Contextual Search and Bilateral Trade with a Trace-Square-Root Potential

Romain Cosson* Federico Fusco† Anupam Gupta‡ Stefano Leonardi†
Renato Paes Leme§ Matteo Russo¶

Abstract

We present a unified potential method for contextual search and bilateral trade with one-bit feedback and obtain order-tight algorithms for both problems. Our framework starts from a Bayesian perspective in which the unknown target (for contextual search) and the unknown pair of valuations (for bilateral trade) are sampled from a known prior distribution. In that model, the decision maker bases its prediction on the posterior distribution, obtained by conditioning on feedback received in previous rounds. For Bayesian contextual search, we present an algorithm with a total loss bounded by $O(\operatorname{Tr}\sqrt{\Sigma})$, where Σ is the covariance matrix of the target's prior; and for contextual bilateral trade, our algorithm's regret is bounded by $O(\operatorname{Tr}\sqrt{\Sigma_S} + \operatorname{Tr}\sqrt{\Sigma_B})$, where Σ_S and Σ_B are the covariance matrices of the seller's and buyer's valuations.

We then use ideas from this Bayesian setting and adapt a "dual" of this trace-square-root potential to the classical worst-case setting, where the target and valuations are chosen adversarially from the unit Euclidean ball. For both contextual search and bilateral trade, we obtain tight $O(\sqrt{d})$-regret algorithms. These results match the existing lower bounds up to a constant factor and improve on the previous guarantees of $O(d \log d)$ for contextual search and of $O(d6^d)$ for bilateral trade.

*Computer Science Department, New York University.
†Sapienza University, Rome, Italy.
‡Computer Science Department, New York University, and Google Research.
§Google Research, NYC, United States.
¶EPFL, Lausanne, Switzerland.

1 Introduction

We study parameter search problems in settings where decision makers receive only limited feedback. For example, a pricing algorithm sees the features of the customer, quotes a price for a product described by a feature vector and learns only whether the customer bought it. An idealized dosing procedure receives patient features and learns only whether the dose was too low or too high. A broker quotes prices to both a buyer and a seller, and learns only whether they traded. In each case, there is a vector hidden from the learner, and at each step, the observed features provide a "context": together, the hidden vector and the context determine a target value that the learner tries to match. In this paper, we study two classical problems at the intersection of theoretical computer science and economics that are inspired by such situations: *contextual search* and *contextual bilateral trade*.

Contextual search. The basic problem in this setting is *contextual search*, where the target is an unknown vector $\boldsymbol{z} \in \mathbb{R}^d$. In round t, an adversary reveals the *context* vector $\boldsymbol{x}_t$ with $\|\boldsymbol{x}_t\|_2 \leq 1$. The two together define the target value $\boldsymbol{x}_t^\top \boldsymbol{z}$—this is unknown to the learner, since $\boldsymbol{z}$ is unknown. The learner does observe $\boldsymbol{x}_t$, then proposes a price p_t and incurs symmetric loss $\ell_t = |\boldsymbol{x}_t^\top \boldsymbol{z} - p_t|$. The learner observes neither the target value nor the loss incurred, but only the one-bit feedback $\operatorname{sign}(\boldsymbol{x}_t^\top \boldsymbol{z} - p_t)$—i.e., whether the price was too high or too low. Using this information, the learner aims to bound its total loss $\sum_{t=1}^T \mathbb{E}[\ell_t]$ (also called the "regret") independently of T. The problem of contextual search arose from feature-based pricing and auctions [KL03a, CLPL20, LPLV18]. It is useful to note that, for $d = 1$, the problem reduces to binary search, but higher-dimensional versions become much more challenging. Indeed, a series of works has studied this problem, and the state of the art is an $O(d \log d)$-competitive algorithm [LPLS21]. However, the lower bound was only $\Omega(\sqrt{d})$: can we close this gap?

Contextual bilateral trade. An extension of contextual search is *bilateral trade*, where the learner is a broker who intermediates between a seller and a buyer with hidden vectors $\boldsymbol{s}, \boldsymbol{b} \in \mathbb{R}^d$. In round t, an adversary reveals a context vector $\boldsymbol{x}_t$ with $\|\boldsymbol{x}_t\|_2 \leq 1$ corresponding to valuations $\boldsymbol{x}_t^\top \boldsymbol{s}$ for the seller and $\boldsymbol{x}_t^\top \boldsymbol{b}$ for the buyer—these are unknown to the broker, who does not see $\boldsymbol{s}, \boldsymbol{b}$. The broker then proposes a price p_t and observes whether both agents are willing to trade at that price, i.e., if $\boldsymbol{x}_t^\top \boldsymbol{s} \leq p_t \leq \boldsymbol{x}_t^\top \boldsymbol{b}$. Note that, given the feedback structure, a failed trade does not identify which agent refused. The broker aims to maximize gain from trade, defined as $\sum (\boldsymbol{x}_t^\top \boldsymbol{b} - \boldsymbol{x}_t^\top \boldsymbol{s})_+$, where the sum is taken over all rounds t in which the trade took place, or equivalently, to minimize the total loss $\sum_{t=1}^T \mathbb{E}[\ell_t]$, which is the total gain from trade it failed to realize. The model was introduced by Gaucher et al. [GBC+25], building on the online bilateral trade framework of Cesa-Bianchi et al. [CBCC+21, CBCC+24]. The gaps left by the literature are much larger in this setting: the only upper bound for this one-bit feedback model is $O(d6^d)$ [CFG+26], and the lower bound is again $\Omega(\sqrt{d})$, leaving a huge gap: can we close it?

1.1 Results

In this work, we give tight results for both these problems: we show that both problems admit $O(\sqrt{d})$-regret algorithms, which are tight by the lower bounds mentioned above. In addition to giving tight results, however, our work showcases the use of a potential function—the *square-root trace potential*—which is the central tool in both analyses. Moreover, we also present results for the Bayesian setting (where $\boldsymbol{z}$, $\boldsymbol{s}$, and $\boldsymbol{b}$ are drawn from known prior distributions), which inspire our algorithms for the adversarial setting.

