\mathcal{E}$ on which the empirical trade frequency in every block collected by MIDPOINT-EST is within η of its expectation. It holds with probability at least $1 - 1/T$, so its complement costs at most one unit of regret, and the rest of the analysis takes place on $\mathcal{E}$. On this event, Lemma 4 shows that every midpoint GFT estimate is within $\varepsilon/2$ of its true value. This does not follow from the block accuracy alone, since each utility estimate is computed from an endpoint estimate and inherits its error; the lemma tracks these errors after weighting them by the trade probability, so that each level of the tree adds only $\mathcal{O}(\eta)$.

Combining the Lipschitz property with this accuracy, Lemma 5 controls every interval that the search evaluates. The slack in the retention rule (Algorithm 1, line 8) ensures that the interval containing the optimal price $z^\star$ is never pruned, and every midpoint evaluated at depth j is within a constant multiple of $M2^{-j}$ of $\mathsf{GFT}^\star_{\mathrm{W}}$ (since $\varepsilon \le M2^{-j}$). Pruning therefore does more than preserve the optimum: the intervals that survive are all near-optimal, increasingly so as the tree deepens, and the algorithm can afford to evaluate every one of them.

The regret bound follows from these facts. On $\mathcal{E}$, the regret splits into an exploration part, incurred while the tree is built, and an exploitation part, incurred while $\widehat{z}$ is posted. An exploration round at depth j incurs two gaps: from $\mathsf{GFT}^\star_{\mathrm{W}}$ to the midpoint GFT of its interval, which is $\mathcal{O}(M2^{-j})$ as shown above, and from this midpoint to the pair that MIDPOINT-EST posts in its place, whose prices lie in the same interval, so this gap is another $\mathcal{O}(M2^{-j})$. Since depth j holds at most 2^j intervals, each evaluated for $5n$ rounds, every depth costs $\mathcal{O}(nM)$: the number of intervals at most doubles as their width halves. Over the $J+1$ depths, exploration therefore costs $\widetilde{\mathcal{O}}(K^2)$, since accuracy $\varepsilon = M/K$ requires $n = \widetilde{\mathcal{O}}(K^2)$ rounds per block. During exploitation, stopping at resolution 2^{-J} and selecting $\widehat{z}$ from estimates each lose $\mathcal{O}(M/K)$ per round, or $\mathcal{O}(MT/K)$ in total. Choosing $K = \widetilde{\Theta}(T^{1/3})$ balances the two parts and yields the $\widetilde{\mathcal{O}}(T^{2/3})$ bound of Theorem 1. The full proof is in Appendix B. □

4 GLOBAL BUDGET BALANCE AGAINST THE DISTRIBUTION BENCHMARK

In this section, we study GFT maximization under GBB, with the best feasible distribution over price pairs as the benchmark. We design an online learning algorithm (Algorithm 2) that satisfies GBB and achieves $\widetilde{\mathcal{O}}(T^{2/3})$ regret. This improves the $\widetilde{\mathcal{O}}(T^{3/4})$ bound of Lunghi et al. (2026b) and matches the $\Omega\left(T^{2/3}\right)$ lower bound of Chen et al. (2026) up to polylogarithmic factors. The benchmark may mix price pairs that break even only in combination, so learning their GFT alone is insufficient: the algorithm must also learn their profit. The algorithm first accumulates a profit reserve, then uses it to finance exploration and subsidized trades. During the second stage, it maintains candidate valuation distributions consistent with the feedback, selects an optimistic feasible pricing distribution, and probes nearby prices to refine the candidates.

Algorithm 2 TWO-STAGE-GBB

Input: T, M.
$\tau \leftarrow$ PROFIT-MAX $\left(T^{1/3}, \beta, T\right)$ // the prefix rounds
if $K = 0$ **then**
  post $(0, 1)$ in every remaining round
run GLOBAL-OFU$(K, \mathcal{F}, \tau)$ // the suffix rounds

Posting a pair (p, q) repeatedly reveals its trade probability and hence its expected profit, but does not identify the utilities needed to compute its GFT. To recover these utilities, we extend the recursion of Lemma 1 to an arbitrary fixed opposite-side price.

Lemma 2. *For a valuation pair $(s, b) \sim \mu_S \otimes \mu_B$ and an interval $I = [u, v] \subseteq [0, 1]$ with $u < v$ and width $|I| = v - u$, let $X \sim \mathrm{Unif}(I)$ be independent of (s, b). Then, for any fixed buyer price $q \in [0, 1]$ and any fixed seller price $p \in [0, 1]$,*

$$\Pr[s \le v,\ b \ge q] \cdot U_S(v) = \Pr[s \le u,\ b \ge q] \cdot U_S(u) + |I| \cdot \Pr[s \le X,\ b \ge q],$$
$$\Pr[s \le p,\ b \ge u] \cdot U_B(u) = \Pr[s \le p,\ b \ge v] \cdot U_B(v) + |I| \cdot \Pr[s \le p,\ b \ge X].$$

The three probabilities in each recursion can be estimated by posting the two endpoint prices and a uniform price in the interval, while keeping the opposite-side price fixed. By eq. (1), the two recursions, together with the trade probability, determine the GFT and the profit of (p, q).

Applying the seller recursion directly on $[0, p]$ would recover the weighted utility at p from $U_S(0) = 0$. The difficulty is the cost of obtaining this information: even if (p, q) belongs to a good pricing distribution, a uniform probe over $[0, p]$ can be far from p and lose a constant amount of GFT. Instead, we use dyadic paths whose probes change one price of (p, q) by less than $2|I|$ at a node I. Lipschitz continuity then bounds the resulting loss by $\mathcal{O}(|I|)$ per round. Independence also lets information be shared: the fixed side's acceptance probability cancels in ratios of the three trade probabilities. Thus same-side observations from the same interval can be pooled even when the opposite price changes.

To fund these probes, the algorithm follows the two-stage construction of Bernasconi et al. (2024). It first runs PROFIT-MAX with discretization parameter $T^{1/3}$ and reserve target $\beta \triangleq 1 + 500(1 + M)T^{2/3}[\log(17T^2)]^3$. Let τ be the round at which this stage stops; if $\tau < T$, it has accumulated at least β. This reserve pays for exploration and any losses during the second stage.[1] The remaining horizon determines the dyadic resolution

$$K \triangleq \max\left\{2^j : j \ge 3,\ 10^{10} 2^{3j} \left[\log(17T^2)\right]^4 \le \tfrac{T-\tau}{2}\right\},$$

with $\max \varnothing \triangleq 0$. If $K = 0$, the algorithm posts $(0, 1)$ in every remaining round and terminates. Otherwise, it runs GLOBAL-OFU on the grid $\mathcal{G}_K \triangleq \left\{\left(\frac{i}{K}, \frac{j}{K}\right) : i, j = 0, \ldots, K\right\}$, using a dyadic tree of depth $J \triangleq \log_2 K$.

The second stage learns the unknown independent valuation law through the candidate family

$$\mathcal{F} \triangleq \left\{\theta = \mu_{S,\theta} \otimes \mu_{B,\theta} : \left\|\frac{d\mu_{S,\theta}}{dx}\right\|_\infty \le M,\ \left\|\frac{d\mu_{B,\theta}}{dx}\right\|_\infty \le M\right\}.$$

A subscript θ means that a probability, utility, GFT or profit is evaluated under that candidate law. The previous mechanism of Lunghi et al. (2026b) reduces GFT exploration to $2K$ groups but retains a K^2-armed profit bandit, whose $\widetilde{\mathcal{O}}(K\sqrt{T})$ cost leads to the $T^{3/4}$ rate. We use the shared information in $\mathcal{F}$ to learn GFT and profit together.

GLOBAL-OFU **subroutine.** GLOBAL-OFU begins with boundary exploration. At $(p, 0)$, the trade bit reports whether the seller accepts p; at $(1, q)$, it reports whether the buyer accepts q. With tolerance $\eta \triangleq 1/[80K(J + 1)]$, the algorithm samples each of these $2K + 1$ boundary grid pairs for $n \triangleq \lceil 128 \log(17T^2)/\eta \rceil$ rounds. The resulting marginal estimates define an initial candidate set $\mathcal{F}_0$ through the KL constraints in Algorithm 3, where $\mathrm{KL}(x, y) \triangleq x \log \frac{x}{y} + (1 - x) \log \frac{1-x}{1-y}$.

[1] Unlike the initial budget in bandits with knapsacks, this reserve must be earned through trade. With a sublinear budget-violation allowance, Lunghi et al. (2026a) instead use that allowance to fund subsidies without first earning a reserve.

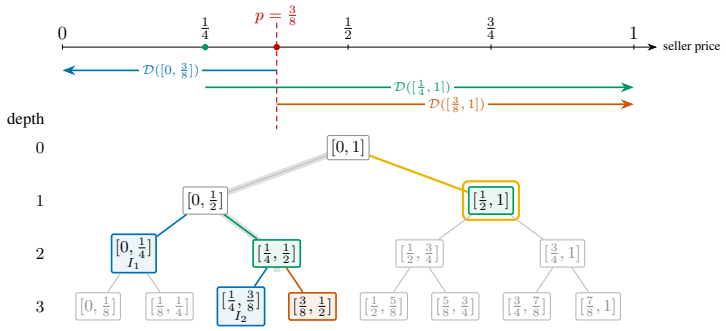

Figure 1: PATH at $p = \frac{3}{8}$ with $J = 3$: blue marks $\mathcal{D}([0, \frac{3}{8}])$, green $\mathcal{D}([\frac{1}{4}, 1])$ and orange $\mathcal{D}([\frac{3}{8}, 1])$; the yellow interval is contained in both $\mathcal{D}([\frac{1}{4}, 1])$ and $\mathcal{D}([\frac{3}{8}, 1])$, and the pale trunk the intervals that straddle $\frac{3}{8}$.

For a pair (p, q), write $\mathcal{S}(p, q)$ and $\mathcal{B}(p, q)$ for its seller and buyer *paths*, each a collection of dyadic intervals. To define these paths, consider the binary tree obtained by repeatedly bisecting $[0, 1]$ down to depth J. Its nodes are the dyadic intervals

$$\left[i2^{-j}, (i+1)2^{-j}\right], \qquad 0 \le j \le J, \quad 0 \le i < 2^j.$$

For a grid-aligned interval $[a, b]$ with $a = 0$ or $b = 1$, let $\mathcal{D}([a, b])$ contain the largest tree intervals lying entirely inside $[a, b]$: an interval is included only if no larger tree interval inside $[a, b]$ contains it. These intervals have disjoint interiors and together cover $[a, b]$. For example, $\mathcal{D}([0, \frac{3}{8}]) = \left\{[0, \frac{1}{4}], [\frac{1}{4}, \frac{3}{8}]\right\}, \mathcal{D}([\frac{1}{4}, 1]) = \left\{[\frac{1}{4}, \frac{1}{2}], [\frac{1}{2}, 1]\right\}$. In the first example, $[0, \frac{1}{2}]$ extends beyond $\frac{3}{8}$, so we use $[0, \frac{1}{4}]$ and then $[\frac{1}{4}, \frac{3}{8}]$ to cover the remainder. Figure 1 illustrates the construction at $p = 3/8$. The full construction and the role of each partition are given in Appendix D.

The seller path starts with $\mathcal{D}([0, p])$ and adds $\mathcal{D}([v, 1])$ at the right endpoint v of each interval $[u, v] \in \mathcal{D}([0, p])$. The collection $\mathcal{S}(p, q)$ is the union of the intervals in all these partitions. The buyer path mirrors this construction: $\mathcal{B}(p, q)$ contains the intervals in $\mathcal{D}([q, 1])$ and in $\mathcal{D}([0, u])$ for every $[u, v] \in \mathcal{D}([q, 1])$. The intervals in $\mathcal{D}([0, p])$ form a chain along which Lemma 2 propagates the known utility $U_S(0) = 0$ to the target price p. To evaluate the probability weights in this recursion, we also need the acceptance probabilities at the interval endpoints. The shared observations estimate ratios such as $\Pr[s \le u] / \Pr[s \le v]$, which alone do not determine these probabilities. For each right endpoint v, the additional partition $\mathcal{D}([v, 1])$ connects it to the boundary 1, where the seller accepts with probability $\Pr[s \le 1] = 1$. Multiplying the successive ratios along this partition recovers $\Pr[s \le v]$. The buyer path similarly recovers buyer acceptance probabilities from the known value $\Pr[b \ge 0] = 1$. By independence, multiplying the seller and buyer acceptance probabilities gives the joint trade probabilities needed in the recursion and the GFT calculation.

For an interval $I = [u, v]$ and an independent $X \sim \mathrm{Unif}(I)$, the three local probes are the following price pairs, where arm 0 is the *anchor*, and $a_{I,\ell}$ denotes the event that price pair ℓ makes a successful trade:

	$\ell = 0$	$\ell = 1$	$\ell = 2$
$I \in \mathcal{S}(p, q)$	(v, q)	(u, q)	(X, q)
$I \in \mathcal{B}(p, q)$	(p, u)	(p, v)	(p, X)

At epoch k, following the principle of optimism in the face of uncertainty (OFU), the subroutine selects a candidate $\theta_k \in \mathcal{F}_{k-1}$ and a distribution γ_k on $\mathcal{G}_K$ that maximize GFT while breaking even under θ_k. In Appendix E, we show that the OFU problem admits a maximizer by reducing it to a finite-dimensional problem over a compact set, and give a polynomial-time algorithm that solves it approximately. Moreover, γ_k can be chosen with at most two support points. For any candidate-indexed quantity X_θ, write its upper and lower envelopes as

$$\overline{X}_k \triangleq \sup_{\theta \in \mathcal{F}_{k-1}} X_\theta, \qquad \underline{X}_k \triangleq \inf_{\theta \in \mathcal{F}_{k-1}} X_\theta.$$

To decide where to explore, the algorithm measures disagreement about the local probability ratios,

$$\phi_k\left((p, q), I\right) \triangleq \overline{\Pr}_k[a_{I,0}] \cdot \max_{\ell \in \{1,2\}} \sup_{\theta, \theta' \in \mathcal{F}_{k-1}} \left(\frac{\Pr_\theta[a_{I,\ell}]}{\Pr_\theta[a_{I,0}]} - \frac{\Pr_{\theta'}[a_{I,\ell}]}{\Pr_{\theta'}[a_{I,0}]} \right).$$

Algorithm 3 GLOBAL-OFU

Input: $K, \mathcal{F}, \tau$.
$t \leftarrow \tau + 1$; $\mathcal{F}_0 \leftarrow \mathcal{F}$; $N_{I,\ell} \leftarrow 0$ for every typed I and every ℓ
$\eta \leftarrow 1/[80K(J+1)]$
for all $(p,q) \in \mathcal{G}_K$ with $q = 0$ or $p = 1$ **do** /* boundary exploration phase */
  $\widehat{\Pr}[s \le p,\ b \ge q] \leftarrow 0$
  for $i = 1, \dots, n$ **do**
    if $\Pi_{t-1} < 1$ or $t > \frac{T+\tau}{2}$ **then**
      post $(0,1)$ in every remaining round
    post (p,q); $\widehat{\Pr}[s \le p,\ b \ge q] \mathrel{+}= \frac{1}{n} Z_t\,(p,q)$
  $\mathcal{F}_0 \leftarrow \left\{\theta \in \mathcal{F}_0 : \mathrm{KL}\left(\widehat{\Pr}[s \le p,\ b \ge q]\,, \Pr_\theta[s \le p,\ b \ge q]\right) \le \frac{\eta}{128}\right\}$
for $k = 1, 2, \dots$ **do** /* refinement phase */
  if $\mathcal{F}_{k-1} = \emptyset$ or $t > \frac{T+\tau}{2}$ **then**
    post $(0,1)$ in every remaining round
  Solve the following problem:

$$(\theta_k, \gamma_k) = \underset{\substack{\theta \in \mathcal{F}_{k-1},\ \gamma \in \Delta(\mathcal{G}_K) \\ |\mathrm{supp}(\gamma)| \le 2}}{\arg\max} \mathbb{E}_{(p,q)\sim\gamma}\left[\mathsf{GFT}_\theta\,(p,q)\right]$$

$$\text{s.t. } \mathbb{E}_{(p,q)\sim\gamma}\left[(q-p) \cdot \Pr_\theta[s \le p,\ b \ge q]\right] \ge 0$$

  for all $(p,q) \in \mathrm{supp}(\gamma_k)$ **do**
    $(\mathcal{S}(p,q), \mathcal{B}(p,q)) \leftarrow$ PATH (p,q)
  $((p,q), I) \leftarrow \underset{\substack{(p,q)\in\mathrm{supp}(\gamma_k) \\ I \in \mathcal{S}(p,q)\cup\mathcal{B}(p,q)}}{\arg\max} \gamma_k(p,q) \cdot \phi_k\left((p,q), I\right)$
  if $\gamma_k(p,q) \cdot \phi_k\left((p,q), I\right) \le \eta$ **then**
    break
  $\mathcal{F}_k \leftarrow$ LOCAL-REFINE $(\gamma_k, (p,q), I, \mathcal{F}_{k-1})$
    // Probe locally when the sampled pair is (p,q).
    // Update shared trade counts and filter candidates.
while $t \le T$ and $\Pi_{t-1} \ge 1$ **do** /* exploitation phase */
  draw $(p_t, q_t) \sim \gamma_k$ and post it
post $(0,1)$ in every remaining round

The algorithm chooses the pair and path node maximizing $\gamma_k(p,q) \cdot \phi_k((p,q), I)$. Multiplying by $\gamma_k(p,q)$ accounts for the weight of the pair in the selected distribution.

Given γ_k, the selected pair (p,q), its node I and the current candidate set, LOCAL-REFINE runs a block of refinement rounds. Each round first draws a price pair from γ_k; if it draws (p,q), it replaces that pair by a uniformly chosen pair of I, and otherwise posts the drawn pair. The subroutine updates the shared counts $N_{I,\ell}$ of trades generated by each arm and retains candidates consistent with the relative proportions of these counts. It returns the updated family $\mathcal{F}_k$ for the next epoch. The block length adapts to the accumulated count $N_I \triangleq \sum_\ell N_{I,\ell}$ and the estimated chance of obtaining a trade. The full subroutine is given in Algorithm 5 in Appendix C.2. The complete algorithm and its second stage are given in Algorithms 2 and 3.

When the maximum weighted uncertainty over the current support and its paths is at most the predetermined tolerance η, the algorithm stops refining and uses γ_k for exploitation. The analysis bounds its GFT shortfall and its expected profit deficit by $\mathcal{O}(1/K)$ per round. GBB is enforced separately: the second stage switches permanently to $(0,1)$ if the cumulative profit is below one, and exploration also stops if its allotted time is exhausted or the candidate family becomes empty. These rules keep the cumulative profit nonnegative on every sample path, while the reserve ensures that the GBB rule does not interrupt learning or exploitation on the high-probability event of the analysis.

Finally, we establish the regret guarantee for Algorithm 2. The full proof is deferred to Appendix C.

Theorem 2. *For GFT maximization in repeated bilateral trade with one-bit feedback, where the valuations are i.i.d. draws from a product distribution with bounded density, Algorithm 2 satisfies GBB and achieves regret $\widetilde{\mathcal{O}}(T^{2/3})$ against the best feasible distribution over price pairs.*

Due to the page limit, we defer the conclusion and directions for future work to Appendix F.

AI USE STATEMENT

We used generative AI tools as an interactive aid in developing and checking parts of the proofs, as well as in improving the presentation of the paper. All mathematical arguments and claims were independently reviewed and verified by the authors, who take full responsibility for the content of the paper.

REFERENCES

Yossi Azar, Amos Fiat, and Federico Fusco. An α-regret analysis of adversarial bilateral trade. *Artificial Intelligence*, 337:104231, 2024. doi: 10.1016/j.artint.2024.104231. URL https://doi.org/10.1016/j.artint.2024.104231.

Moshe Babaioff, Amitai Frey, and Noam Nisan. Learning to maximize gains from trade in small markets. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pp. 195. ACM, 2024. doi: 10.1145/3670865.3673463. URL https://doi.org/10.1145/3670865.3673463.

Francesco Bacchiocchi, Matteo Castiglioni, Roberto Colomboni, and Alberto Marchesi. Online bilateral trade with minimal feedback: Don't waste seller's time. In *Advances in Neural Information Processing Systems 38*, 2025. URL https://openreview.net/forum?id=tH8Q2TRSZf.

François Bachoc, Nicolò Cesa-Bianchi, Tommaso Cesari, and Roberto Colomboni. Fair online bilateral trade. In *Advances in Neural Information Processing Systems 37*, pp. 37241–37263, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/419ce7cb77aa90868c16882ce4257a69-Paper-Conference.pdf.

François Bachoc, Tommaso Cesari, and Roberto Colomboni. A parametric contextual online learning theory of brokerage. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 2293–2309. PMLR, 2025a. URL https://proceedings.mlr.press/v267/bachoc25a.html.

François Bachoc, Tommaso Cesari, and Roberto Colomboni. A tight regret analysis of non-parametric repeated contextual brokerage. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 2836–2844. PMLR, 2025b. URL https://proceedings.mlr.press/v258/bachoc25a.html.

Martino Bernasconi, Matteo Castiglioni, Andrea Celli, and Federico Fusco. No-regret learning in bilateral trade via global budget balance. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pp. 247–258. ACM, 2024. doi: 10.1145/3618260.3649653. URL https://doi.org/10.1145/3618260.3649653.

Avrim Blum, Tuomas Sandholm, and Martin Zinkevich. Online algorithms for market clearing. *Journal of the ACM*, 53(5):845–879, 2006. doi: 10.1145/1183907.1183913. URL https://doi.org/10.1145/1183907.1183913.

Liad Blumrosen and Yehonatan Mizrahi. Approximating gains-from-trade in bilateral trading. In *Web and Internet Economics – 12th International Conference, WINE 2016*, pp. 400–413. Springer, 2016. doi: 10.1007/978-3-662-54110-4_28. URL https://doi.org/10.1007/978-3-662-54110-4_28.

Natasa Bolić, Tommaso Cesari, and Roberto Colomboni. An online learning theory of brokerage. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, pp. 216–224, 2024.

Johannes Brustle, Yang Cai, Fa Wu, and Mingfei Zhao. Approximating gains from trade in two-sided markets via simple mechanisms. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pp. 589–590. ACM, 2017. doi: 10.1145/3033274.3085148. URL https://doi.org/10.1145/3033274.3085148.

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

Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. Bilateral trade: A regret minimization perspective. *Mathematics of Operations Research*, 49(1): 171–203, 2024b. doi: 10.1287/moor.2023.1351. URL https://doi.org/10.1287/moor.2023.1351.

Houshuang Chen, Yaonan Jin, Pinyan Lu, and Chihao Zhang. Tight regret bounds for fixed-price bilateral trade. In *53rd International Colloquium on Automata, Languages, and Programming (ICALP 2026)*, volume 374 of *Leibniz International Proceedings in Informatics*, pp. 57:1–57:21. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026. doi: 10.4230/LIPIcs.ICALP.2026.57. URL https://doi.org/10.4230/LIPIcs.ICALP.2026.57.

Riccardo Colini-Baldeschi, Paul W. Goldberg, Bart de Keijzer, Stefano Leonardi, and Stefano Turchetta. Fixed price approximability of the optimal gain from trade. In *Web and Internet Economics – 13th International Conference, WINE 2017*, pp. 146–160. Springer, 2017. doi: 10.1007/978-3-319-71924-5_11. URL https://doi.org/10.1007/978-3-319-71924-5_11.

Yuan Deng, Jieming Mao, Balasubramanian Sivan, and Kangning Wang. Approximately efficient bilateral trade. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 718–721. ACM, 2022. doi: 10.1145/3519935.3520054. URL https://doi.org/10.1145/3519935.3520054.

Simone Di Gregorio, Paul Dütting, Federico Fusco, and Chris Schwiegelshohn. Nearly tight regret bounds for profit maximization in bilateral trade. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science*, pp. 1570–1594. IEEE, 2025. doi: 10.1109/FOCS63196.2025.00083. URL https://doi.org/10.1109/FOCS63196.2025.00083.

Paul Dütting, Federico Fusco, Philip Lazos, Stefano Leonardi, and Rebecca Reiffenhäuser. Efficient two-sided markets with limited information. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pp. 1452–1465. ACM, 2021. doi: 10.1145/3406325.3451076. URL https://doi.org/10.1145/3406325.3451076.

Yiding Feng, Mengfan Ma, Bo Peng, and Zongqi Wan. Searching for optimal prices in two-sided markets. *arXiv preprint arXiv:2602.11691*, 2026. URL https://arxiv.org/abs/2602.11691.

Solenne Gaucher, Martino Bernasconi, Matteo Castiglioni, Andrea Celli, and Vianney Perchet. Feature-based online bilateral trade. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=xnF2U0ro7b.

Kathleen M. Hagerty and William P. Rogerson. Robust trading mechanisms. *Journal of Economic Theory*, 42(1):94–107, 1987. doi: 10.1016/0022-0531(87)90104-9. URL https://doi.org/10.1016/0022-0531(87)90104-9.

Yaonan Jin. Tight regret bounds for bilateral trade under semi feedback. *arXiv preprint arXiv:2601.16412*, 2026. URL https://arxiv.org/abs/2601.16412.

Zi Yang Kang, Francisco Pernice, and Jan Vondrák. Fixed-price approximations in bilateral trade. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 2964–2985. Society for Industrial and Applied Mathematics, 2022. URL https://arxiv.org/abs/2107.14327.

Anna Lunghi, Matteo Castiglioni, and Alberto Marchesi. Online two-sided markets: Many buyers enhance learning. *arXiv preprint arXiv:2503.01529*, 2025. URL https://arxiv.org/abs/2503.01529.

Anna Lunghi, Matteo Castiglioni, and Alberto Marchesi. Better regret rates in bilateral trade via sublinear budget violation. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 6494–6536. Society for Industrial and Applied Mathematics, 2026a. doi: 10.1137/1.9781611978971.233. URL https://doi.org/10.1137/1.9781611978971.233.

Anna Lunghi, Mattia Piccinato, Matteo Castiglioni, and Alberto Marchesi. A stronger benchmark for online bilateral trade: From fixed prices to distributions. In *Proceedings of the Forty-Third International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*. PMLR, 2026b. URL https://arxiv.org/abs/2602.05681.

Roger B. Myerson and Mark A. Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281, 1983. doi: 10.1016/0022-0531(83)90048-0. URL https://doi.org/10.1016/0022-0531(83)90048-0.

A Further Related Work

Table 2 collects the known regret bounds against the best fixed price, and Table 3 those of GBB one-bit-feedback mechanisms against both benchmarks; our results, shown in bold, close the two remaining gaps. In both tables, IV and BD abbreviate independent valuations and bounded density, respectively.

Table 2: Regret bounds of fixed-price mechanisms against the best fixed price, up to polylogarithmic factors.

Feedback	Valuations	SBB	WBB	GBB
Full	i.i.d.+IV i.i.d.	$T^{1/2}$ Cesa-Bianchi et al., 2024b; Bernasconi et al., 2024; Chen et al., 2026		
	adversarial	T Cesa-Bianchi et al., 2024b; Azar et al., 2024		$T^{1/2}$ Bernasconi et al., 2024
Semi	i.i.d.+IV i.i.d.	$T^{2/3}$ Chen et al., 2026; Lunghi et al., 2025		
	adversarial	T Azar et al., 2024		$T^{2/3}$ Chen et al., 2026; Jin, 2026
Two-bit	i.i.d.+IV+BD	$T^{2/3}$ Cesa-Bianchi et al., 2024b; Chen et al., 2026		
One-bit	i.i.d.+IV+BD	T Cesa-Bianchi et al., 2024b	$\boldsymbol{T^{2/3}}$ Cesa-Bianchi et al., 2024b; **this paper**	$T^{2/3}$ Chen et al., 2026
One- or two-bit	i.i.d.+BD	T Cesa-Bianchi et al., 2024b	$T^{3/4}$ Cesa-Bianchi et al., 2024b; Chen et al., 2026	
	i.i.d.+IV	T Cesa-Bianchi et al., 2024b; Chen et al., 2026		$T^{2/3}$ Chen et al., 2026
	i.i.d. adversarial	T Cesa-Bianchi et al., 2024b; Azar et al., 2024; Chen et al., 2026		$T^{3/4}$ Chen et al., 2026; Bernasconi et al., 2024

Note: Two valuation classes listed in one row share the same bounds. BD denotes a bounded joint density, which under IV is equivalent to bounded marginal densities up to constants. Unrestricted i.i.d. and adaptive σ-smooth valuations are incomparable; our product class is $1/M^2$-smooth.

Table 3: Regret bounds of GBB one-bit-feedback mechanisms against the two benchmarks, up to polylogarithmic factors.

Valuations	Best fixed price	Best feasible distribution
i.i.d.+IV+BD	$T^{2/3}$ Chen et al., 2026	$\boldsymbol{T^{2/3}}$ Chen et al., 2026; **this paper**
i.i.d.+BD	$T^{3/4}$ Cesa-Bianchi et al., 2024b; Chen et al., 2026	$T^{3/4}$ Lunghi et al., 2026b
i.i.d.	$T^{3/4}$ Chen et al., 2026; Bernasconi et al., 2024	T Lunghi et al., 2026b
adversarial	$T^{3/4}$ Chen et al., 2026; Bernasconi et al., 2024	T Bernasconi et al., 2024

Cesa-Bianchi et al. (2024b) gave a WBB one-bit mechanism with $\widetilde{\mathcal{O}}(T^{3/4})$ regret for i.i.d. valuations with a bounded joint density.[2] Their $\Omega\left(T^{2/3}\right)$ lower-bound construction already uses independent valuations with bounded densities.[3] Subsequent work (Cesa-Bianchi et al., 2024a) improved the lower-bound construction and achieved the tight regret bound, even under the more informative two-bit feedback; however, the corresponding hard instances are correlated and therefore do not cover the more structured case of independent valuation distributions. This left open the tight regret rate for WBB one-bit-feedback fixed-price mechanisms under independent valuations with bounded densities, which is the first of the two questions this paper settles.

The distribution benchmark was introduced by Bernasconi et al. (2024), who showed that no GBB mechanism attains sublinear α-regret against it on adversarial valuations for any $\alpha < \frac{36}{35}$, even under full feedback. For stochastic valuations, Lunghi et al. (2026b) proved that the regret is linear without a bounded density, even under two-bit feedback, and gave a one-bit mechanism with $\widetilde{\mathcal{O}}(T^{3/4})$ regret when the joint density is bounded. Under independent valuations with bounded densities, the best known upper bound against the distribution benchmark thus remained $\widetilde{\mathcal{O}}(T^{3/4})$, above the $\Omega\left(T^{2/3}\right)$ lower bound inherited from Chen et al. (2026); closing this gap is the second of the two questions this paper settles.

B Deferred Proofs for the WBB Algorithm

B.1 The Recursive Decomposition

The proof of Lemma 1 uses the following observation.

Observation 1. *For a real-valued random variable V and an interval $[a, b]$ with $a < b$, let $X \sim \mathrm{Unif}[a, b]$ be independent of V. Then we have $\int_a^b \Pr[V \le x]\,\mathrm{d}x = (b-a)\cdot\Pr[V \le X]$.*

Lemma 1. *For a valuation pair $(s, b) \sim \mu_S \otimes \mu_B$ and prices p, z, q with $0 \le p < z < q \le 1$, let $X \sim \mathrm{Unif}[p, z]$ and $Y \sim \mathrm{Unif}[z, q]$ be independent of (s, b). Then we have*

$$\underbrace{\Pr[s \le z,\, b \ge z]}_{\text{(I)}} U_S(z) = \underbrace{\Pr[s \le p,\, b \ge z]}_{\text{(II)}} U_S(p) + (z-p)\underbrace{\Pr[s \le X,\, b \ge z]}_{\text{(III)}},$$

$$\underbrace{\Pr[s \le z,\, b \ge z]}_{\text{(I)}} U_B(z) = \underbrace{\Pr[s \le z,\, b \ge q]}_{\text{(IV)}} U_B(q) + (q-z)\underbrace{\Pr[s \le z,\, b \ge Y]}_{\text{(V)}}.$$

Proof. We prove the seller identity. If $\Pr[s \le z,\, b \ge z] = 0$, both sides vanish, so assume it is positive. By the definition of the utility,

$$\begin{aligned}
U_S(z) &= \mathbb{E}\left[z - s \mid s \le z\right] \\
\triangleright \text{ law of total expectation} \quad &= \tfrac{\Pr[s\le p]\cdot\mathbb{E}[z-s\mid s\le p]+\Pr[p<s\le z]\cdot\mathbb{E}[z-s\mid p<s\le z]}{\Pr[s\le z]} \\
&= \tfrac{\Pr[s\le p]\cdot U_S(p)+(z-p)\cdot\Pr[s\le p]+\Pr[p<s\le z]\cdot\mathbb{E}[z-s\mid p<s\le z]}{\Pr[s\le z]} \\
&= \tfrac{\Pr[s\le p]\cdot U_S(p)+\int_p^z \Pr[s\le x]\,\mathrm{d}x}{\Pr[s\le z]} \\
&= \tfrac{\Pr[s\le p]\cdot U_S(p)+(z-p)\cdot\Pr[s\le X]}{\Pr[s\le z]} \qquad \text{Observation 1} \\
&= \tfrac{\Pr[b\ge z]\cdot\left(\Pr[s\le p]\cdot U_S(p)+(z-p)\cdot\Pr[s\le X]\right)}{\Pr[b\ge z]\cdot\Pr[s\le z]} \\
\triangleright \text{ independent valuations} \quad &= \tfrac{\Pr[s\le p,\, b\ge z]\cdot U_S(p)+(z-p)\cdot\Pr[s\le X,\, b\ge z]}{\Pr[s\le z,\, b\ge z]}.
\end{aligned}$$

This is equivalent to the seller identity in Lemma 1. The buyer identity follows symmetrically by exchanging the seller's lower-tail decomposition with the buyer's upper-tail decomposition and using $Y \sim \mathrm{Unif}[z, q]$. □

[2] Their WBB reward is the agents' total utility, equal to standard GFT minus the nonnegative spread. Since the objectives coincide at common prices, their upper bound transfers to our benchmark.