Worst-case guarantees. For worst-case guarantees, the unknown target $\boldsymbol{z} \in \mathbb{R}^d$ and the valuation vectors $\boldsymbol{s}, \boldsymbol{b} \in \mathbb{R}^d$ are assumed to be selected from the Euclidean unit ball by the adversary. The adversary knows the learner's algorithm, but not the random bits it uses. This is the main point of view taken in prior works on both contextual search and bilateral trade. We obtain horizon-independent $O(\sqrt{d})$ expected-loss bounds for both problems, matching the existing lower bounds up to constant factors. Table 1 summarizes our worst-case guarantees and compares them with prior work for contextual search and bilateral trade.

Result	Contextual search	Bilateral trade
Standard lower bound [LPLS21, CFG+26]	$\sqrt{d}$	$\sqrt{d}/6$
Paes Leme and Schneider [PLS18]	$O(d^4)$	—
Liu, Paes Leme, and Schneider [LPLS21]	$O(d \log d)$	—
Cosson et al. [CFG+26]	—	$O(d\, 6^d)$
This paper (Thms. 2.4 and 3.4)	$O(\sqrt{d})$ (Alg. 2)	$O(\sqrt{d})$ (Alg. 4)

Table 1: Worst-case guarantees for unit-norm contexts and hidden vectors.

Bayesian guarantees. For Bayesian guarantees, we assume that the unknown vectors $\boldsymbol{z}, \boldsymbol{s}, \boldsymbol{b} \in \mathbb{R}^d$ are sampled from a known prior distribution with finite second moments. This point of view deviates from the existing literature but is particularly illuminating: it yields simpler algorithms and a first potential function, based on posterior covariance, that will guide the adversarial analysis. For known priors with finite second moments, our bounds depend only on the prior covariance matrices and hold at every horizon (Table 2).

Result	Contextual search	Bilateral trade
This paper (Thms. 2.1 and 3.2)	$O(\operatorname{Tr}\sqrt{\Sigma})$ (Alg. 1)	$O(\operatorname{Tr}\sqrt{\Sigma_S} + \operatorname{Tr}\sqrt{\Sigma_B})$ (Alg. 3)

Table 2: Bayesian rates for unit-norm contexts and priors with finite second moments. Σ is the prior covariance in contextual search, and Σ_S, Σ_B are the seller's and buyer's marginal prior covariances in bilateral trade.

Our worst-case guarantees and our Bayesian guarantees are clearly related, since $\operatorname{Tr}\sqrt{\Sigma} = O(\sqrt{d})$ when Σ is the covariance matrix associated with the uniform distribution in the Euclidean ball. Yet neither directly implies the other: in particular, the Bayesian setting allows for distributions with unbounded support or with skewed covariance matrices, such as Gaussian distributions. Thus, the Bayesian setting is not merely an intermediate conceptual step toward worst-case competitive guarantees, but is of independent interest.

1.2 Our Techniques

We now outline the intuition behind our algorithms and the essential ingredients of their analysis.

1.2.1 Our Algorithms

In the Bayesian setting, the learner is given a prior distribution on the hidden parameters: the vector $\boldsymbol{z}$ for contextual search, or the pair $(\boldsymbol{b}, \boldsymbol{s})$ for contextual bilateral trade. Using this, at any time t, the learner has access to the *posterior* distribution of the unknown parameters $\boldsymbol{z}$ or $\boldsymbol{s}, \boldsymbol{b}$

conditioned on all the past observations. Our Bayesian algorithms (Algorithms 1 and 3) compute a proposed price p_t based only on that posterior distribution and on the context vector $\boldsymbol{x}_t$; then they update the posterior distribution with the corresponding feedback. Our Bayesian algorithm for contextual search (Algorithm 1) is remarkably simple: it posts prices $\mu_t \pm \sigma_t/2$, where μ_t is the posterior mean of $\boldsymbol{x}_t^\top \boldsymbol{z}$, and σ_t is the posterior standard deviation of $\boldsymbol{x}_t^\top \boldsymbol{z}$. The bilateral trade algorithm (Algorithm 3) similarly uses the posterior, now of the two projected values $\boldsymbol{x}_t^\top \boldsymbol{s}$ and $\boldsymbol{x}_t^\top \boldsymbol{b}$, and posts a common price to both agents. This price is chosen to balance trading and learning, allowing us to bound total regret by the learner's initial uncertainty (Lemma 3.1).

In the worst-case setting, the learner has no posterior distribution on the unknown parameters, so it maintains a suitable "proxy" for the potential used in the Bayesian analysis. For contextual search, it maintains two parameters, a vector $\boldsymbol{m}_t \in \mathbb{R}^d$ and a matrix $R_t \in \mathbb{R}^{d \times d}$, which we call the posterior *mean proxy* and the posterior *square-root covariance proxy*. Informally, these parameters can be interpreted as an estimate of the unknown parameter (for $\boldsymbol{m}_t$) and a measure of uncertainty in that estimate in all directions (for R_t). Importantly, we do not maintain explicit uncertainty sets as in previous work—in particular, the unknown parameter need not lie within the ellipsoid $\boldsymbol{m}_t + R_t B_d$, where B_d is the unit Euclidean ball (as, e.g., in the ellipsoid method). Instead, our proofs directly show that the per-step regret can be charged to the decrease in a certain potential. From this perspective, the worst-case contextual search algorithm (Algorithm 2) posts prices $m_t \pm r_t/2$, where $m_t = \boldsymbol{x}_t^\top \boldsymbol{m}_t$, and $r_t = \|R_t \boldsymbol{x}_t\|_2$—a rule remarkably similar to the Bayesian rule in Algorithm 1. For bilateral trade (Algorithm 4), $\boldsymbol{m}_t \in \mathbb{R}^{2d}$ stores separate d-dimensional mean proxies for the seller and buyer, and the matrix $R_t \in \mathbb{R}^{2d \times 2d}$ is block diagonal, with one $d \times d$ block for each agent's square-root covariance proxy. The algorithm samples a common price and two possible successor states from a finite lottery, then keeps the successor selected by the trade bit. Whereas contextual search admits explicit state updates, here we apply the minimax theorem to guarantee the existence of a single lottery that works for every possible hidden pair of buyer and seller.