[3] The buyer distribution is fixed and known, so (p, p) weakly dominates any WBB pair (p, q) in standard GFT without reducing information. Chen et al. (2026) later extend the lower bound to GBB under semi feedback.

B.2 The Estimation Subroutine

Algorithm 4 specifies the subroutine MIDPOINT-EST used by Algorithm 1. Here $\mathrm{clip}_{[a,\,b]}(\cdot)$ truncates an estimate to $[a, b]$.

Algorithm 4 MIDPOINT-EST

Input: p, q, $\widehat{U}_S(p)$, $\widehat{U}_B(q)$.

`/* (I)--(V) are the probabilities in Lemma 1 with` $z = \frac{p+q}{2}$ `*/`

initialize each of the five empirical trade frequencies below to zero
for $i = 1, \dots, n$ **do** `// estimate (I)`
  post $\left(\frac{p+q}{2}, \frac{p+q}{2}\right)$; $\widehat{\Pr}\left[s \le \frac{p+q}{2},\, b \ge \frac{p+q}{2}\right] \mathrel{+}= \frac{1}{n} Z_t\left(\frac{p+q}{2}, \frac{p+q}{2}\right)$
for $i = 1, \dots, n$ **do** `// estimate (II)`
  post $\left(p, \frac{p+q}{2}\right)$; $\widehat{\Pr}\left[s \le p,\, b \ge \frac{p+q}{2}\right] \mathrel{+}= \frac{1}{n} Z_t\left(p, \frac{p+q}{2}\right)$
for $i = 1, \dots, n$ **do** `// estimate (IV)`
  post $\left(\frac{p+q}{2}, q\right)$; $\widehat{\Pr}\left[s \le \frac{p+q}{2},\, b \ge q\right] \mathrel{+}= \frac{1}{n} Z_t\left(\frac{p+q}{2}, q\right)$
for $i = 1, \dots, n$ **do** `// estimate (III)`
  draw $X \sim \mathrm{Unif}\left[p, \frac{p+q}{2}\right]$, post $\left(X, \frac{p+q}{2}\right)$; $\widehat{\Pr}\left[s \le X,\, b \ge \frac{p+q}{2}\right] \mathrel{+}= \frac{1}{n} Z_t\left(X, \frac{p+q}{2}\right)$
for $i = 1, \dots, n$ **do** `// estimate (V)`
  draw $Y \sim \mathrm{Unif}\left[\frac{p+q}{2}, q\right]$, post $\left(\frac{p+q}{2}, Y\right)$; $\widehat{\Pr}\left[s \le \frac{p+q}{2},\, b \ge Y\right] \mathrel{+}= \frac{1}{n} Z_t\left(\frac{p+q}{2}, Y\right)$

`/* seller recursion from (I)--(III) */`

$$\widehat{U}_S\left(\tfrac{p+q}{2}\right) \leftarrow \mathrm{clip}_{\left[0,\,\frac{p+q}{2}\right]}\left(\frac{\widehat{\Pr}\left[s \le p,\, b \ge \frac{p+q}{2}\right] \widehat{U}_S(p) + \frac{q-p}{2}\widehat{\Pr}\left[s \le X,\, b \ge \frac{p+q}{2}\right]}{\widehat{\Pr}\left[s \le \frac{p+q}{2},\, b \ge \frac{p+q}{2}\right] + \eta}\right)$$

`/* buyer recursion from (I), (IV), and (V) */`

$$\widehat{U}_B\left(\tfrac{p+q}{2}\right) \leftarrow \mathrm{clip}_{\left[0,\,1-\frac{p+q}{2}\right]}\left(\frac{\widehat{\Pr}\left[s \le \frac{p+q}{2},\, b \ge q\right] \widehat{U}_B(q) + \frac{q-p}{2}\widehat{\Pr}\left[s \le \frac{p+q}{2},\, b \ge Y\right]}{\widehat{\Pr}\left[s \le \frac{p+q}{2},\, b \ge \frac{p+q}{2}\right] + \eta}\right)$$

$$\widehat{\mathsf{GFT}}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right) \leftarrow \mathrm{clip}_{[0,\,1]}\left(\left(\widehat{\Pr}\left[s \le \tfrac{p+q}{2},\, b \ge \tfrac{p+q}{2}\right] + \eta\right) \cdot \left(\widehat{U}_S\left(\tfrac{p+q}{2}\right) + \widehat{U}_B\left(\tfrac{p+q}{2}\right)\right)\right)$$

return $\widehat{\mathsf{GFT}}\left(\frac{p+q}{2}, \frac{p+q}{2}\right),\ \widehat{U}_S\left(\frac{p+q}{2}\right),\ \widehat{U}_B\left(\frac{p+q}{2}\right)$

B.3 Regret Analysis of the Algorithm

This subsection gives the full proof of Theorem 1, following the proof sketch after its statement. We first need the regularity of the diagonal GFT.

Lemma 3. *The map $z \mapsto \mathsf{GFT}(z, z)$ is M-Lipschitz on $[0, 1]$.*

Proof. Fix $0 \le x < y \le 1$. A trade occurs at the common price z exactly when $s \le z \le b$, so

$$\mathbb{1}\{s \le y \le b\} - \mathbb{1}\{s \le x \le b\} = \mathbb{1}\{x < s \le y \le b\} - \mathbb{1}\{s \le x \le b < y\}.$$

Both events on the right imply $s \le b$. Multiplying by $b - s$ and taking expectations,

$$\begin{aligned} \mathsf{GFT}(y, y) - \mathsf{GFT}(x, x) &\le \mathbb{E}\left[(b - s) \cdot \mathbb{1}\{x < s \le y \le b\}\right] \\ &\triangleright\ b - s \le 1 \quad \le \Pr[x < s \le y] \\ &\triangleright\ \text{density of } s \text{ at most } M \quad \le M(y - x), \end{aligned}$$

and symmetrically $\mathsf{GFT}(x, x) - \mathsf{GFT}(y, y) \le \Pr[x \le b < y] \le M(y - x)$. Hence

$$|\mathsf{GFT}(y, y) - \mathsf{GFT}(x, x)| \le M(y - x).$$

□

Lemma 4. *There is an event $\mathcal{E}$ with $\Pr[\mathcal{E}] \ge 1 - \frac{1}{T}$ such that every pair (p, q) evaluated by the algorithm satisfies*

$$\left|\widehat{\mathsf{GFT}}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right) - \mathsf{GFT}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right)\right| < \frac{\varepsilon}{2}.$$

Proof. Let $\mathcal{E}$ be the event on which every completed block's empirical trade frequency is within η of its conditional population mean. The complete dyadic tree contains fewer than $2K$ pairs and hence fewer than $10K$ possible blocks. Conditional on any block being opened and on its starting history, its n observations are independent Bernoulli variables with the corresponding conditional population mean. Hoeffding's inequality and a conditional union bound give

$$\begin{aligned} \Pr[\mathcal{E}^c] &\le 10K \cdot 2e^{-2n\eta^2} \\ \triangleright\ 2n\eta^2 \ge \log(20KT) \quad &\le \tfrac{1}{T}. \end{aligned}$$

We first control the seller-side contribution; the buyer-side argument is symmetric. On $\mathcal{E}$, the regularized midpoint probability exceeds its population counterpart by at most 2η, so the population weighted utility lies in the corresponding clipping interval. Together with $0 \le U_S(p) \le p$, this gives

$$\begin{aligned} &\left|\left(\widehat{\Pr}\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] + \eta\right) \cdot \widehat{U}_S\left(\tfrac{p+q}{2}\right) - \Pr\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] \cdot U_S\left(\tfrac{p+q}{2}\right)\right| \\ &\le \left|\widehat{\Pr}\left[s \le p,\ b \ge \tfrac{p+q}{2}\right] \cdot \widehat{U}_S(p) + \tfrac{q-p}{2} \cdot \widehat{\Pr}\left[s \le X,\ b \ge \tfrac{p+q}{2}\right]\right. \\ &\qquad \left. - \Pr\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] \cdot U_S\left(\tfrac{p+q}{2}\right)\right| \\ &\le \Pr\left[s \le p,\ b \ge \tfrac{p+q}{2}\right] \cdot \left|\widehat{U}_S(p) - U_S(p)\right| + \eta \cdot \widehat{U}_S(p) + \tfrac{q-p}{2}\eta \qquad \text{Lemma 1} \\ &\le \Pr[s \le p,\ b \ge p] \cdot \left|\widehat{U}_S(p) - U_S(p)\right| + \tfrac{3}{2}\eta. \end{aligned} \tag{2}$$

Combining eq. (2) with midpoint-block concentration gives

$$\begin{aligned} &\Pr\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] \cdot \left|\widehat{U}_S\left(\tfrac{p+q}{2}\right) - U_S\left(\tfrac{p+q}{2}\right)\right| \\ &\le \Pr[s \le p,\ b \ge p] \cdot \left|\widehat{U}_S(p) - U_S(p)\right| + \tfrac{7}{2}\eta. \end{aligned} \tag{3}$$

Since the estimates at 0 and 1 are exact and every other endpoint is the midpoint of an ancestor estimated by MIDPOINT-EST, induction using eq. (3) and its buyer-side analogue gives

$$\begin{aligned} \Pr[s \le p,\ b \ge p] \cdot \left|\widehat{U}_S(p) - U_S(p)\right| &\le \frac{7}{2} j\eta, \\ \Pr[s \le q,\ b \ge q] \cdot \left|\widehat{U}_B(q) - U_B(q)\right| &\le \frac{7}{2} j\eta. \end{aligned} \tag{4}$$

Since the final clipping in Algorithm 4 is nonexpansive, applying eq. (4) to the seller and buyer sides gives

$$\begin{aligned} &\left|\widehat{\mathsf{GFT}}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right) - \mathsf{GFT}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right)\right| \\ &\le \left|\left(\widehat{\Pr}\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] + \eta\right) \cdot \widehat{U}_S\left(\tfrac{p+q}{2}\right) - \Pr\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] \cdot U_S\left(\tfrac{p+q}{2}\right)\right| \\ &\quad + \left|\left(\widehat{\Pr}\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] + \eta\right) \cdot \widehat{U}_B\left(\tfrac{p+q}{2}\right) - \Pr\left[s \le \tfrac{p+q}{2},\ b \ge \tfrac{p+q}{2}\right] \cdot U_B\left(\tfrac{p+q}{2}\right)\right| \\ &\le 2\left(\tfrac{7}{2} j\eta + \tfrac{3}{2}\eta\right) = (7j+3)\eta < \tfrac{\varepsilon}{2}. \end{aligned}$$

Since each utility error is weighted by its trade probability, the argument never divides by the midpoint trade probability and holds even when it is zero. □

Together, Lemmas 3 and 4 imply the following guarantee.

Lemma 5. *On $\mathcal{E}$, for every depth $j \le J$, the optimal diagonal price $z^\star$ satisfies $p \le z^\star \le q$ for at least one pair $(p, q) \in \mathcal{A}_j$. Moreover, every pair (p, q) evaluated at depth $j \le J$ satisfies*

$$\mathsf{GFT}^\star_{\mathrm{W}} - \mathsf{GFT}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right) \le \frac{5}{2} M 2^{-j} + 3\varepsilon.$$

Proof. On $\mathcal{E}$, we first show that an interval containing the optimizer is never pruned. The root interval contains $z^\star$. For $j < J$, let $(p_j^\star, q_j^\star) \in \mathcal{A}_j$ bracket $z^\star$, i.e., $p_j^\star \le z^\star \le q_j^\star$, and write $z_j^\star = (p_j^\star + q_j^\star)/2$. Every interval in $\mathcal{A}_j$ has width 2^{-j}, so $\left|z_j^\star - z^\star\right| \le 2^{-j-1}$ and

$$\widehat{\mathsf{GFT}}\left(z_j^\star, z_j^\star\right) + 2\varepsilon + M 2^{-j-1} \ge \mathsf{GFT}\left(z_j^\star, z_j^\star\right) + \tfrac{3\varepsilon}{2} + M 2^{-j-1} \qquad \text{Lemma 4}$$

$$\geq \mathsf{GFT}^{\star}_{\mathrm{W}} + \tfrac{3\varepsilon}{2} \qquad \text{Lemma 3}$$
$$\geq \max_{(p,q)\in\mathcal{A}_j} \widehat{\mathsf{GFT}}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right). \qquad \text{Lemma 4}$$

Hence the retention rule in Algorithm 1, line 8, keeps the interval, and one of its children again contains the optimizer. Induction gives an optimizer-containing interval at every depth.

For an interval (p, q) retained at depth $j < J$, using Lemma 3 in the last step,

$$\begin{aligned}
\mathsf{GFT}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right) &\geq \widehat{\mathsf{GFT}}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right) - \tfrac{\varepsilon}{2} && \text{Lemma 4}\\
\triangleright \text{ Algorithm 1, line 8} \quad &\geq \widehat{\mathsf{GFT}}\left(z_j^{\star}, z_j^{\star}\right) - \tfrac{5\varepsilon}{2} - M2^{-j-1} &&\\
&\geq \mathsf{GFT}\left(z_j^{\star}, z_j^{\star}\right) - 3\varepsilon - M2^{-j-1} && \text{Lemma 4}\\
&\geq \mathsf{GFT}^{\star}_{\mathrm{W}} - M2^{-j} - 3\varepsilon. && (5)
\end{aligned}$$

Let $z = (p+q)/2$ be the evaluated midpoint of $(p, q) \in \mathcal{A}_j$, so $z = 1/2$ at depth 0; for $j \geq 1$, let $\bar{z}$ be its retained parent's midpoint in $\mathcal{A}_{j-1}$.

$$\begin{aligned}
\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(1/2, 1/2\right) &\leq \tfrac{M}{2} \leq \tfrac{5}{2}M + 3\varepsilon, && \text{Lemma 3}\\
\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(z, z\right) &= \mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(\bar{z}, \bar{z}\right) + \mathsf{GFT}\left(\bar{z}, \bar{z}\right) - \mathsf{GFT}\left(z, z\right) &&\\
&\leq M2^{-(j-1)} + 3\varepsilon + M\left|z - \bar{z}\right| && \text{eq. (5); Lemma 3}\\
\triangleright\ |z - \bar{z}| = 2^{-j-1} \quad &\leq M2^{-(j-1)} + 3\varepsilon + M2^{-j-1} = \tfrac{5}{2}M2^{-j} + 3\varepsilon. &&
\end{aligned}$$

□

Putting the preceding estimates together, we now prove the regret guarantee of Algorithm 1.