Although the focus of the paper is not on computational complexity, we note that our worst-case contextual search algorithm (Algorithm 2) uses $O(d^2)$ memory throughout the execution and performs only $O(d^2)$ operations per round, making it significantly more practical than previous approaches for contextual search that rely on costly geometric routines such as computing intrinsic volumes or barycenters of high-dimensional convex bodies [PLS18, LPLS21].

1.2.2 Analyses via the Trace-Sqrt Potential

We bound the regret of all our algorithms via a potential-based analysis, using a pair of potential functions we feel may be useful in other contexts. It will be illuminating to start with the Bayesian problem for contextual search. Here we have a random vector $\boldsymbol{z} \in \mathbb{R}^d$ with finite second moment; this represents the hidden vector that a learner does not know; however, we do know the distribution from which $\boldsymbol{z}$ is drawn, and hence know its mean $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{z}]$ and covariance $\Sigma = \mathrm{Cov}(\boldsymbol{z}) \succeq 0$. In order to quantify the uncertainty about $\boldsymbol{z}$, we use the *trace-sqrt potential* of the covariance:

$$\Psi(\Sigma) = \mathrm{Tr}\sqrt{\Sigma} = \sum_{i=1}^{d} \sigma_i, \tag{1}$$

where $\sigma_i = \sqrt{\lambda_i}$ is the square root of the i-th eigenvalue of Σ, commonly called the i-th semiaxis length; see Figure 1 for an illustration. The figure highlights a key point: the potential naturally captures a notion of "distance" and "spread", as opposed to other potentials (like the logarithm of the volume) which would need various tricks to relate regret to the reduction in the volume, incurring losses along the way.

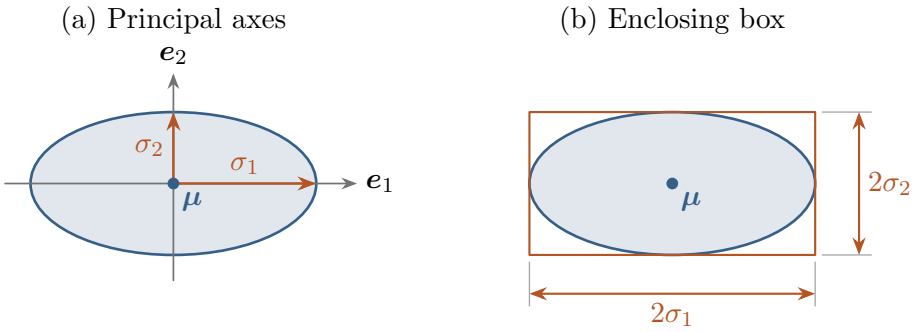


Figure 1: The covariance ellipsoid associated with Σ. Geometrically, the trace-square-root potential $\Psi = \sigma_1 + \sigma_2$ equals the 1-dimensional intrinsic volume (i.e., half the perimeter) of the smallest enclosing box of the ellipsoid, divided by 2. This perspective relates our potential to early geometric approaches to contextual search [PLS18], which relied on intrinsic volumes.

Whatever the learner observes next — the feedback from a round, a new context, or any other information revealed about $\boldsymbol{z}$ — is a *signal*: a random variable Y defined jointly with $\boldsymbol{z}$. After observing Y, the remaining uncertainty is described by $\boldsymbol{z} \mid Y$; let the new covariance matrix be Σ_Y. This means the potential moves from $\Psi(\Sigma)$ to $\Psi(\Sigma_Y)$.

The effect of signals on the potential. The first important observation is this: For any signal Y, the law of total covariance (which states that the total covariance equals the expected posterior covariance plus the covariance of the posterior mean) gives

$$\operatorname{Cov}(\boldsymbol{z}) \succeq \mathbb{E}\left[\operatorname{Cov}(\boldsymbol{z} \mid Y)\right]. \tag{2}$$

Now, the monotonicity and concavity of the operator $A \mapsto \sqrt{A}$ on the positive-semidefinite cone (see, e.g., Example 3.48 in [VB04]) allow us to infer that the potential is monotone in the Löwner order. This, combined with (2), means that

$$\Psi(\Sigma) \geq \mathbb{E}\left[\Psi(\Sigma_Y)\right] \tag{3}$$

and hence the potential is non-increasing in expectation. Of course, this expected decrease is not good enough—we need the potential decrease to pay for our regret.

Potential drops versus directional uncertainty. The key ingredient in relating the potential drop to our regret is the following claim, which is a strengthening of the matrix square-root function's monotonicity and concavity:

Lemma 1.1 (Gap Function)**.** *For any vector $\boldsymbol{x} \in \mathbb{R}^d$ with $||\boldsymbol{x}|| \leq 1$, the "gap" function*

$$g_{\boldsymbol{x}}(A) = \operatorname{Tr}\sqrt{A} - \sqrt{\boldsymbol{x}^\top A \boldsymbol{x}}$$

is monotone and concave on the positive-semidefinite cone.