Theorem 1. *For GFT maximization in repeated bilateral trade with one-bit feedback, where the valuations are i.i.d. draws from a product distribution with bounded density,* TREE-SEARCH *(Algorithm 1) satisfies WBB and achieves regret $\widetilde{\mathcal{O}}(T^{2/3})$ against the best fixed price.*

Proof of Theorem 1. By the tower property and splitting over the good event $\mathcal{E}$,

$$\begin{aligned}
\mathsf{REG}_{\mathrm{W}} = \mathbb{E}\left[\textstyle\sum_{t\in[T]}\left(\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(p_t, q_t\right)\right)\cdot \mathbb{1}\left\{\mathcal{E}\right\}\right]\\
+ \mathbb{E}\left[\textstyle\sum_{t\in[T]}\left(\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(p_t, q_t\right)\right)\cdot \mathbb{1}\left\{\mathcal{E}^c\right\}\right].
\end{aligned}$$

By Lemma 4, the second term is at most $T \cdot \Pr[\mathcal{E}^c] \leq 1$.

On $\mathcal{E}$, the first term decomposes into four contributions: midpoint suboptimality, off-diagonal posting loss, finite-resolution loss, and final selection loss.

$$\begin{aligned}
\text{Regret on } \mathcal{E} &= \mathbb{E}\left[\textstyle\sum_{t\in[T]}\left(\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(p_t, q_t\right)\right)\cdot \mathbb{1}\left\{\mathcal{E}\right\}\right]\\
&= \mathbb{E}\left[\textstyle\sum_{j=0}^{J}\sum_{(p,q)\in\mathcal{A}_j}\sum_{t\,\in\text{ rounds of }(p,q)}\left(\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right)\right)\cdot \mathbb{1}\left\{\mathcal{E}\right\}\right] && (6)\\
&\quad + \mathbb{E}\left[\textstyle\sum_{j=0}^{J}\sum_{(p,q)\in\mathcal{A}_j}\sum_{t\,\in\text{ rounds of }(p,q)}\left(\mathsf{GFT}\left(\tfrac{p+q}{2}, \tfrac{p+q}{2}\right) - \mathsf{GFT}\left(p_t, q_t\right)\right)\cdot \mathbb{1}\left\{\mathcal{E}\right\}\right] && (7)\\
&\quad + \mathbb{E}\left[\textstyle\sum_{t\,\in\text{ exploitation rounds}}\left(\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(z_J^{\star}, z_J^{\star}\right)\right)\cdot \mathbb{1}\left\{\mathcal{E}\right\}\right] && (8)\\
&\quad + \mathbb{E}\left[\textstyle\sum_{t\,\in\text{ exploitation rounds}}\left(\mathsf{GFT}\left(z_J^{\star}, z_J^{\star}\right) - \mathsf{GFT}\left(\widehat{z}, \widehat{z}\right)\right)\cdot \mathbb{1}\left\{\mathcal{E}\right\}\right]. && (9)
\end{aligned}$$

We bound the four terms in order.

$$\begin{aligned}
\textit{eq. (6)} &\leq \mathbb{E}\left[\textstyle\sum_{j=0}^{J}\sum_{(p,q)\in\mathcal{A}_j} 5n\left(\tfrac{5}{2}M2^{-j} + 3\varepsilon\right)\right] && \text{Lemma 5}\\
\triangleright\ |\mathcal{A}_j| \leq 2^j \quad &\leq \textstyle\sum_{j=0}^{J} 5n\,2^j\left(\tfrac{5}{2}M2^{-j} + 3\varepsilon\right)\\
&= \tfrac{25}{2}nM(J+1) + 15n\varepsilon\cdot\textstyle\sum_{j=0}^{J}2^j\\
\triangleright\ \textstyle\sum_{j=0}^{J}2^j < 2K \quad &< \tfrac{25}{2}nM(J+1) + 30n\varepsilon K
\end{aligned}$$

$$
\begin{aligned}
\triangleright\, \varepsilon K = M \quad &= \tfrac{25}{2} nM(J+1) + 30nM. \\
\textit{eq. (7)} \quad &\le \mathbb{E}\left[\textstyle\sum_{j=0}^{J} \sum_{(p,q)\in\mathcal{A}_j} 5n \cdot \tfrac{1}{2} M 2^{-j}\right] \\
\triangleright\, |\mathcal{A}_j| \le 2^j \quad &\le \textstyle\sum_{j=0}^{J} 5n\, 2^j \cdot \tfrac{1}{2} M 2^{-j} \\
&= \tfrac{5}{2} nM(J+1). \\
\textit{eq. (8)} \quad &\le T \cdot \mathbb{E}\left[\left(\mathsf{GFT}^{\star}_{\mathrm{W}} - \mathsf{GFT}\left(z_J^{\star}, z_J^{\star}\right)\right) \cdot \mathbb{1}\left\{\mathcal{E}\right\}\right] \\
\triangleright\, |z_J^{\star} - z^{\star}| \le 2^{-J-1} \quad &\le \tfrac{1}{2} TM 2^{-J} \qquad \text{Lemma 3} \\
\triangleright\, 2^{-J} < 2K^{-1} \quad &< TM/K. \\
\textit{eq. (9)} \quad &\le T \cdot \mathbb{E}\left[\max\left\{0,\ \mathsf{GFT}\left(z_J^{\star}, z_J^{\star}\right) - \mathsf{GFT}\left(\widehat{z}, \widehat{z}\right)\right\} \cdot \mathbb{1}\left\{\mathcal{E}\right\}\right] \\
\triangleright\, \widehat{z} \text{ maximizes } \widehat{\mathsf{GFT}} \text{ over } \mathcal{A}_J \quad &\le T\varepsilon \qquad \text{Lemma 4} \\
\triangleright\, \varepsilon = M/K \quad &= TM/K.
\end{aligned}
$$

Adding the four bounds together with the regret on the bad event,

$$
\begin{aligned}
\mathsf{REG}_{\mathrm{W}} &\le 15nM(J+1) + 30nM + \tfrac{2TM}{K} + 1 \\
&\le 45nM(J+1) + \tfrac{2TM}{K} + 1.
\end{aligned}
$$

Then, by substituting $K = T^{1/3}(\log T)^{-4/3}$, $J = \lfloor \log_2 K \rfloor$, $\eta = \frac{M}{16K(J+1)}$, and $n = \left\lceil \frac{\log(20KT)}{2\eta^2} \right\rceil$, it follows that

$$
\begin{aligned}
\mathsf{REG}_{\mathrm{W}} &= \mathcal{O}\big(K^2(J+1)^3 \log(20KT) + \tfrac{T}{K}\big) \\
&= \mathcal{O}\big(T^{2/3}(\log T)^{4/3}\big) = \widetilde{\mathcal{O}}(T^{2/3}).
\end{aligned}
$$

This finishes the proof of Theorem 1. □

C Deferred Proofs for the GBB Algorithm

This appendix gives the refinement subroutine and the full proof of Theorem 2.

C.1 The Recursive Decomposition

Lemma 2. *For a valuation pair $(s,b) \sim \mu_S \otimes \mu_B$ and an interval $I = [u,v] \subseteq [0,1]$ with $u < v$ and width $|I| = v - u$, let $X \sim \mathrm{Unif}(I)$ be independent of (s,b). Then, for any fixed buyer price $q \in [0,1]$ and any fixed seller price $p \in [0,1]$,*

$$
\begin{aligned}
\Pr[s \le v,\ b \ge q] \cdot U_S(v) &= \Pr[s \le u,\ b \ge q] \cdot U_S(u) + |I| \cdot \Pr[s \le X,\ b \ge q], \\
\Pr[s \le p,\ b \ge u] \cdot U_B(u) &= \Pr[s \le p,\ b \ge v] \cdot U_B(v) + |I| \cdot \Pr[s \le p,\ b \ge X].
\end{aligned}
$$

Proof. For every $x \in [0,1]$, the definition of the seller's conditional utility gives

$$
\Pr[s \le x] \cdot U_S(x) = \int_0^x \Pr[s \le t]\ \mathrm{d}t.
$$

By independence and Observation 1, subtracting this identity at u from that at v and multiplying by $\Pr[b \ge q]$ yields the seller recursion. The buyer recursion follows symmetrically. □

C.2 The Refinement Subroutine

Algorithm 5 specifies the subroutine Local-Refine used by Algorithm 3. The count $N_{I,\ell}$ records trades obtained when arm ℓ is probed at the typed interval I, and $N_I = \sum_\ell N_{I,\ell}$ is the total. These counts are shared across calls, including calls with different opposite-side prices, because independence makes the normalized arm probabilities depend only on the valuation distribution on the side of I. The upper and lower probability envelopes in the subroutine are taken over its input candidate family $\mathcal{F}$, with the epoch subscript suppressed. When $N_I = 0$, the estimator imposes no empirical restriction. For a candidate with zero anchor probability, we interpret the ratio constraints in cross-multiplied form and set its arm ratios to 0, so that every arm ratio, including those in ϕ_k, is well defined and lies in $[0,1]$. All branches that post $(0,1)$ for the remaining rounds terminate the algorithm.

Algorithm 5 LOCAL-REFINE

function LOCAL-REFINE($\gamma, (p,q), I, \mathcal{F}$)
  Input: γ, $(p,q) \in \mathrm{supp}(\gamma)$, I, $\mathcal{F}$.
  $N_I \leftarrow \sum_\ell N_{I,\ell}$
  $m \leftarrow \left\lceil \dfrac{4\max\{\eta n, N_I\}}{\gamma(p,q)\cdot \Pr[a_{I,0}]} \right\rceil$
  if $t + m - 1 > \frac{T+\tau}{2}$ **then**
    post $(0,1)$ in every remaining round
  for $i = 1, \ldots, m$ **do** // the refinement rounds of this call
    if $\Pi_{t-1} < 1$ **then**
      post $(0,1)$ in every remaining round
    draw $(p_t, q_t) \sim \gamma$
    if $(p_t, q_t) = (p,q)$ **then**
      draw $\ell \sim \mathrm{Unif}\{0,1,2\}$ and $X \sim \mathrm{Unif}(I)$
      $(p_t, q_t) \leftarrow$ the prices of arm ℓ at I
      post (p_t, q_t)
      $N_{I,\ell} \leftarrow N_{I,\ell} + 1$ if $Z_t(p_t, q_t) = 1$
    else
      post (p_t, q_t)
  return $\left\{\theta \in \mathcal{F} : \left| \dfrac{N_{I,\ell}}{\sum_{\ell'} N_{I,\ell'}} - \dfrac{\Pr_\theta[a_{I,\ell}]}{\sum_{\ell'} \Pr_\theta[a_{I,\ell'}]} \right| \right.$
    $\left. \leq \sqrt{\dfrac{\log(1224KT(\sum_{\ell'} N_{I,\ell'} + 1)^2)}{2\sum_{\ell'} N_{I,\ell'}}} \text{ for all } \ell \right\}$

C.3 REGRET ANALYSIS OF THE ALGORITHM

We first outline the proof of Theorem 2. As in the WBB analysis, Lipschitz continuity controls the cost of local exploration: Lemma 6 and Corollary 1 show that probing a node I changes GFT and profit by $\mathcal{O}(|I|)$. Let $\theta^\star \in \mathcal{F}$ denote the true valuation law. We work on the confidence event $\mathcal{E}$ of Lemma 7, on which the true law survives every elimination and each completed call in the unstopped suffix adds at least $\max\{\eta n, N_I\}$ trades. The growing counts shrink the confidence bounds for the local probability ratios.

Since $\theta^\star$ survives, optimism bounds the gap between the grid benchmark (the best feasible distribution on $\mathcal{G}_K$, defined below) and γ_k by the range of its GFT over the candidate family. Lemma 8 controls both this range and the profit range by $\mathcal{O}(J)\cdot\gamma_k(p,q)\cdot\phi_k((p,q),I)$ at the selected pair and node. At termination, the weighted uncertainty is at most η, giving the $\mathcal{O}(1/K)$ GFT shortfall and expected profit deficit in Corollary 2.

The $(2K+1)n$ boundary exploration rounds cost $\widetilde{\mathcal{O}}(K^2)$ regret. During refinement, a round incurs the gap from the grid benchmark to γ_k and the loss from replacing its selected pair by a local probe. Lemma 9 bounds the accumulated weighted uncertainty at each node, giving $\widetilde{\mathcal{O}}(K^2)$ for the first contribution after summing over the nodes. The same bound and the selection threshold limit each node to $\widetilde{\mathcal{O}}(1/\eta^2)$ refinement rounds. Since nodes at each depth have total width $\mathcal{O}(1)$, the local probing losses also sum to $\widetilde{\mathcal{O}}(K^2)$. Corollary 3 ensures that exploration fits within its allotted half of the remaining horizon.

The GBB rule keeps the cumulative profit nonnegative on every sample path. On $\mathcal{E}$, apart from the additional failure event in Lemma 10, the reserve covers the cumulative deficit of the second stage, so this rule never intervenes. The first stage costs $\widetilde{\mathcal{O}}(T^{2/3})$ regret by Lunghi et al. (2026b, Lemma 7.2) and the choice of β. Discretization and exploitation contribute $\widetilde{\mathcal{O}}((T-\tau)/K)$; choosing the largest admissible $K = \widetilde{\Theta}((T-\tau)^{1/3})$ balances this with the $\widetilde{\mathcal{O}}(K^2)$ exploration cost. The cases $\tau = T$ and $K = 0$ are handled directly in the proof. As in the WBB analysis, the confidence, GBB and first-stage guarantees fail only with probability $\mathcal{O}(1/T)$, which costs $\mathcal{O}(1)$ regret, and combining them gives the claimed bound.

Throughout this section, $\mathsf{REG}_{\mathrm{G}}$ is the regret against $\mathsf{GFT}^{\star}_{\theta^\star} = \mathsf{GFT}^{\star}_{\mathrm{G}}$, and we write ℓ_t for the arm sampled when round t is a probe. The corresponding grid benchmark is

$$\mathsf{GFT}^{\star}_{K,\theta} \triangleq \max_{\gamma \in \Delta(\mathcal{G}_K)} \mathbb{E}_{(p,q)\sim\gamma}\left[\mathsf{GFT}_\theta(p,q)\right] \quad \text{s.t.} \quad \mathbb{E}_{(p,q)\sim\gamma}\left[\mathsf{PRO}_\theta(p,q)\right] \geq 0.$$

For the bounded-density class used here, the discretization bound of Lunghi et al. (2026b, Lemma 5.2) reads

$$\mathsf{GFT}^{\star}_{\theta^\star} \leq \mathsf{GFT}^{\star}_{K,\theta^\star} + \mathcal{O}\big(\tfrac{\log(17T^2)}{K}\big).$$

Thus it suffices to compete with distributions supported on the finite grid $\mathcal{G}_K$, up to the discretization error above.

The first stage runs the PROFIT-MAX algorithm of Bernasconi et al. (2024), which keeps the cumulative profit nonnegative on every sample path and accumulates a reserve of β by the time it stops,

$$\Pi_\tau \geq \beta \quad (\tau < T), \qquad \Pi_t \geq 0 \quad (0 \leq t \leq \tau).$$

Its regret against the distribution benchmark is bounded by the guarantee of Lunghi et al. (2026b, Lemma 7.2), which holds with probability at least $1 - \frac{2}{17T}$.

We now turn to the suffix. GBB moves one price of an off-diagonal pair at a time and must also track the profit, so we first extend Lemma 3 off the diagonal.

Lemma 6. *For every $\theta \in \mathcal{F}$, $\mathsf{GFT}_\theta(p,q)$ is M-Lipschitz and $\mathsf{PRO}_\theta(p,q)$ is $(1+M)$-Lipschitz in each of p and q on $[0,1]$, the other price held fixed.*

Proof. Fix q and $0 \leq p < p' \leq 1$. The two pairs differ only on trades with $p < s \leq p'$, so

$$\begin{aligned} \left|\mathsf{GFT}_\theta(p',q) - \mathsf{GFT}_\theta(p,q)\right| &= \left|\mathbb{E}_\theta\left[(b-s)\cdot \mathbb{1}\left\{p < s \leq p',\ b \geq q\right\}\right]\right| \\ \triangleright\ |b-s| \leq 1 \quad &\leq \Pr_\theta[p < s \leq p'] \\ \triangleright\ \text{density of } s \text{ at most } M \quad &\leq M\left|p - p'\right|. \end{aligned}$$

The expected-profit identity gives

$$\begin{aligned} \left|\mathsf{PRO}_\theta(p',q) - \mathsf{PRO}_\theta(p,q)\right| &= \left|(q-p')\cdot \Pr_\theta[s \leq p',\ b \geq q] - (q-p)\cdot \Pr_\theta[s \leq p,\ b \geq q]\right| \\ &\leq \left|p-p'\right| \cdot \Pr_\theta[s \leq p',\ b \geq q] \\ &\quad + \left|q-p\right| \left|\Pr_\theta[s \leq p',\ b \geq q] - \Pr_\theta[s \leq p,\ b \geq q]\right| \\ \triangleright\ \text{density at most } M \quad &\leq (1+M)\left|p-p'\right|. \end{aligned}$$

The buyer price is symmetric, with $\Pr_\theta[q \leq b < q']$ in place of $\Pr_\theta[p < s \leq p']$. □

An arm of a typed node I on a path of (p,q) moves one coordinate of (p,q) by less than $2\left|I\right|$, so Lemma 6 gives the following.