We defer the proof to Lemma B.1 in Appendix B, but let us just mention that it follows by writing $g_{\boldsymbol{x}}$ as the infimum of a family of affine functions:

$$2g_{\boldsymbol{x}}(A) = \inf_{R \succ 0} \left\{ \operatorname{Tr} R - \|R\boldsymbol{x}\|_2 + \operatorname{Tr}\left[\left(R^{-1} - \frac{\boldsymbol{x}\boldsymbol{x}^\top}{\|R\boldsymbol{x}\|_2}\right) A\right] \right\}, \tag{4}$$

where $\boldsymbol{x}\boldsymbol{x}^\top/\|R\boldsymbol{x}\|_2 = 0$ when $\boldsymbol{x} = 0$. (Somewhat surprisingly, we will return to this expression in a moment!) Now the potential drop can be bounded below by the decrease in the directional standard deviation, using (2), Lemma 1.1, and Jensen's inequality:

Lemma 1.2 (Potential decrease lemma)**.** *Let $\boldsymbol{z} \in \mathbb{R}^d$ be a random vector with finite second moment and covariance matrix $\Sigma \succeq 0$, and let Y be any jointly defined random variable. Fix $\boldsymbol{x} \in \mathbb{R}^d$ with $\|\boldsymbol{x}\|_2 \le 1$. Then*

$$\underbrace{\operatorname{Tr}(\sqrt{\Sigma}) - \mathbb{E}\left[\operatorname{Tr}(\sqrt{\Sigma_Y})\right]}_{\textit{expected potential drop}} \ge \underbrace{\sqrt{\boldsymbol{x}^\top\Sigma\boldsymbol{x}} - \mathbb{E}\left[\sqrt{\boldsymbol{x}^\top\Sigma_Y\boldsymbol{x}}\right]}_{\textit{drop in directional std. dev.}}. \tag{5}$$

To get intuition for the quantity on the right-hand side, define the "uncertainty" along a specific direction $\boldsymbol{x}$ (which should be thought of as the context vector) by the "directional" standard deviation $\mathrm{sd}_{\boldsymbol{x}}(\boldsymbol{z}) = \sqrt{\boldsymbol{x}^\top\Sigma\boldsymbol{x}}$. So the right-hand side of (5) is indeed the drop in the standard deviation in the direction of the context vector. In a very concrete sense, this reduces the high-dimensional object (related to Σ) to a one-dimensional object (the reduction in standard deviation of a scalar random variable).

The Bayesian analysis. All that remains is to relate the loss incurred by our algorithm to this drop. Indeed, we show precisely this for our Bayesian contextual search algorithm: if the context vector is $\boldsymbol{x}_t$, then our algorithm's expected regret is at most a constant times the drop in standard deviation (see the proof of Theorem 2.1). Using the above inequalities and summing over all time steps, we get that the expected total regret is at most a constant times the expected total drop in potential Ψ. Since the potential is non-negative, this is at most the same constant times the initial potential, which is $\operatorname{Tr}\sqrt{\Sigma}$. This completes the Bayesian proof.

For contextual bilateral trade, our potential is the sum of two trace-sqrt potentials—one for each of the unknown buyer and seller vectors, $\Psi_{\mathrm{tr}} = \operatorname{Tr}\sqrt{\Sigma_B} + \operatorname{Tr}\sqrt{\Sigma_S}$—and we bound the lost gain from trade analogously.

Moving to the adversarial problem. Since the Bayesian problem has regret $O(\sqrt{d})$, it is tempting to use the minimax theorem to infer the existence of an algorithm for the adversarial case. We take a slightly different route for contextual search. We go back to the proof of Lemma 1.1, where we wrote $g_{\boldsymbol{x}}(\Sigma)$ as the infimum of affine functions in (4). A similar proof shows that

$$2\underbrace{\operatorname{Tr}\sqrt{\Sigma}}_{=\Psi(\Sigma)} = \inf_{\boldsymbol{m}\in\mathbb{R}^d,\, R\succ 0} \mathbb{E}\big[\underbrace{\operatorname{Tr} R + (\boldsymbol{z}-\boldsymbol{m})^\top R^{-1}(\boldsymbol{z}-\boldsymbol{m})}_{=\Phi(\boldsymbol{m},R,\boldsymbol{z})}\big]. \tag{6}$$

The function Ψ on the left was our Bayesian potential, whereas the function Φ on the right is what we use as our potential for the adversarial problem—the potentials for the "dual" Bayesian/adversarial problems are duals to each other!

We then show that our algorithm, which captures its uncertainty about $\boldsymbol{z}$ (which is now an adversarially chosen quantity) using the posterior "proxy mean" $\boldsymbol{m}_t$ and "proxy covariance" matrix R_t^2, satisfies

$$\text{expected cost of the algorithm at any time} \quad \le \quad O(1)\cdot\text{expected change in } \Phi.$$

For bilateral trade, we again sum the seller's and buyer's quadratic potentials. While in contextual search we can achieve the displayed bound through an explicit state update, in bilateral trade, a

one-round minimax argument (Lemma 3.3) instead guarantees a finite lottery of prices and successor states satisfying the bound for every hidden pair, up to arbitrarily small slack.

1.3 Related Work

Contextual search was introduced as a multidimensional generalization of binary search [PLS18], providing a common framework for earlier contextual pricing problems at the interface of computer science and economics [KL03b, ARS14, CLPL20, Jav17]. Most existing approaches to contextual search are inspired by *cutting-plane methods* in optimization, where the algorithm maintains a polytope of feasible targets that shrinks as the learner acquires information (in each round, a half-space of targets is incompatible with the feedback received). Since computing a candidate price based on the barycenter or Steiner point of the feasible polytope is insufficient [LPLS21], [PLS18] relied on tools from integral geometry. The loss bound was improved to $O(d \log d)$ in [LPLS21], by posting a price corresponding to the barycenter of an augmented version of the feasible polytope.

Our proxy covariance matrix approach is reminiscent of *ellipsoid methods* in optimization, which maintain a quadratic approximation of the feasible region [YN76, Sho77]. Ellipsoid methods have previously been applied to contextual search, e.g., in [CLPL20, LPLV18, GGK+21b, GBC+25]. A key novelty of our approach is that our proxy ellipsoid (or constant-factor dilatation of it) does not contain the feasible region. In fact, it expands in rounds where it receives "exterior" feedback! Of course, on average over several rounds, our proxy ellipsoid concentrates on the target parameters.

Related work on contextual search has also explored other loss functions, such as the *pricing loss*, which measures the revenue shortfall relative to knowing the buyer's valuation, for which [LPLS21] obtain $O(d \log \log T + d \log d)$ regret. Another criterion counts predictions whose error exceeds a prescribed tolerance ϵ; for this criterion, [LPLV18] give an algorithm making at most $O(d \log(d/\epsilon))$ errors.

The geometric viewpoint also connects contextual search to inverse optimization and low-regret cutting-plane methods [BPS17, GGK+21a, BFL25, STBO25]. In online inverse linear optimization, the learner recommends an action from a feasible set and then observes the action that maximizes the unknown linear objective over that set. Under Euclidean unit-ball normalization, [Sak26, GGL+26] obtain a horizon-independent $O(\sqrt{d})$ expected-regret bound for this problem. The distinction in feedback is important: the optimal action feedback is more informative than the one-bit contextual search feedback.

The problem of bilateral trade goes back to [MS83], which, interestingly, originally focused on a one-dimensional Bayesian setting somewhat related to ours, albeit with a single round ($T = 1$). Cesa-Bianchi et al. [CBCC+21, CBCC+24] studied a sequential version of the problem where the algorithm competes with the best fixed price in hindsight, with valuations sampled from a fixed unknown distribution; here, the regret unavoidably depends on the number of rounds T, but the dependence can be made sublinear. The contextual (multi-dimensional) problem was introduced by [GBC+25]; their noiseless algorithm, which requires two-bit feedback, has $\tilde{O}(d^2 \log T)$ regret. In the weaker one-bit feedback setting, the best time-independent result achieves $O(d\, 6^d)$ regret [CFG+26].

In variants of bilateral trade, the broker can post two distinct prices: p for the buyer and q for the seller. This extra degree of freedom can facilitate maximizing gain from trade. To restrict that advantage, it is usually assumed that the algorithm respects "global budget balance" [BCCF24], i.e., that it does not run persistent deficits, or "per-round budget balance" [CCC+24], which imposes

$q \le p$ in every round. Allowing the broker to post distinct prices also enables other measures of performance for bilateral trade, such as broker profit maximization [GBC+25, GDFS25]. Our algorithms do not use this flexibility.

1.4 Notation

We use $\|\cdot\|_2$ for the Euclidean norm, $B_2^d = \{\boldsymbol{v} \in \mathbb{R}^d : \|\boldsymbol{v}\|_2 \le 1\}$ for the Euclidean ball, and $A \succeq B$ for the positive-semidefinite Löwner order on symmetric matrices. The positive-semidefinite square root of a positive-semidefinite matrix is unique and denoted by $\Sigma^{1/2}$. We write $\mathrm{Tr}(\cdot)$ for trace, $\mathrm{Cov}(\cdot)$ for covariance, and $\mathrm{sd}(\cdot)$ for standard deviation. We set $\mathrm{sign}(0) = 1$, and write $a_+ = \max\{a, 0\}$ for the positive part.

We usually use bold symbols to denote vectors in $\mathbb{R}^d$. Specifically, we use $\boldsymbol{\mu}$ for a distributional mean and $\boldsymbol{m}$ for the algorithmic mean proxy, with scalar projections $\mu = \boldsymbol{x}^\top \boldsymbol{\mu}$ and $m = \boldsymbol{x}^\top \boldsymbol{m}$ along a context $\boldsymbol{x}$. A covariance matrix is denoted by Σ, with directional standard deviation $\sigma = \sqrt{\boldsymbol{x}^\top \Sigma \boldsymbol{x}}$.

The observed history at time t, denoted by $\mathcal{F}_{t-1}$, contains the contexts, posted prices, feedback, and the learner's randomization from previous rounds. We write $\mathbb{E}_t = \mathbb{E}[\cdot \mid \mathcal{F}_{t-1}, \boldsymbol{x}_t]$ for expectation conditional on this history and the current context, before the learner's fresh randomization. For both contextual search and bilateral trade, contexts may adapt to this history but are chosen before the learner's fresh randomization. In the Bayesian setting, each context is conditionally independent of the hidden vector or pair given $\mathcal{F}_{t-1}$, so observing it does not change the posterior.

Additional notation is summarized in the table in Appendix A.

2 Contextual Search

In contextual search, the learner is looking for an unknown target $\boldsymbol{z} \in \mathbb{R}^d$. At round t, the learner receives a context $\boldsymbol{x}_t$ with $\|\boldsymbol{x}_t\|_2 \le 1$, posts $P_t \in \mathbb{R}$, suffers loss

$$\ell_t = |\boldsymbol{x}_t^\top \boldsymbol{z} - P_t|, \qquad \text{and receives feedback} \qquad Y_t = \mathrm{sign}(\boldsymbol{x}_t^\top \boldsymbol{z} - P_t),$$

without observing either the target value or the loss. We seek bounds on the total loss $\sum_{t=1}^T \mathbb{E}[\ell_t]$ that are independent of T. In this section, we first present a Bayesian algorithm whose expected total loss is at most $24\ \mathrm{Tr}\sqrt{\Sigma}$ for any known prior on $\boldsymbol{z}$ with finite second moment and covariance matrix Σ (Theorem 2.1). We then give an algorithm that needs no prior and has expected total loss at most $12\sqrt{d}$ when the hidden vector $\boldsymbol{z}$ lies in the Euclidean unit ball (Theorem 2.4). Both guarantees hold for every horizon T. For completeness, we recall the classical $\Omega(\sqrt{d})$ lower bound for contextual search in Appendix E.