Corollary 1. *Let I be a typed node on a path of $(p,q) \in \mathcal{G}_K$, and let (p',q') be an arm of I. For every $\theta \in \mathcal{F}$,*

$$\left|\mathsf{GFT}_\theta(p',q') - \mathsf{GFT}_\theta(p,q)\right| < 2M\left|I\right|,$$

and

$$\left|\mathsf{PRO}_\theta(p',q') - \mathsf{PRO}_\theta(p,q)\right| < 2(1+M)\left|I\right|.$$

The next four lemmas assume $K \geq 8$, or equivalently, that the set defining K is nonempty. We analyze the unstopped suffix, the suffix run on the same sample path with the GBB and horizon rules of Algorithms 3 and 5 removed; it still ends at round T, so its number of calls and every count are at most T. Corollary 3 and Lemma 10 show that on $\mathcal{E}$ neither rule is triggered, except on an event of probability at most $\frac{1}{17T}$. In either suffix, every call of Algorithm 5 has $\underline{\Pr}_k[a_{I,0}] > 0$ (see the proof of Lemma 9), so the block length in line 3 is well defined.

Lemma 7. *There is an event $\mathcal{E}$ with $\Pr[\mathcal{E}] \geq 1 - \frac{8}{17T}$ on which*

(i) $\theta^\star \in \mathcal{F}_k$ for every $k \geq 0$;

(ii) every completed call of Algorithm 5 in the unstopped suffix adds at least $\max\{\eta n, N_I\}$ *trades to the count* N_I *at its node.*

Proof. By the definition of $\mathcal{F}_0$ in Algorithm 3, line 9, and a union bound over the $2K+1$ boundary probes,

$$\begin{aligned}\Pr[\theta^\star \notin \mathcal{F}_0] &\le \textstyle\sum_{(p,q)} \Pr\Big[\mathrm{KL}\big(\widehat{\Pr}[s \le p,\ b \ge q]\,, \Pr_{\theta^\star}[s \le p,\ b \ge q]\big) > \tfrac{\eta}{128}\Big] \\ \triangleright \text{ Chernoff bound} \quad &\le 2(2K+1)e^{-n\eta/128} \\ \triangleright\ n\eta/128 \ge \log(17T^2) \quad &\le \tfrac{2(2K+1)}{17T^2} \\ \triangleright\ K \le T \quad &\le \tfrac{6}{17T}.\end{aligned}$$

Fix a round in an estimator call at $(p,q), I$. Algorithm 5 gives

$$\begin{aligned}\Pr[\ell_t = \ell,\ Z_t = 1 \mid \mathcal{H}_{t-1}] &= \tfrac{\gamma_k(p,q)}{3} \cdot \Pr_{\theta^\star}[a_{I,\ell}]\,, \\ \tfrac{\Pr[\ell_t=\ell,\ Z_t=1 \mid \mathcal{H}_{t-1}]}{\sum_{\ell'=0}^{2} \Pr[\ell_t=\ell',\ Z_t=1 \mid \mathcal{H}_{t-1}]} &= \tfrac{\Pr_{\theta^\star}[a_{I,\ell}]}{\sum_{\ell'=0}^{2} \Pr_{\theta^\star}[a_{I,\ell'}]}.\end{aligned}$$

Since this conditional law does not depend on the scheduling, Hoeffding's inequality and a union bound give

$$\begin{aligned}\Pr\left[\exists\, I, N_I, \ell : \left|\tfrac{N_{I,\ell}}{N_I} - \tfrac{\Pr_{\theta^\star}[a_{I,\ell}]}{\sum_{\ell'=0}^{2} \Pr_{\theta^\star}[a_{I,\ell'}]}\right| > \sqrt{\tfrac{\log(1224KT(N_I+1)^2)}{2N_I}}\right] &\le 12K \cdot \textstyle\sum_{N_I \ge 1} 2 \\ &\quad \times \tfrac{1}{1224KT(N_I+1)^2} \\ &= \tfrac{1}{51T} \cdot \textstyle\sum_{N_I \ge 1} \tfrac{1}{(N_I+1)^2} \\ &< \tfrac{1}{51T}.\end{aligned}$$

The intersection of these two events fails with probability less than $\frac{6}{17T} + \frac{1}{51T} < \frac{7}{17T}$. On it, $\theta^\star$ satisfies every constraint imposed by a completed elimination, so $\theta^\star \in \mathcal{F}_k$ for every $k \ge 0$, which is item (i).

For item (ii), fix a call that starts with N_I counts. Summing the conditional arm law over the three arms and the m rounds of the call,

$$\begin{aligned}\mathbb{E}\left[\textstyle\sum_\ell N_{I,\ell} - N_I \mid \mathcal{H}_{t-1}\right] &= \tfrac{m \cdot \gamma_k(p,q)}{3} \cdot \textstyle\sum_\ell \Pr_{\theta^\star}[a_{I,\ell}] \\ \triangleright\ \theta^\star \in \mathcal{F}_{k-1} \quad &\ge \tfrac{m \cdot \gamma_k(p,q) \cdot \underline{\Pr}_k[a_{I,0}]}{3} \\ \triangleright \text{ Algorithm 5, line 3} \quad &\ge \tfrac{4\max\{\eta n, N_I\}}{3}.\end{aligned}$$

Since $\eta n \ge 128 \log(17T^2)$, a conditional Chernoff bound and a union bound over the at most T calls give

$$\begin{aligned}T \cdot \Pr\left[\textstyle\sum_\ell N_{I,\ell} < N_I + \max\{\eta n, N_I\},\ \theta^\star \in \mathcal{F}_{k-1} \mid \mathcal{H}_{t-1}\right] &\le T \cdot \exp\big\{-\tfrac{\max\{\eta n, N_I\}}{24}\big\} \\ &\le T \cdot \exp\big\{-\tfrac{16}{3}\log(17T^2)\big\} \\ &\le \tfrac{1}{17T}.\end{aligned}$$

Let $\mathcal{E}$ be the event that item (i) holds and every completed call made while $\theta^\star \in \mathcal{F}_{k-1}$ adds at least $\max\{\eta n, N_I\}$ trades to N_I. On $\mathcal{E}$, item (i) ensures that every call is made while $\theta^\star \in \mathcal{F}_{k-1}$, so item (ii) holds as well. A union bound gives $\Pr[\mathcal{E}^c] < \frac{7}{17T} + \frac{1}{17T} = \frac{8}{17T}$. □

The following observations about PATH, illustrated in Figure 1, are used below.

Observation 2. *Index the intervals of a partition returned by* $\mathcal{D}$ *in increasing order as* $I_1, I_2, \ldots$*, so that* $\min I_{r+1} = \max I_r$*. Then* $a_{I_r,0} = a_{I_{r+1},1} = \{s \le \max I_r,\ b \ge q\}$ *for seller-typed* I_r, I_{r+1}*, and* $a_{I_r,1} = a_{I_{r+1},0}$ *on the buyer side, with the endpoints exchanged.*

Observation 3. *For* $(p,q) \in \mathcal{G}_K$ *the intervals of* $\mathcal{D}([0,p])$ *have pairwise distinct depths, and so do those of* $\bigcup_{I \in \mathcal{D}([0,p])} \mathcal{D}([\max I, 1])$*; hence* $|\mathcal{S}(p,q)| \le 2(J+1)$*, and* $|\mathcal{B}(p,q)| \le 2(J+1)$ *on the buyer side.*

Observation 4. *A binary tree of depth J has $2^{J+1}-1 = 2K-1$ nodes, so the seller and buyer trees carry fewer than $4K$ typed nodes between them and, with three arms at each node, fewer than $12K$ node–arm pairs.*

Lemma 8. *Let $((p,q),I)$ maximize the weighted pair–node uncertainty in epoch k (Algorithm 3, line 16). Then*

$$\overline{\mathsf{GFT}}_k(\gamma_k) - \underline{\mathsf{GFT}}_k(\gamma_k) \le 16(J+1)\cdot\gamma_k(p,q)\cdot\phi_k((p,q),I),$$

and

$$\overline{\mathsf{PRO}}_k(\gamma_k) - \underline{\mathsf{PRO}}_k(\gamma_k) \le 4(J+1)\cdot\gamma_k(p,q)\cdot\phi_k((p,q),I).$$

Proof. We give the seller side; the buyer side is the same argument with $\Pr_\theta[s \le \cdot\,]$ replaced by $\Pr_\theta[b \ge \cdot\,]$ and the two endpoints of each interval exchanged. Index $\mathcal{D}([0,p])$ as in Observation 2. The arms of I_r in Algorithm 5 satisfy a recursion in r:

$$\begin{aligned}\Pr_\theta[a_{I_r,0}]\,U_{S,\theta}(\max I_r) &= \Pr_\theta[a_{I_r,1}]\,U_{S,\theta}(\min I_r) + |I_r|\cdot\Pr_\theta[a_{I_r,2}] && \text{Lemma 2}\\ &= \Pr_\theta\left[a_{I_{r-1},0}\right]U_{S,\theta}(\max I_{r-1}) + |I_r|\cdot\Pr_\theta[a_{I_r,2}]\,. && \text{Observation 2}\end{aligned}$$

Summed over r, this recursion telescopes, and its base term $\Pr_\theta[a_{I_1,1}]\,U_{S,\theta}(\min I_1)$ vanishes because $U_{S,\theta}(0)=0$:

$$\Pr_\theta[s\le p,\ b\ge q]\,U_{S,\theta}(p) = \sum_r |I_r|\cdot\Pr_\theta[a_{I_r,2}]\,. \tag{10}$$

By Observation 2, the arm-1 ratios telescope along a path in the same way, leaving one marginal in each denominator, so the two paths returned by Algorithm 7 reconstruct the joint trade probability:

$$\begin{aligned}\textstyle\prod_{I\in\mathcal{D}([p,1])}\frac{\Pr_\theta[a_{I,1}]}{\Pr_\theta[a_{I,0}]}\cdot\prod_{I\in\mathcal{D}([0,q])}\frac{\Pr_\theta[a_{I,1}]}{\Pr_\theta[a_{I,0}]} &= \frac{\Pr_\theta[s\le p,\ b\ge q]}{\Pr_\theta[b\ge q]}\cdot\frac{\Pr_\theta[s\le p,\ b\ge q]}{\Pr_\theta[s\le p]}\\ &= \Pr_\theta[s\le p,\ b\ge q]\,.\end{aligned} \tag{11}$$

Under the zero-anchor convention of Appendix C.2, this identity also holds when the trade probability is zero, since a factor whose arm-1 event is the trade event itself then vanishes. Fix $(p_i,q_i)\in\operatorname{supp}(\gamma_k)$ and $\theta,\theta'\in\mathcal{F}_{k-1}$, let r index $\mathcal{D}([0,p_i])$, and let I'_r range over the reconstruction cover $\mathcal{D}([\max I_r,1])\cup\mathcal{D}([0,q_i])$ of the r-th anchor. By Observation 3, this cover has at most $2(J+1)$ intervals, all in $\mathcal{S}(p_i,q_i)\cup\mathcal{B}(p_i,q_i)$; the maxima below range over this set. Write x_ρ and y_ρ for the ρ-th factors of the product in eq. (11) under θ and θ', all nonnegative, and interchange the candidates if necessary so that $\prod_\rho x_\rho \ge \prod_\rho y_\rho$. With $z_\rho \triangleq \min\{x_\rho,y_\rho\}\le y_\rho$,

$$\begin{aligned}\left|\textstyle\prod_\rho x_\rho - \prod_\rho y_\rho\right| &\le \textstyle\prod_\rho x_\rho - \prod_\rho z_\rho\\ \triangleright\ 0\le z_{\rho'}\le x_{\rho'} \quad &\le \textstyle\sum_\rho\left(\prod_{\rho'\ne\rho}x_{\rho'}\right)(x_\rho - z_\rho)\\ \triangleright\ x_\rho - z_\rho \le |x_\rho - y_\rho| \quad &\le \textstyle\sum_\rho\left(\prod_{\rho'\ne\rho}x_{\rho'}\right)\cdot|x_\rho - y_\rho|\,,\end{aligned} \tag{12}$$

so every coefficient involves only the candidate θ. Then every anchor satisfies

$$\begin{aligned}\left|\Pr_\theta[a_{I_r,0}] - \Pr_{\theta'}[a_{I_r,0}]\right| &= \left|\textstyle\prod_{I'_r}\frac{\Pr_\theta\left[a_{I'_r,1}\right]}{\Pr_\theta\left[a_{I'_r,0}\right]} - \prod_{I'_r}\frac{\Pr_{\theta'}\left[a_{I'_r,1}\right]}{\Pr_{\theta'}\left[a_{I'_r,0}\right]}\right| && \text{eq. (11)}\\ &\le \textstyle\sum_{I'_r}\overline{\Pr}_k[a_{I'_r,0}] && \text{eq. (12)}\\ &\quad\times\max_{\ell\in\{1,2\}}\left|\frac{\Pr_\theta\left[a_{I'_r,\ell}\right]}{\Pr_\theta\left[a_{I'_r,0}\right]} - \frac{\Pr_{\theta'}\left[a_{I'_r,\ell}\right]}{\Pr_{\theta'}\left[a_{I'_r,0}\right]}\right|\\ &\le \textstyle\sum_{I'_r}\phi_k\left((p_i,q_i),I'_r\right)\\ &\le 2(J+1)\max_{I'}\phi_k\left((p_i,q_i),I'\right).\end{aligned} \tag{13}$$

Arm 2 differs from the anchor by one further ratio, whose disagreement is bounded by ϕ_k at $\ell=2$:

$$\begin{aligned}\left|\Pr_\theta[a_{I_r,2}] - \Pr_{\theta'}[a_{I_r,2}]\right| &= \left|\Pr_\theta[a_{I_r,0}]\cdot\tfrac{\Pr_\theta[a_{I_r,2}]}{\Pr_\theta[a_{I_r,0}]} - \Pr_{\theta'}[a_{I_r,0}]\cdot\tfrac{\Pr_{\theta'}[a_{I_r,2}]}{\Pr_{\theta'}[a_{I_r,0}]}\right|\\ &\le \tfrac{\Pr_\theta[a_{I_r,2}]}{\Pr_\theta[a_{I_r,0}]}\cdot\left|\Pr_\theta[a_{I_r,0}] - \Pr_{\theta'}[a_{I_r,0}]\right|\\ &\quad + \Pr_{\theta'}[a_{I_r,0}]\cdot\left|\tfrac{\Pr_\theta[a_{I_r,2}]}{\Pr_\theta[a_{I_r,0}]} - \tfrac{\Pr_{\theta'}[a_{I_r,2}]}{\Pr_{\theta'}[a_{I_r,0}]}\right|\\ \triangleright\ a_{I_r,2}\subseteq a_{I_r,0} \quad &\le \left|\Pr_\theta[a_{I_r,0}] - \Pr_{\theta'}[a_{I_r,0}]\right| + \overline{\Pr}_k[a_{I_r,0}]\cdot\left|\tfrac{\Pr_\theta[a_{I_r,2}]}{\Pr_\theta[a_{I_r,0}]} - \tfrac{\Pr_{\theta'}[a_{I_r,2}]}{\Pr_{\theta'}[a_{I_r,0}]}\right|\\ &\le 2(J+1)\max_{I'}\phi_k\left((p_i,q_i),I'\right) + \max_{I'}\phi_k\left((p_i,q_i),I'\right) \quad \text{eq. (13)}\\ &= (2J+3)\max_{I'}\phi_k\left((p_i,q_i),I'\right).\end{aligned}$$