2.1 Bayesian Contextual Search

In the Bayesian setting, the target $\boldsymbol{z} \in \mathbb{R}^d$ has a known prior with finite second moment. That distribution is updated with the information acquired by the learner in previous rounds. Just before round t, we denote by

$$\boldsymbol{\mu}_t = \mathbb{E}[\boldsymbol{z} \mid \mathcal{F}_{t-1}], \qquad \Sigma_t = \mathrm{Cov}(\boldsymbol{z} \mid \mathcal{F}_{t-1})$$

the posterior mean and covariance of the target. For the current context $\boldsymbol{x}_t$, we denote the scalar target by $Z = \boldsymbol{x}_t^\top \boldsymbol{z}$, its conditional mean by $\mu_t = \boldsymbol{x}_t^\top \boldsymbol{\mu}_t$ and its standard deviation by $\sigma_t = (\boldsymbol{x}_t^\top \Sigma_t \boldsymbol{x}_t)^{1/2}$. Algorithm 1 queries one of two prices centered at μ_t, as illustrated in Figure 2.

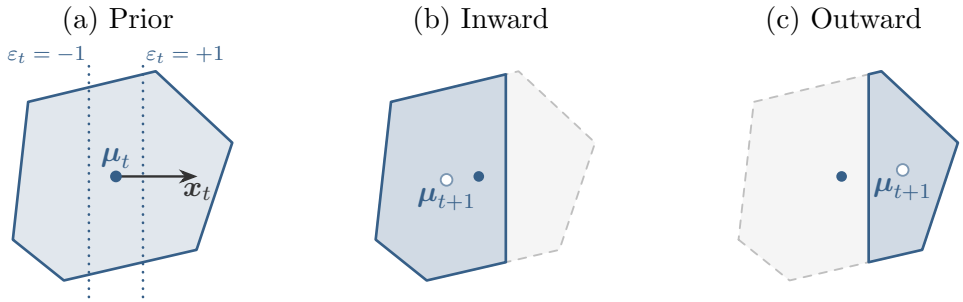


Figure 2: (a) Bayesian prior; (b) posterior for $\varepsilon_t = 1$ with $Y_t = -1$; (c) posterior for $\varepsilon_t = 1$ with $Y_t = 1$.

Algorithm 1 Bayesian Contextual Search

Given a known prior on $\boldsymbol{z}$ with finite second moment, repeat the following at each round $t = 1, 2, \ldots$:

1. Compute the posterior moments $\boldsymbol{\mu}_t, \Sigma_t$ of $\boldsymbol{z}$ and their projections on the context vector $\boldsymbol{x}_t$,
$$\mu_t = \boldsymbol{x}_t^\top \boldsymbol{\mu}_t, \qquad \sigma_t = (\boldsymbol{x}_t^\top \Sigma_t \boldsymbol{x}_t)^{1/2}.$$
2. Draw a fresh independent uniform sign $\varepsilon_t \in \{-1, 1\}$ and post the price
$$P_t = \mu_t + \frac{\sigma_t}{2}\varepsilon_t.$$
3. Observe $Y_t = \operatorname{sign}(\boldsymbol{x}_t^\top \boldsymbol{z} - P_t)$ and compute the posterior distribution on $\boldsymbol{z}$.

Theorem 2.1 (Bayesian Contextual Search)**.** *For any known prior on $\boldsymbol{z} \in \mathbb{R}^d$ with finite second moment and covariance Σ_1, Algorithm 1 satisfies, for every horizon $T \geq 1$,*

$$\sum_{t=1}^{T} \mathbb{E}\left[\ell_t\right] \leq 24 \operatorname{Tr} \sqrt{\Sigma_1}. \tag{7}$$

The expectation is over the prior and the randomization of the learner and the contexts.

Proof of Theorem 2.1. We use the following one-dimensional contraction bound for the standard deviation, stated in Lemma 2.2. The lemma, whose proof is elementary, implies that a single bit of information suffices to shrink the posterior standard deviation of a scalar distribution by a constant factor. The proof of this lemma is given in Appendix C.2.

Lemma 2.2 (One-bit deviation contraction)**.** *Let Z be a square-integrable real random variable, with $\mu = \mathbb{E}[Z]$ and $\sigma = \operatorname{sd}(Z)$. Let ε be uniform on $\{-1, 1\}$ and independent of Z. Define*

$$P = \mu + \frac{\sigma}{2}\varepsilon, \qquad Y = \operatorname{sign}(Z - P).$$

Writing the posterior standard deviation as $\sigma' = \operatorname{sd}(Z \mid P, Y)$, we have $\mathbb{E}[\sigma'] \leq \frac{23}{24}\sigma$.

The intuition is the following: because the threshold is placed at $\pm\frac{1}{2}\sigma$ away from the mean, the middle region $[\mu - \frac{1}{2}\sigma, \mu + \frac{1}{2}\sigma]$ can hold at most $\frac{1}{4}$ of the variance σ^2; at least one of the two outer tails must hold a large fraction of the variance. Conditioning on whether Z falls in that tail creates a win-win situation:

- If the tail has large probability, then every point in the tail is at least $\frac{1}{2}\sigma$ away from the mean, so learning whether Z is in the tail causes a large shift in the posterior mean, which reduces the variance (by the law of total variance).
- If the tail has small probability, then isolating the tail removes a large chunk of variance from the "main body" of the distribution (shrinking the main body's standard deviation), while the tail itself occurs so rarely that its own conditional standard deviation—which grows only like $1/\sqrt{p}$ while its probability shrinks like p—contributes almost nothing ($\propto \sqrt{p}$) to the expectation.

We now move on to the core argument. First, we bound the loss in a round of Algorithm 1 by σ_t, the uncertainty along the context vector. If $\sigma_t = 0$, then $P_t = Z = \mu_t$ almost surely, so the loss is zero. For $\sigma_t > 0$, we have

$$\mathbb{E}_t[\ell_t] = \mathbb{E}_t[|Z - P_t|] = \mathbb{E}_t\left[\max\left\{|Z-\mu_t|, \frac{\sigma_t}{2}\right\}\right] \le \frac{\sigma_t}{2} + \frac{\mathbb{E}_t\left[(Z-\mu_t)^2\right]}{2\sigma_t} = \sigma_t,$$

where we used $\max\{|u|, \sigma_t/2\} \le \sigma_t/2 + u^2/2\sigma_t$.