Splitting the GFT into its seller, buyer, and profit parts and applying the triangle inequality,

$$\begin{aligned}
&\left|\mathsf{GFT}_\theta(p_i,q_i)-\mathsf{GFT}_{\theta'}(p_i,q_i)\right| \\
&\le \left|\Pr_\theta[s\le p_i,\ b\ge q_i]\,U_{S,\theta}\left(p_i\right)-\Pr_{\theta'}[s\le p_i,\ b\ge q_i]\,U_{S,\theta'}\left(p_i\right)\right| &(14)\\
&\quad+\left|\Pr_\theta[s\le p_i,\ b\ge q_i]\,U_{B,\theta}\left(q_i\right)-\Pr_{\theta'}[s\le p_i,\ b\ge q_i]\,U_{B,\theta'}\left(q_i\right)\right| &(15)\\
&\quad+\left|\mathsf{PRO}_\theta(p_i,q_i)-\mathsf{PRO}_{\theta'}(p_i,q_i)\right|. &(16)
\end{aligned}$$

By eq. (10) the two seller-side weighted utilities are the corresponding sums over r, so

$$\begin{aligned}
\textit{eq. (14)} &= \left|\textstyle\sum_r |I_r|\cdot\left(\Pr_\theta[a_{I_r,2}]-\Pr_{\theta'}[a_{I_r,2}]\right)\right| \\
&\le \textstyle\sum_r |I_r|\cdot\left|\Pr_\theta[a_{I_r,2}]-\Pr_{\theta'}[a_{I_r,2}]\right| \\
\triangleright\ \text{the arm-2 bound}\quad &\le (2J+3)\cdot\textstyle\sum_r |I_r|\max_{I'}\phi_k\left((p_i,q_i),I'\right) \\
\triangleright\ \textstyle\sum_r|I_r|=p_i\le 1\quad &\le (2J+3)\max_{I'}\phi_k\left((p_i,q_i),I'\right),
\end{aligned}$$

and eq. (15) satisfies the same bound with the endpoints exchanged. The last interval of $\mathcal{D}([0,p_i])$ has $\max I_r = p_i$, so its anchor is the trade event $\{s\le p_i,\ b\ge q_i\}$; the expected-profit identity and eq. (13) then give

$$\begin{aligned}
\textit{eq. (16)} &= |q_i-p_i|\ \left|\Pr_\theta[a_{I_r,0}]-\Pr_{\theta'}[a_{I_r,0}]\right| \\
&\le 2(J+1)\max_{I'}\phi_k\left((p_i,q_i),I'\right),
\end{aligned}$$

so that

$$\left|\mathsf{GFT}_\theta(p_i,q_i)-\mathsf{GFT}_{\theta'}(p_i,q_i)\right|\le(6J+8)\max_{I'}\phi_k\left((p_i,q_i),I'\right).$$

Multiplying by $\gamma_k(p_i,q_i)$, summing over the at most two atoms of $\mathrm{supp}(\gamma_k)$, and applying the maximality of $(p,q),I$ in Algorithm 3, line 16,

$$\begin{aligned}
\overline{\mathsf{GFT}}_k(\gamma_k)-\underline{\mathsf{GFT}}_k(\gamma_k) &\le 2(6J+8)\cdot\gamma_k(p,q)\cdot\phi_k((p,q),I)\\
&\le 16(J+1)\cdot\gamma_k(p,q)\cdot\phi_k((p,q),I).
\end{aligned}$$

The same argument for expected profit gives

$$\overline{\mathsf{PRO}}_k(\gamma_k)-\underline{\mathsf{PRO}}_k(\gamma_k)\le 4(J+1)\cdot\gamma_k(p,q)\cdot\phi_k((p,q),I).$$

□

Corollary 2. *Let $\widehat{k}$ be the epoch at which the refinement loop terminates. On $\mathcal{E}$,*

$$\mathsf{GFT}^\star_{K,\theta^\star}-\mathbb{E}_{(p,q)\sim\gamma_{\widehat{k}}}\left[\mathsf{GFT}_{\theta^\star}(p,q)\right]\le\frac{1}{5K},$$

and

$$\mathbb{E}_{(p,q)\sim\gamma_{\widehat{k}}}\left[\mathsf{PRO}_{\theta^\star}(p,q)\right]\ge-\frac{1}{20K}.$$

Proof. An optimal γ for $\mathsf{GFT}^\star_{K,\theta^\star}$ is supported on at most two price pairs. Since $\theta^\star\in\mathcal{F}_{\widehat{k}-1}$, pairing it with this distribution gives a feasible point of the joint OFU program. The optimality of the OFU choice therefore gives

$$\begin{aligned}
\mathsf{GFT}^\star_{K,\theta^\star}-\mathbb{E}_{(p,q)\sim\gamma_{\widehat{k}}}\left[\mathsf{GFT}_{\theta^\star}(p,q)\right] &\le \mathbb{E}_{(p,q)\sim\gamma_{\widehat{k}}}\left[\mathsf{GFT}_{\theta_{\widehat{k}}}(p,q)-\mathsf{GFT}_{\theta^\star}(p,q)\right]\\
\triangleright\ \theta_{\widehat{k}},\theta^\star\in\mathcal{F}_{\widehat{k}-1}\quad &\le \overline{\mathsf{GFT}}_{\widehat{k}}(\gamma_{\widehat{k}})-\underline{\mathsf{GFT}}_{\widehat{k}}(\gamma_{\widehat{k}})\\
\triangleright\ \text{Algorithm 3, line 17}\quad &\le 16(J+1)\eta=\tfrac{1}{5K}. \qquad \text{Lemma 8}
\end{aligned}$$

The same two memberships, together with the feasibility of $(\theta_{\widehat{k}},\gamma_{\widehat{k}})$, give

$$\begin{aligned}
\mathbb{E}_{(p,q)\sim\gamma_{\widehat{k}}}\left[\mathsf{PRO}_{\theta^\star}(p,q)\right] &\ge \underline{\mathsf{PRO}}_{\widehat{k}}(\gamma_{\widehat{k}})\\
&\ge \mathbb{E}_{(p,q)\sim\gamma_{\widehat{k}}}\left[\mathsf{PRO}_{\theta_{\widehat{k}}}(p,q)\right]-\left(\overline{\mathsf{PRO}}_{\widehat{k}}(\gamma_{\widehat{k}})-\underline{\mathsf{PRO}}_{\widehat{k}}(\gamma_{\widehat{k}})\right)\\
&\ge -\left(\overline{\mathsf{PRO}}_{\widehat{k}}(\gamma_{\widehat{k}})-\underline{\mathsf{PRO}}_{\widehat{k}}(\gamma_{\widehat{k}})\right)\\
\triangleright\ \text{Algorithm 3, line 17}\quad &\ge -4(J+1)\eta=-\tfrac{1}{20K}. \qquad \text{Lemma 8}
\end{aligned}$$

□

For the refinement rounds, the next lemma bounds the uncertainty charged to each typed node.

Lemma 9. *On $\mathcal{E}$, every typed node I satisfies*

$$\textstyle\sum_{\text{calls at } I} m \cdot \gamma_k(p,q) \cdot \phi_k\left((p,q), I\right) \le \frac{7945 \log(17T^2)}{\eta}.$$

Proof. Suppose the selected node is a seller node $I = [u, v] \in \mathcal{S}(p, q)$; buyer nodes are symmetric. By independence, $\Pr_\theta[a_{I,0}] = \Pr_\theta[s \le v] \cdot \Pr_\theta[b \ge q]$, and the boundary probes at $(v, 0)$ and $(1, q)$ constrain the two factors:

$$\mathrm{KL}\big(\widehat{\Pr}[s \le v,\ b \ge 0]\,, \Pr_\theta[s \le v]\big) \le \tfrac{\eta}{128}, \qquad \mathrm{KL}\big(\widehat{\Pr}[s \le 1,\ b \ge q]\,, \Pr_\theta[b \ge q]\big) \le \tfrac{\eta}{128}.$$

As $\overline{\Pr}_k[a_{I,0}] > \eta$ by Algorithm 3, line 17, and each factor is at most one, both marginals of the maximizing candidate exceed η; by the two bounds above, each empirical marginal then exceeds $\frac{7\eta}{8}$. Since $\mathrm{KL}(x, y) = x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y} \ge x\left(\frac{y}{x} - 1 - \log \frac{y}{x}\right)$, every survivor's seller marginal satisfies

$$\begin{aligned}
\tfrac{\Pr_\theta[s\le v]}{\widehat{\Pr}[s\le v,\ b\ge 0]} - 1 - \log \tfrac{\Pr_\theta[s\le v]}{\widehat{\Pr}[s\le v,\ b\ge 0]} &\le \tfrac{\mathrm{KL}\left(\widehat{\Pr}[s\le v,\ b\ge 0], \Pr_\theta[s\le v]\right)}{\widehat{\Pr}[s\le v,\ b\ge 0]} \\
&\le \tfrac{\eta}{128\,\widehat{\Pr}[s\le v,\ b\ge 0]} \\
\triangleright\ \widehat{\Pr}[s \le v,\ b \ge 0] > \tfrac{7\eta}{8} \quad &< \tfrac{1}{112} \\
&< 2^{-1/4} - 1 + \tfrac{\log 2}{4} \\
&< 2^{1/4} - 1 - \tfrac{\log 2}{4}.
\end{aligned}$$

The map $x \mapsto x - 1 - \log x$ decreases on $(0, 1]$ and increases on $[1, \infty)$, so that ratio lies strictly between $2^{-1/4}$ and $2^{1/4}$; the buyer marginal is the same with the two probes exchanged. In particular $\underline{\Pr}_k[a_{I,0}] > 0$ and

$$\begin{aligned}
\tfrac{\overline{\Pr}_k[a_{I,0}]}{\underline{\Pr}_k[a_{I,0}]} &= \sup_{\theta,\theta' \in \mathcal{F}_{k-1}} \tfrac{\Pr_\theta[s\le v]\cdot\Pr_\theta[b\ge q]}{\Pr_{\theta'}[s\le v]\cdot\Pr_{\theta'}[b\ge q]} \\
&\le \sup_{\theta,\theta' \in \mathcal{F}_{k-1}} \tfrac{\Pr_\theta[s\le v]}{\Pr_{\theta'}[s\le v]} \cdot \sup_{\theta,\theta' \in \mathcal{F}_{k-1}} \tfrac{\Pr_\theta[b\ge q]}{\Pr_{\theta'}[b\ge q]} \\
&< \left(\sqrt[4]{2} \cdot \sqrt[4]{2}\right) \cdot \left(\sqrt[4]{2} \cdot \sqrt[4]{2}\right) \\
&= 2.
\end{aligned}$$

It remains to bound the call length m in terms of $\phi_k\left((p, q), I\right)$. Fix $\ell \in \{1, 2\}$, $\theta, \theta' \in \mathcal{F}_{k-1}$ and $N_I \ge 1$. Rewriting both ratios in terms of the normalized arm probabilities,

$$\begin{aligned}
\tfrac{\Pr_\theta[a_{I,\ell}]}{\Pr_\theta[a_{I,0}]} - \tfrac{\Pr_{\theta'}[a_{I,\ell}]}{\Pr_{\theta'}[a_{I,0}]} &= \tfrac{\sum_{\ell'} \Pr_\theta[a_{I,\ell'}]}{\Pr_\theta[a_{I,0}]} \Big[\Big(\tfrac{\Pr_\theta[a_{I,\ell}]}{\sum_{\ell'} \Pr_\theta[a_{I,\ell'}]} - \tfrac{\Pr_{\theta'}[a_{I,\ell}]}{\sum_{\ell'} \Pr_{\theta'}[a_{I,\ell'}]} \Big) \\
&\quad + \tfrac{\Pr_{\theta'}[a_{I,\ell}]}{\Pr_{\theta'}[a_{I,0}]} \cdot \Big(\tfrac{\Pr_{\theta'}[a_{I,0}]}{\sum_{\ell'} \Pr_{\theta'}[a_{I,\ell'}]} - \tfrac{\Pr_\theta[a_{I,0}]}{\sum_{\ell'} \Pr_\theta[a_{I,\ell'}]} \Big) \Big] \\
\triangleright\ a_{I,1}, a_{I,2} \subseteq a_{I,0} \quad &\le 3\Big(\Big| \tfrac{\Pr_\theta[a_{I,\ell}]}{\sum_{\ell'} \Pr_\theta[a_{I,\ell'}]} - \tfrac{\Pr_{\theta'}[a_{I,\ell}]}{\sum_{\ell'} \Pr_{\theta'}[a_{I,\ell'}]} \Big| + \Big| \tfrac{\Pr_\theta[a_{I,0}]}{\sum_{\ell'} \Pr_\theta[a_{I,\ell'}]} - \tfrac{\Pr_{\theta'}[a_{I,0}]}{\sum_{\ell'} \Pr_{\theta'}[a_{I,\ell'}]} \Big| \Big) \\
\triangleright\ \text{Algorithm 5, line 17} \quad &\le 3 \cdot 2 \cdot 2\sqrt{\tfrac{\log(1224KT(N_I+1)^2)}{2N_I}} \\
&= 12\sqrt{\tfrac{\log(1224KT(N_I+1)^2)}{2N_I}}.
\end{aligned}$$

Taking the maximum over ℓ and the supremum over θ, θ' in the definition of ϕ_k and clipping at one,

$$\begin{aligned}
\phi_k\left((p,q), I\right) &= \overline{\Pr}_k[a_{I,0}] \\
&\quad \times \max_{\ell \in \{1,2\}} \sup_{\theta,\theta' \in \mathcal{F}_{k-1}} \left(\tfrac{\Pr_\theta[a_{I,\ell}]}{\Pr_\theta[a_{I,0}]} - \tfrac{\Pr_{\theta'}[a_{I,\ell}]}{\Pr_{\theta'}[a_{I,0}]} \right) \\
&\le \overline{\Pr}_k[a_{I,0}] \cdot 12\sqrt{\tfrac{\log(1224KT(N_I+1)^2)}{2N_I}} \\
\triangleright\ \phi_k\left((p,q), I\right) \le \overline{\Pr}_k[a_{I,0}] \text{ and } N_I \le T \quad &\le \overline{\Pr}_k[a_{I,0}] \cdot \min\left\{1,\ 12\sqrt{\tfrac{3\log(17T^2)}{2N_I}}\right\}. \qquad (17)
\end{aligned}$$

For $N_I \geq 1$, the selection rule also bounds the node's count:

$$\begin{aligned} N_I &\leq 216\log(17T^2)\cdot\left(\tfrac{\overline{\Pr}_k[a_{I,0}]}{\phi_k((p,q),I)}\right)^2 && \text{eq. (17)} \\ \triangleright \text{ Algorithm 3, line 17} \quad &< 216\log(17T^2)\cdot\left(\tfrac{\gamma_k(p,q)\cdot\overline{\Pr}_k[a_{I,0}]}{\eta}\right)^2 \\ &\leq \tfrac{216\log(17T^2)}{\eta^2}. \end{aligned}$$

The rule of Algorithm 5, line 3, gives

$$\begin{aligned} m\cdot\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right) &\leq \left(\tfrac{4\max\{\eta n,N_I\}}{\gamma_k(p,q)\cdot\underline{\Pr}_k[a_{I,0}]}+1\right)\cdot\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right) \\ &\leq \tfrac{4\max\{\eta n,N_I\}\cdot\phi_k((p,q),I)}{\underline{\Pr}_k[a_{I,0}]}+\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right) \\ \triangleright\ \overline{\Pr}_k[a_{I,0}]<2\cdot\underline{\Pr}_k[a_{I,0}] \quad &\leq 8\max\{\eta n,N_I\}\cdot\tfrac{\phi_k((p,q),I)}{\overline{\Pr}_k[a_{I,0}]}+\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right) \\ &\leq 8\max\{\eta n,N_I\}\cdot\tfrac{\phi_k((p,q),I)}{\overline{\Pr}_k[a_{I,0}]}+1. \end{aligned} \tag{18}$$

By Lemma 7, the first call at a fixed typed node I raises N_I to at least ηn. Summing over all calls at I,

$$\begin{aligned} \textstyle\sum_{\text{calls at } I} m\cdot\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right) &\leq 8\cdot\textstyle\sum_{\text{calls at } I}\max\{\eta n,N_I\}\cdot\tfrac{\phi_k((p,q),I)}{\overline{\Pr}_k[a_{I,0}]}+\tfrac{\log(17T^2)}{\eta} \\ &\qquad\qquad \text{eq. (18)} \\ &\leq 8\eta n+96\sqrt{\tfrac{3}{2}\log(17T^2)}\cdot\textstyle\sum_{\text{later calls at } I}\sqrt{N_I}+\tfrac{\log(17T^2)}{\eta} \\ &\qquad\qquad \text{eq. (17)} \\ \triangleright\ N_I \text{ at least doubles} \quad &\leq 8\eta n+384\sqrt{\tfrac{3}{2}\log(17T^2)}\cdot\max_{\text{calls at } I}\sqrt{N_I}+\tfrac{\log(17T^2)}{\eta} \\ \triangleright\ N_I<\tfrac{216\log(17T^2)}{\eta^2} \quad &\leq 8\eta n+\tfrac{6913\log(17T^2)}{\eta} \\ \triangleright\ \eta n\leq 129\log(17T^2) \quad &\leq \tfrac{7945\log(17T^2)}{\eta}. \end{aligned}$$

□

Corollary 3. *On $\mathcal{E}$, the unstopped suffix completes its boundary exploration and refinement rounds by round $\frac{T+\tau}{2}$.*

Proof. Every selected call has $\eta<\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right)$, so

$$\begin{aligned} (2K+1)n+\textstyle\sum_{\text{all calls}} m &\leq (2K+1)n \\ &\quad+\tfrac{1}{\eta}\cdot\textstyle\sum_{\text{all calls}} m\cdot\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right) \\ \triangleright\ \text{fewer than } 4K \text{ typed nodes} \quad &< (2K+1)n+\tfrac{4K\cdot 7945\log(17T^2)}{\eta^2} && \text{Lemma 9} \\ \triangleright\ \eta^{-1}=80K(J+1) \text{ and } J+1\leq 2\log(17T^2) \quad &< 10^9K^3\left[\log(17T^2)\right]^3 \\ \triangleright\ \text{definition of } K \quad &\leq \tfrac{T-\tau}{2}. \end{aligned}$$

Here each call is counted at the length m fixed at its start, and Lemma 9 applies to all calls, as its proof uses item (ii) of Lemma 7 only for earlier calls, which are completed. Since the suffix starts at round $\tau+1$, the boundary exploration and refinement rounds end by round $\frac{T+\tau}{2}$, so no call is cut off and the horizon rule is never triggered. □

Lemma 10. *Algorithm 2 satisfies $\Pi_t\geq 0$ for every $0\leq t\leq T$ on every sample path. On $\mathcal{E}$, except on a further event of probability at most $\frac{1}{17T}$, $\min_{\tau\leq t\leq T}\Pi_t>1$, so the two suffixes coincide.*

Proof. A suffix round changes the cumulative profit by $(q_t-p_t)\cdot Z_t\geq -1$. If $\Pi_{t-1}\geq 1$, then $\Pi_t\geq\Pi_{t-1}-1\geq 0$; if $0\leq\Pi_{t-1}<1$, the GBB rule posts $(0,1)$ and $\Pi_t=\Pi_{t-1}\geq 0$. The prefix guarantee $\Pi_t\geq 0$ for $t\leq\tau$ provides the induction base, so GBB holds on every sample path.

It remains to show that on $\mathcal{E}$ the GBB rule is never triggered. The largest decrease of the cumulative profit from Π_τ decomposes into boundary exploration, refinement, exploitation, and martingale deviation:

$$\max_{\tau<t\le T}\big(\Pi_\tau-\Pi_t\big) \le \sum_{t\,\in\,\text{boundary exploration rounds}} \max\{0,-\mathbb{E}\left[\mathsf{PRO}_{\theta^\star}(p_t,q_t)\mid\mathcal{H}_{t-1}\right]\} \tag{19}$$

$$+\sum_{t\,\in\,\text{refinement rounds}} \max\{0,-\mathbb{E}\left[\mathsf{PRO}_{\theta^\star}(p_t,q_t)\mid\mathcal{H}_{t-1}\right]\} \tag{20}$$

$$+\sum_{t\,\in\,\text{exploitation rounds}} \max\{0,-\mathbb{E}\left[\mathsf{PRO}_{\theta^\star}(p_t,q_t)\mid\mathcal{H}_{t-1}\right]\} \tag{21}$$

$$+\max_{\tau<t\le T}\sum_{r=\tau+1}^{t}\big[\mathbb{E}\left[(q_r-p_r)\cdot Z_r\mid\mathcal{H}_{r-1}\right]-(q_r-p_r)\cdot Z_r\big]. \tag{22}$$

Each of the $(2K+1)n$ boundary rounds posts a grid pair, so the expected-profit identity gives

$$\begin{aligned}\textit{eq. (19)} &\le \sum_{t\,\in\,\text{boundary exploration rounds}} |q_t-p_t|\cdot\Pr_{\theta^\star}[s\le p_t,\ b\ge q_t]\\ &\le (2K+1)n.\end{aligned}$$

In a refinement round at a typed node I, Corollary 1 bounds the profit change from (p,q) to the posted arm (p',q') by $2(1+M)\,|I|$. The deficit of such a round splits into an action perturbation, a model mismatch, and the optimistic profit:

$$\begin{aligned}\textit{eq. (20)} &\le \sum_{\text{all calls}} m\cdot\max\big\{0,\ \mathbb{E}_{(p,q)\sim\gamma_k}\left[\mathsf{PRO}_{\theta^\star}(p,q)\right]\\ &\qquad -\mathbb{E}\left[\mathsf{PRO}_{\theta^\star}(p',q')\mid\mathcal{H}_{t-1}\right]\\ &\qquad +\mathbb{E}_{(p,q)\sim\gamma_k}\left[\mathsf{PRO}_{\theta_k}(p,q)\right]-\mathbb{E}_{(p,q)\sim\gamma_k}\left[\mathsf{PRO}_{\theta^\star}(p,q)\right]\\ &\qquad -\mathbb{E}_{(p,q)\sim\gamma_k}\left[\mathsf{PRO}_{\theta_k}(p,q)\right]\big\}\\ \triangleright\ \text{feasibility of } (\theta_k,\gamma_k)\quad &\le \sum_{\text{all calls}} m\cdot\big(\overline{\mathsf{PRO}}_k(\gamma_k)-\underline{\mathsf{PRO}}_k(\gamma_k)+2(1+M)\,|I|\big)\\ &\le 4(J+1)\cdot\sum_{\text{all calls}} m\cdot\gamma_k(p,q)\cdot\phi_k((p,q),I) \qquad \text{Lemma 8}\\ &\quad +2(1+M)\cdot\sum_{\text{typed } I}|I|\cdot\sum_{\text{calls at } I} m\\ \triangleright\ \eta<\gamma_k(p,q)\cdot\phi_k\left((p,q),I\right)\quad &\le \tfrac{127120K(J+1)\log(17T^2)}{\eta}+\tfrac{31780(1+M)(J+1)\log(17T^2)}{\eta^2}. \qquad \text{Lemma 9}\end{aligned}$$

After the refinement loop terminates at Algorithm 3, line 17, every exploitation round draws from the resulting $\gamma_{\widehat{k}}$, and there are at most $T-\tau$ of them, so

$$\begin{aligned}\textit{eq. (21)} &\le (T-\tau)\cdot\max\left\{0,\ -\mathbb{E}_{(p,q)\sim\gamma_{\widehat{k}}}\left[\mathsf{PRO}_{\theta^\star}(p,q)\right]\right\}\\ &\le \tfrac{T-\tau}{20K}. \qquad \text{Corollary 2}\end{aligned}$$

The martingale differences in eq. (22) have range at most 2, so the maximal Azuma–Hoeffding inequality gives, outside an event of probability at most $\frac{1}{17T}$,

$$\textit{eq. (22)} \le 2\sqrt{2(T-\tau)\log(17T^2)}.$$

By the maximality of K among the admissible dyadic resolutions,

$$\tfrac{T-\tau}{2} \in \left[10^{10}K^3\big[\log(17T^2)\big]^4,\ 8\cdot10^{10}K^3\big[\log(17T^2)\big]^4\right).$$

Adding the four bounds and substituting n, η and J gives

$$\begin{aligned}\max_{\tau<t\le T}\big(\Pi_\tau-\Pi_t\big) &< 499(1+M)(T-\tau)^{2/3}\big[\log(17T^2)\big]^3\\ &< 500(1+M)T^{2/3}\big[\log(17T^2)\big]^3\\ \triangleright\ \text{definition of } \beta\quad &\le \beta-1.\end{aligned}$$

Combining this with $\Pi_\tau\ge\beta$ gives $\min_{\tau\le t\le T}\Pi_t>1$, so the GBB rule is never triggered; with Corollary 3, the two suffixes coincide. □

Theorem 2. *For GFT maximization in repeated bilateral trade with one-bit feedback, where the valuations are i.i.d. draws from a product distribution with bounded density, Algorithm 2 satisfies GBB and achieves regret $\widetilde{\mathcal{O}}(T^{2/3})$ against the best feasible distribution over price pairs.*

Proof of Theorem 2. Splitting the horizon at the stopping time τ of the first stage, the regret decomposes into a prefix term and a suffix term:

$$\mathsf{REG}_{\mathrm{G}} = \mathbb{E}\left[\tau \cdot \mathsf{GFT}^{\star}_{\theta^\star} - \textstyle\sum_{t\in[\tau]} \mathsf{GFT}_{\theta^\star}(p_t, q_t)\right] \tag{23}$$

$$+ \mathbb{E}\left[(T-\tau) \cdot \mathsf{GFT}^{\star}_{\theta^\star} - \textstyle\sum_{t=\tau+1}^{T} \mathsf{GFT}_{\theta^\star}(p_t, q_t)\right]. \tag{24}$$

For the prefix term, the PROFIT-MAX regret guarantee holds outside an event of probability at most $\frac{2}{17T}$. Since the per-round regret is at most 2, this event contributes at most $2T \cdot \frac{2}{17T} < 1$, so

$$\textit{eq. (23)} \le \widetilde{\mathcal{O}}\big(\beta + \tfrac{T}{T^{1/3}} + \sqrt{T^{1/3} T \log(17T^2)}\big) + 1 \quad \text{Lunghi et al., 2026b, Lemma 7.2}$$

$$\triangleright \text{ definition of } \beta \quad \le \widetilde{\mathcal{O}}(T^{2/3}).$$

For the suffix term, since τ is a stopping time with respect to the history, the event $\{\tau < t\}$ is determined by $\mathcal{H}_{t-1}$, and the tower property gives

$$\textit{eq. (24)} = \mathbb{E}\left[(T-\tau) \cdot \mathsf{GFT}^{\star}_{\theta^\star} - \textstyle\sum_{t=\tau+1}^{T} \mathbb{E}\left[\mathsf{GFT}_{\theta^\star}(p_t, q_t) \mid \mathcal{H}_{t-1}\right]\right]. \tag{25}$$

The failure events of Lemmas 7 and 10 arise only when $K \ge 8$ and have total probability at most $\frac{8}{17T} + \frac{1}{17T} = \frac{9}{17T}$, so they contribute at most $2T \cdot \frac{9}{17T} < 2$ to eq. (25). It remains to bound the quantity inside the expectation in eq. (25) outside these events, and we distinguish three cases.

If $\tau = T$, this quantity is zero. If $\tau < T$ and $K = 0$, the suffix posts the safe action $(0, 1)$, whose GFT is zero under the bounded-density model, so this quantity equals $(T-\tau) \cdot \mathsf{GFT}^{\star}_{\theta^\star}$, and

$$(T-\tau) \cdot \mathsf{GFT}^{\star}_{\theta^\star} \le T - \tau$$

$$\triangleright \text{ the defining inequality for } K \text{ fails at } 2^3 \quad < 2 \cdot 10^{10} \cdot 8^3 \big[\log(17T^2)\big]^4 = \mathcal{O}\big(\big[\log(17T^2)\big]^4\big).$$

Finally, let $K \ge 8$. Outside the failure events, the two suffixes coincide by Lemma 10. The suffix rounds split into boundary exploration, recursive refinement, and terminal exploitation, so the quantity inside the expectation decomposes into a discretization term and one term per suffix phase:

$$(T-\tau) \cdot \mathsf{GFT}^{\star}_{\theta^\star} - \textstyle\sum_{t=\tau+1}^{T} \mathbb{E}\left[\mathsf{GFT}_{\theta^\star}(p_t, q_t) \mid \mathcal{H}_{t-1}\right] = (T-\tau) \cdot \big(\mathsf{GFT}^{\star}_{\theta^\star} - \mathsf{GFT}^{\star}_{K,\theta^\star}\big) \tag{26}$$

$$+ \textstyle\sum_{t \in \text{ boundary exploration rounds}} \big[\mathsf{GFT}^{\star}_{K,\theta^\star} - \mathbb{E}\left[\mathsf{GFT}_{\theta^\star}(p_t, q_t) \mid \mathcal{H}_{t-1}\right]\big] \tag{27}$$

$$+ \textstyle\sum_{t \in \text{ refinement rounds}} \big[\mathsf{GFT}^{\star}_{K,\theta^\star} - \mathbb{E}\left[\mathsf{GFT}_{\theta^\star}(p_t, q_t) \mid \mathcal{H}_{t-1}\right]\big] \tag{28}$$

$$+ \textstyle\sum_{t \in \text{ exploitation rounds}} \big[\mathsf{GFT}^{\star}_{K,\theta^\star} - \mathbb{E}\left[\mathsf{GFT}_{\theta^\star}(p_t, q_t) \mid \mathcal{H}_{t-1}\right]\big]. \tag{29}$$

We bound the four terms in order.

$$\textit{eq. (26)} \le \mathcal{O}\big(\tfrac{(T-\tau)\log(17T^2)}{K}\big). \quad \text{Lunghi et al., 2026b, Lemma 5.2}$$

$$\triangleright \text{ per-round regret} \le 2 \quad \textit{eq. (27)} \le 2(2K+1)n.$$

$$\textit{eq. (28)} \le 16(J+1) \cdot \textstyle\sum_{\text{all calls}} m \cdot \gamma_k(p,q) \cdot \phi_k((p,q), I) \quad \text{Lemma 8}$$

$$+ 2M \cdot \textstyle\sum_{\text{typed } I} |I| \cdot \sum_{\text{calls at } I} m \quad \text{Corollary 1}$$

$$\triangleright\ \eta < \gamma_k(p,q) \cdot \phi_k\left((p,q), I\right) \quad \le \tfrac{508480K(J+1)\log(17T^2)}{\eta} + \tfrac{31780M(J+1)\log(17T^2)}{\eta^2}. \quad \text{Lemma 9}$$

$$\textit{eq. (29)} \le \tfrac{T-\tau}{5K}. \quad \text{Corollary 2}$$

Adding the four bounds and substituting n and η gives

$$(T-\tau) \cdot \mathsf{GFT}^{\star}_{\theta^\star} - \textstyle\sum_{t=\tau+1}^{T} \mathbb{E}\left[\mathsf{GFT}_{\theta^\star}(p_t, q_t) \mid \mathcal{H}_{t-1}\right] \le \mathcal{O}\big(\tfrac{(T-\tau)\log(17T^2)}{K}\big)$$

$$+ \mathcal{O}\big(K^2\big[\log(17T^2)\big]^4\big)$$

$$\triangleright \text{ maximality of } K \quad = \mathcal{O}\big((T-\tau)^{2/3}\big[\log(17T^2)\big]^3\big).$$

In all three cases, the quantity inside the expectation in eq. (25) is therefore $\widetilde{\mathcal{O}}(T^{2/3})$ outside the failure events, and together with their contribution this bounds the suffix term eq. (24) by $\widetilde{\mathcal{O}}(T^{2/3})$. Adding the bound on the prefix term gives

$$\mathsf{REG}_{\mathrm{G}} \le \widetilde{\mathcal{O}}(T^{2/3}).$$

Pathwise GBB follows from the prefix guarantee when $\tau = T$, from the terminal safe action when $K = 0$, and from Lemma 10 otherwise. This finishes the proof of Theorem 2. □

D THE DYADIC PATH CONSTRUCTION

An interval $I = [u, v]$ is dyadic of depth j when $|I| = 2^{-j}$ and $2^j u$ is an integer; the dyadic intervals of depth at most J are the nodes of the binary tree that halves $[0, 1]$ down to width 2^{-J}. Algorithm 6 returns the coarsest partitions of $[0, x]$ and $[x, 1]$ into such nodes.