Next, we pay for each round's loss with a decrease in the posterior potential $\Psi_t = \operatorname{Tr}\sqrt{\Sigma_t}$, where $\Sigma_t = \operatorname{Cov}(\boldsymbol{z} \mid \mathcal{F}_{t-1})$. Let $\sigma_t = (\boldsymbol{x}_t^\top \Sigma_t \boldsymbol{x}_t)^{1/2}$ and $\sigma_t' = (\boldsymbol{x}_t^\top \Sigma_{t+1} \boldsymbol{x}_t)^{1/2}$ be the directional standard deviations of the target projection before and after the observation; we apply Lemma 2.2 to the random variable $\boldsymbol{x}_t^T \boldsymbol{z}$ and Lemma 1.2 to the potential Ψ_t to obtain the potential inequality

$$\mathbb{E}_t[\ell_t] \le \sigma_t \overset{\text{Lemma 2.2}}{\le} 24\big(\sigma_t - \mathbb{E}_t[\sigma_t']\big) \overset{\text{Lemma 1.2}}{\le} 24\big(\Psi_t - \mathbb{E}_t[\Psi_{t+1}]\big).$$

Telescoping this inequality yields the theorem:

$$\sum_{t=1}^{T} \mathbb{E}[\ell_t] \le 24\big(\Psi_1 - \mathbb{E}[\Psi_{T+1}]\big) \le 24\Psi_1. \qquad \square$$

2.2 Adversarial Contextual Search

We now consider a fixed unknown target $\boldsymbol{z}$ in the unit Euclidean ball. Our worst-case algorithm, Algorithm 2, maintains two parameters: a vector $\boldsymbol{m}_t \in \mathbb{R}^d$, the mean proxy, and a symmetric positive-definite scale matrix $R_t \in \mathbb{R}^{d\times d}$, the square-root covariance proxy. This algorithm mirrors the Bayesian algorithm, Algorithm 1, in that $\boldsymbol{m}_t$ plays the role of $\boldsymbol{\mu}_t$ and R_t plays the role of $\Sigma_t^{1/2}$ (their one-dimensional projections along the context axis are denoted by m_t and r_t). The key nuance is that, without Bayesian posterior updates, Algorithm 2 specifies explicit updates for both proxies $\boldsymbol{m}_t, R_t$ that depend only on the posted price and the feedback received, as specified in (10). Figure 3 illustrates one round of the algorithm.

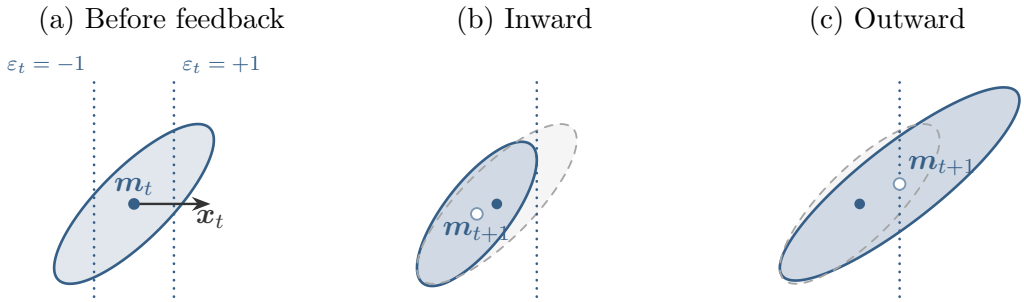


Figure 3: Proxy ellipsoids: (a) before feedback; (b) after $\varepsilon_t = 1$ with $Y_t = -1$; (c) after $\varepsilon_t = 1$ with $Y_t = 1$.

Algorithm 2 Adversarial Contextual Search

Initialize

$$\boldsymbol{m}_1 = 0, \qquad R_1 = \frac{1}{\sqrt{d}} I_d. \tag{8}$$

At each round $t = 1, 2, \ldots$:

1. Receive $\boldsymbol{x}_t$, compute $m_t = \boldsymbol{x}_t^\top \boldsymbol{m}_t$, $\boldsymbol{u}_t = R_t \boldsymbol{x}_t$, $a_t = \boldsymbol{x}_t^\top \boldsymbol{u}_t$, and $r_t = \|\boldsymbol{u}_t\|_2$.
2. Draw a fresh independent uniform sign $\varepsilon_t \in \{-1, 1\}$ and post

$$P_t = m_t + \frac{r_t}{2}\varepsilon_t. \tag{9}$$

3. Observe Y_t and update

$$\boldsymbol{m}_{t+1} = \boldsymbol{m}_t + \frac{Y_t r_t}{a_t(3 - \varepsilon_t Y_t)} \boldsymbol{u}_t, \qquad R_{t+1} = R_t + \frac{\varepsilon_t Y_t}{a_t(3 - \varepsilon_t Y_t)} \boldsymbol{u}_t \boldsymbol{u}_t^\top. \tag{10}$$

Before providing our regret guarantee, we first verify that the algorithm is well-defined, meaning that the updates in (10) preserve the positive definiteness of the scale matrix $R_t \succ 0$. This observation is obvious in the case where $\varepsilon_t Y_t = 1$, because adding a positive semidefinite matrix preserves positive definiteness; so it suffices to verify that positive definiteness is preserved whenever $\varepsilon_t Y_t = -1$.