Since $K = 2^J$, every grid price is an endpoint of a depth-J node. For such a price v, the coarsest dyadic partition $\mathcal{D}([0, v])$ is the binary expansion of v: it has one node of width 2^{-j} for each position j at which the binary expansion of v has a one, from the widest at 0 to the narrowest at v. Symmetrically, $\mathcal{D}([u, 1])$ is the binary expansion of $1 - u$, laid out from u to 1 with the narrowest node first. PATH (Algorithm 7) combines the two. Writing $\mathcal{D}([0, p]) = \{I_1, \ldots, I_m\}$ from left to right, the seller path $\mathcal{S}(p, q)$ consists of these nodes together with the partitions $\mathcal{D}([\max I_r, 1])$, $r = 1, \ldots, m$. The right endpoints $\max I_r$ are the successive truncations of the binary expansion of p, the last being p itself. The buyer path $\mathcal{B}(p, q)$ is the mirror image, built from $\mathcal{D}([q, 1])$ and the partitions $\mathcal{D}([0, \min I])$ for $I \in \mathcal{D}([q, 1])$. At $p = 0$ or $q = 1$, the corresponding path is $\mathcal{D}([0, 1])$, as specified in Algorithm 7.

Algorithm 6 $\mathcal{D}$

function $\mathcal{D}([u, v])$
  Input: $[u, v] \subseteq [0, 1]$ with $u = 0$ or $v = 1$, and $2^J u, 2^J v$ integers.
  if $u = v$ **then**
    return $\emptyset$
  if $u = 0$ **then**
    $j \leftarrow \min\left\{j' \geq 0 : 2^{j'} v \in \mathbb{N}\right\}$
    return $\mathcal{D}([0, v - 2^{-j}]) \cup \left\{[v - 2^{-j}, v]\right\}$
  else
    $j \leftarrow \min\left\{j' \geq 0 : 2^{j'} u \in \mathbb{N}\right\}$
    return $\left\{[u, u + 2^{-j}]\right\} \cup \mathcal{D}([u + 2^{-j}, 1])$

Algorithm 7 PATH

function PATH(p, q)
  Input: $(p, q) \in \mathcal{G}_K$.
  if $p = 0$ **then**
    $\mathcal{S}(p, q) \leftarrow \mathcal{D}([0, 1])$
  else
    $\mathcal{S}(p, q) \leftarrow \emptyset$
    for all $[u, v] \in \mathcal{D}([0, p])$ **do**
      $\mathcal{S}(p, q) \leftarrow \mathcal{S}(p, q) \cup \{[u, v]\} \cup \mathcal{D}([v, 1])$
  if $q = 1$ **then**
    $\mathcal{B}(p, q) \leftarrow \mathcal{D}([0, 1])$
  else
    $\mathcal{B}(p, q) \leftarrow \emptyset$
    for all $[u, v] \in \mathcal{D}([q, 1])$ **do**
      $\mathcal{B}(p, q) \leftarrow \mathcal{B}(p, q) \cup \{[u, v]\} \cup \mathcal{D}([0, u])$
  return $(\mathcal{S}(p, q), \mathcal{B}(p, q))$

In Figure 1, $p = \frac{3}{8} = (0.011)_2$ has the nonzero digits 2^{-2} and 2^{-3}, so $\mathcal{D}([0, \frac{3}{8}]) = \{I_1, I_2\}$ with $I_1 = [0, \frac{1}{4}]$ and $I_2 = [\frac{1}{4}, \frac{3}{8}]$. Their right endpoints are the truncations $\frac{1}{4}$ and $\frac{3}{8}$; since $1 - \frac{1}{4} = (0.11)_2$ and $1 - \frac{3}{8} = (0.101)_2$, the partitions on the right are $\mathcal{D}([\frac{1}{4}, 1]) = \left\{[\frac{1}{4}, \frac{1}{2}], [\frac{1}{2}, 1]\right\}$ and $\mathcal{D}([\frac{3}{8}, 1]) = \left\{[\frac{3}{8}, \frac{1}{2}], [\frac{1}{2}, 1]\right\}$. The seller path therefore has five distinct nodes, with $[\frac{1}{2}, 1]$ shared by the two additional partitions. Consecutive nodes of a partition share a price pair: on the seller side, the anchor of one is arm 1 of the next (Observation 2).

The partition $\mathcal{D}([0, p])$ and the partitions $\mathcal{D}([\max I_r, 1])$ play different roles. The former carries the recursion: since consecutive nodes share a price pair, the identity from Lemma 2 telescopes along $I_1, \ldots, I_m$ from $U_S(0) = 0$ to the weighted utility at p (eq. (10)). Each step is weighted by the trade probability of its anchor, but the algorithm observes a node only through the relative trade frequencies of its arms; the latter partitions supply these weights. Under independence, the ratio of the trade probabilities of arms 1 and 0 depends on the seller's marginal alone. These ratios telescope

along $\mathcal{D}([\max I_r, 1])$ from $\Pr[s \le 1] = 1$ to $\Pr[s \le \max I_r]$, which, with the buyer's partition of $[0, q]$, gives the anchor probability of I_r (eq. (11)). These additional partitions are also dyadic, so their probes change one price by less than $2\,|I|$ at a node I and their statistics can be shared. With the buyer side mirrored, the GFT and the profit of (p, q) are determined by the conditional trade probabilities of the arms on its two paths.

E Finite-Dimensional Implementation of Global-OFU

The OFU problem in Algorithm 3, line 13, depends on θ only through the GFT, the profit and the arm probabilities at grid pairs, and we pose the same problem over the following values. For $K \ge 8$, write $C_i = [(i-1)/K, i/K]$, $i = 1, \dots, K$, and retain the $4K$ cell values

$$\left(\mu_{S,\theta}(C_i),\ \int_{C_i} s\,\mathrm{d}\mu_{S,\theta}(s),\ \mu_{B,\theta}(C_i),\ \int_{C_i} b\,\mathrm{d}\mu_{B,\theta}(b)\right)_{i=1}^{K}.$$

We call the two problems equivalent when they have the same optimal value and (θ, γ) is optimal for the original exactly when the cell values of θ, together with γ, are optimal for the reduced one.

Lemma 11. *For every epoch k with $\mathcal{F}_{k-1} \neq \emptyset$, the OFU problem in Algorithm 3, line 13, is equivalent to a problem over the $4K$ cell values whose feasible set is compact.*

Proof. For each C_i, the seller coordinates satisfy $0 \le \mu_{S,\theta}(C_i) \le M/K$, $\sum_{i=1}^{K} \mu_{S,\theta}(C_i) = 1$, and

$$\frac{i-1}{K}\mu_{S,\theta}(C_i) + \frac{\mu_{S,\theta}(C_i)^2}{2M} \le \int_{C_i} s\,\mathrm{d}\mu_{S,\theta}(s) \le \frac{i}{K}\mu_{S,\theta}(C_i) - \frac{\mu_{S,\theta}(C_i)^2}{2M}. \tag{30}$$

Indeed,

$$\begin{aligned} \textstyle\int_{C_i}\left(s - \frac{i-1}{K}\right)\mathrm{d}\mu_{S,\theta}(s) &= \textstyle\int_0^{1/K} \mu_{S,\theta}\left(\left[\frac{i-1}{K} + x,\ \frac{i}{K}\right]\right)\mathrm{d}x \\ \triangleright \text{ density at most } M \quad &\ge \textstyle\int_0^{\mu_{S,\theta}(C_i)/M} \left(\mu_{S,\theta}(C_i) - Mx\right)\mathrm{d}x \\ &= \textstyle\frac{\mu_{S,\theta}(C_i)^2}{2M}. \end{aligned}$$

Replacing $s - \frac{i-1}{K}$ by $\frac{i}{K} - s$ gives the upper bound in eq. (30).

Conversely, prescribe $4K$ cell values satisfying these constraints for both marginals. If $\mu_{S,\theta}(C_i) = 0$, then $\int_{C_i} s\,\mathrm{d}\mu_{S,\theta}(s) = 0$; put no mass in C_i. Otherwise, construct $\mu_{S,\theta'}$ with density M on

$$\left[\frac{\int_{C_i} s\,\mathrm{d}\mu_{S,\theta}(s)}{\mu_{S,\theta}(C_i)} - \frac{\mu_{S,\theta}(C_i)}{2M},\ \frac{\int_{C_i} s\,\mathrm{d}\mu_{S,\theta}(s)}{\mu_{S,\theta}(C_i)} + \frac{\mu_{S,\theta}(C_i)}{2M}\right] \subseteq C_i, \qquad \text{eq. (30)}$$

and zero elsewhere in C_i. Its length and midpoint give

$$\begin{aligned} \left(\mu_{S,\theta'}(C_i),\ \textstyle\int_{C_i} s\,\mathrm{d}\mu_{S,\theta'}(s)\right) &= M\,\frac{\mu_{S,\theta}(C_i)}{M}\left(1,\ \frac{\int_{C_i} s\,\mathrm{d}\mu_{S,\theta}(s)}{\mu_{S,\theta}(C_i)}\right) \\ &= \left(\mu_{S,\theta}(C_i),\ \textstyle\int_{C_i} s\,\mathrm{d}\mu_{S,\theta}(s)\right). \end{aligned}$$

The same construction for the buyer gives $\theta' = \mu_{S,\theta'} \otimes \mu_{B,\theta'} \in \mathcal{F}$.

Every grid tail and every interval in Algorithm 7 is a union of cells, up to endpoints of zero mass. For $(p, q) \in \mathcal{G}_K$, independence gives

$$\begin{aligned} \Pr_\theta[s \le p,\ b \ge q] &= \left(\textstyle\sum_{i:\, C_i \subseteq [0,p]} \mu_{S,\theta}(C_i)\right)\left(\textstyle\sum_{i:\, C_i \subseteq [q,1]} \mu_{B,\theta}(C_i)\right) \\ &= \Pr_{\theta'}[s \le p,\ b \ge q]\,. \end{aligned}$$

Hence $\mathsf{PRO}_\theta(p, q) = \mathsf{PRO}_{\theta'}(p, q)$, and

$$\begin{aligned} \mathsf{GFT}_\theta(p, q) &= \Pr_\theta[s \le p] \textstyle\int_q^1 b\,\mathrm{d}\mu_{B,\theta}(b) - \Pr_\theta[b \ge q] \int_0^p s\,\mathrm{d}\mu_{S,\theta}(s) \\ \triangleright \text{ preserved cell masses and first moments} \quad &= \mathsf{GFT}_{\theta'}(p, q). \end{aligned}$$

Arms 0 and 1 use grid endpoints. For a seller-typed node $I = [u, v] \in \mathcal{S}(p, q)$,

$$\begin{aligned}\Pr_\theta[a_{I,2}] &= \tfrac{\Pr_\theta[b \ge q]}{|I|} \int_u^v \Pr_\theta[s \le x] \, \mathrm{d}x \\ &= \Pr_\theta[s \le u,\ b \ge q] \\ &\quad + \tfrac{\Pr_\theta[b \ge q]}{|I|} \cdot \textstyle\sum_{i:\, C_i \subseteq I} \left(v\mu_{S,\theta}(C_i) - \int_{C_i} s \, \mathrm{d}\mu_{S,\theta}(s) \right) \\ &= \Pr_{\theta'}[a_{I,2}] \, .\end{aligned}$$

For buyer-typed nodes, use the cell integral $\int_{C_i} (b-u) \, \mathrm{d}\mu_{B,\theta}(b)$, the endpoint term $\Pr_\theta[s \le p,\ b \ge v]$, and the opposite factor $\Pr_\theta[s \le p]$. Thus the candidate constraints, the envelopes and ϕ_k are preserved, so the objective, the profit constraint and membership in $\mathcal{F}_{k-1}$ depend on θ only through its cell values. With the converse construction above, the two problems are equivalent.

The feasible cell values form a compact set. Those of $\mathcal{F}$ are cut out of a bounded box by the non-strict constraints above, and each elimination adds a non-strict inequality in them: the test at Algorithm 3, line 9, defines a sublevel set of a lower semicontinuous function, and the test of Algorithm 5 is non-strict once cross-multiplied. Since γ ranges over a finite union of two-point simplices, the objective and the constraint are continuous, and the point mass at $(0, 1)$ is feasible, the extreme value theorem gives an attained maximum. □

E.1 Approximate Solution of the OFU Problem

We show that, for every epoch k with $\mathcal{F}_{k-1} \neq \emptyset$, the OFU problem in Algorithm 3, line 13, can be solved to within $1/K$ in GFT, with profit at least $-3/(4K)$, through convex programs over one marginal. The moment constraints in eq. (30) are convex, and each KL constraint in Algorithm 3, line 9, involves only one marginal. At every node where Algorithm 5 is called, both factors of the anchor probability are positive for every surviving candidate (see the proof of Lemma 9), so the opposite side's acceptance probability cancels from the ratio constraints of Algorithm 5, which become linear in the cell values of one marginal. Hence, in terms of cell values, $\mathcal{F}_{k-1}$ is the product of the set of its seller marginals and the set of its buyer marginals, both convex sets described by $\mathcal{O}(K)$ variables and $\mathcal{O}(K + T)$ constraints.

By independence, the GFT and the profit at a grid pair (p, q) depend on the seller marginal only through $\Pr_\theta[s \le p]$ and $\mathbb{E}_\theta[s \mathbb{1}\{s \le p\}]$. For γ supported on (p_1, q_1) and (p_2, q_2), once these four numbers and $\gamma(p_1, q_1)$ are fixed, the objective and the profit constraint are linear in the cell values of the buyer marginal. We enumerate the supports, the $(4K)^4$ cubes of side $1/(4K)$ covering $[0, 1]^4$, and the values of $\gamma(p_1, q_1)$ in $\{0, 1/(4K), \ldots, 1\}$. For each cube that contains the four numbers of a seller marginal of $\mathcal{F}_{k-1}$, we fix one such marginal as its representative and, for each value of $\gamma(p_1, q_1)$, maximize the GFT over the buyer marginals of $\mathcal{F}_{k-1}$ subject to profit at least $-3/(4K)$. The output (θ_k, γ_k) is the best solution found.

To bound its error, take an optimal (θ, γ), round $\gamma(p_1, q_1)$ down to a multiple of $1/(4K)$, and replace the seller marginal of θ by the representative of a cube containing its four numbers. The rounding changes the GFT and the profit by at most $1/(2K)$ each, and the replacement changes them by at most $1/(2K)$ and $1/(4K)$, respectively. Since $\mathcal{F}_{k-1}$ is a product, the modified candidate remains in $\mathcal{F}_{k-1}$, and its buyer marginal is feasible for the corresponding program. The output therefore has GFT at least the optimal value minus $1/K$ and profit at least $-3/(4K)$ under $\theta_k \in \mathcal{F}_{k-1}$. With $\mathcal{O}(K^4)$ supports, $\mathcal{O}(K^4)$ cubes and $\mathcal{O}(K)$ values of $\gamma(p_1, q_1)$, solving the OFU problem thus requires $\mathcal{O}(K^9)$ convex programs.

F Conclusion and Future Work

We studied fixed-price bilateral trade under one-bit feedback when the seller and buyer valuations are independent with bounded densities. For WBB mechanisms against the best fixed common price and for GBB mechanisms against the best feasible distribution over price pairs, we gave algorithms achieving $\widetilde{\mathcal{O}}(T^{2/3})$ regret, which improves the previous $\widetilde{\mathcal{O}}(T^{3/4})$ bounds and matches the $\Omega(T^{2/3})$ lower bounds up to logarithmic factors. Both algorithms rely on a recursion that uses independence to carry each side's conditional utility from one endpoint of a price interval to the other through trade probabilities alone. Chained over dyadic intervals, the recursion keeps exploration cheap in both settings. The WBB algorithm refines a dyadic tree of prices while probing only inside the interval being refined, so that a round costs regret of the order of that interval's width; the GBB algorithm applies the recursion along dyadic paths whose statistics are shared across price pairs, so that it learns the valuation law rather than the profit of each pair.

The recursion uses independence only through the product form of the trade probability, so we believe it may be useful beyond the settings studied here. A natural extension is to two-sided markets with several sellers and buyers (Babaioff et al., 2024; Lunghi et al., 2025; Feng et al., 2026), for instance when one price is posted to all sellers and another to all buyers and the accepting traders are matched maximally. At least k trades then occur exactly when at least k sellers and at least k buyers accept, so under independence the probability of this event still factorizes into a seller term and a buyer term, which suggests that a similar recursion may recover each side's utilities from the number of trades alone. Another direction concerns the bounded-density assumption: without it, the linear lower bound of Lunghi et al. (2026b) relies on correlated valuations, whereas for independent valuations Chen et al. (2026) still obtain $\widetilde{\mathcal{O}}(T^{2/3})$ regret for GBB mechanisms against the best fixed price, so it remains open whether the same rate holds against the best feasible distribution. Finally, in contextual settings (Gaucher et al., 2025; Feng et al., 2026), where the valuations depend on observed features, it would be interesting to combine the recursion with a model of how the conditional utilities depend on these features.