Remark 2.3 ($R_t \succ 0$ throughout execution)**.** *Assume $R \succ 0$ and $\boldsymbol{x} \neq 0$, and write $\boldsymbol{u} = R\boldsymbol{x}$ and $a = \boldsymbol{x}^\top R \boldsymbol{x}$. Cauchy–Schwarz in the inner product induced by R gives*

$$(\boldsymbol{v}^\top \boldsymbol{u})^2 \leq (\boldsymbol{v}^\top R \boldsymbol{v})(\boldsymbol{x}^\top R \boldsymbol{x}) \quad \text{for every } \boldsymbol{v} \in \mathbb{R}^d.$$

Hence $\boldsymbol{u}\boldsymbol{u}^\top / a \preceq R$, and the update rule in (10) satisfies

$$R - \frac{\boldsymbol{u}\boldsymbol{u}^\top}{4a} \succeq \frac{3}{4} R \succ 0.$$

This proves the claim by induction over the rounds.

Theorem 2.4 (Adversarial Contextual Search)**.** *For every fixed $\boldsymbol{z}$ with $\|\boldsymbol{z}\|_2 \leq 1$, every horizon $T \geq 1$, and every admissible adaptive context strategy with $\|\boldsymbol{x}_t\|_2 \leq 1$, Algorithm 2 satisfies*

$$\sum_{t=1}^{T} \mathbb{E}\left[\ell_t\right] \leq 12\sqrt{d}. \tag{11}$$

The algorithm uses $O(d^2)$ arithmetic operations and $O(d^2)$ real-valued storage per round.

Proof of Theorem 2.4. For a fixed hidden target, we consider the non-negative potential defined by

$$\Phi_t = \Phi(\boldsymbol{m}_t, R_t, \boldsymbol{z}) = \operatorname{Tr} R_t + (\boldsymbol{z} - \boldsymbol{m}_t)^\top R_t^{-1}(\boldsymbol{z} - \boldsymbol{m}_t). \tag{12}$$

At initialization, the potential satisfies $\Phi_1 = \sqrt{d} + \sqrt{d}\,\|\boldsymbol{z}\|_2^2 \le 2\sqrt{d}$. We use the following potential-decrease bound, which is shown in Appendix C.3.

Lemma 2.5 (One-round potential decrease)**.** *On every round of Algorithm 2,*

$$\mathbb{E}_t[\ell_t] \le 6\big(\Phi_t - \mathbb{E}_t[\Phi_{t+1}]\big). \tag{13}$$

As in the Bayesian case, the proof reduces to one dimension: we define $\varphi(m, r, z) = r + \frac{(m-z)^2}{r}$, which is just a scalar version of the potential $\Phi(\cdot)$. Going from R_t to R_{t+1} is a rank 1 perturbation, so the Sherman–Morrison formula expresses the expected potential drop $\Phi_t - \mathbb{E}_t[\Phi_{t+1}]$ as a multiple of $\varphi(m_t, r_t, z) - \mathbb{E}_t[\varphi(m_{t+1}, r_{t+1}, z)]$, where m_t and r_t are the projected mean and deviation proxies; the proof now follows from elementary calculations.

Summing the potential inequality (13) over rounds yields $\sum_{t=1}^{T} \mathbb{E}[\ell_t] + 6\mathbb{E}[\Phi_{T+1}] \le 6\Phi_1 \le 12\sqrt{d}$, which proves the theorem.

In the algorithm, computing $R_t\boldsymbol{x}_t$ and performing the rank-one update each take $O(d^2)$ arithmetic operations, whereas storing R_t and $\boldsymbol{m}_t$ requires $O(d^2)$ real numbers. Importantly, the algorithm never evaluates Φ_t or computes R_t^{-1}, which only serve in the analysis. □

3 Contextual Bilateral Trade

In bilateral trade, a broker intermediates between a seller and a buyer who are represented by two vectors $\boldsymbol{s}, \boldsymbol{b} \in \mathbb{R}^d$. At round t, the broker receives a context vector $\boldsymbol{x}_t$ with $\|\boldsymbol{x}_t\|_2 \le 1$, proposes a price p_t, and observes whether trade occurs:

$$Y_t = \mathbf{1}\{\boldsymbol{x}_t^\top \boldsymbol{s} \le p_t \le \boldsymbol{x}_t^\top \boldsymbol{b}\}.$$

In particular, a failed trade does not identify the agent who rejected the price. The lost "gain from trade" in round t is defined as $\ell_t = (\boldsymbol{x}_t^\top(\boldsymbol{b} - \boldsymbol{s}))_+(1 - Y_t)$.

In Section 3.1 we present a Bayesian algorithm whose expected total loss is at most $2^{11}\big(\operatorname{Tr}\sqrt{\Sigma_S} + \operatorname{Tr}\sqrt{\Sigma_B}\big)$, for any known prior on the targets with marginal covariance matrices Σ_S and Σ_B (Theorem 3.2). In Section 3.2, we then give an algorithm that needs no prior and has expected total loss at most $2^{12}\sqrt{d} + 1$ when both targets lie in the unit Euclidean ball, so that $\boldsymbol{w} = (\boldsymbol{s}, \boldsymbol{b}) \in K = \mathbb{B}_2^d \times \mathbb{B}_2^d$ (Theorem 3.4). All our guarantees hold for every horizon T. For completeness, we recall the classical $\Omega(\sqrt{d})$ lower bound for bilateral trade in Appendix E.

3.1 Bayesian Contextual Bilateral Trade

In the Bayesian setting, write $W = (\boldsymbol{s}, \boldsymbol{b})$ for the random pair of valuations. Its components may be dependent, and we require only finite second moments. For a distribution π of W, define

$$\Psi_{\mathrm{tr}}(\pi) = \Psi(\operatorname{Cov}(\boldsymbol{s})) + \Psi(\operatorname{Cov}(\boldsymbol{b})) = \operatorname{Tr}\sqrt{\operatorname{Cov}(\boldsymbol{s})} + \operatorname{Tr}\sqrt{\operatorname{Cov}(\boldsymbol{b})}. \tag{14}$$

This is the sum of the two search potentials, not the trace square root of the joint covariance. If π is supported on K, then $\Psi_{\mathrm{tr}}(\pi) \le 2\sqrt{d}$, because each marginal covariance has trace at most one. For support in $[-1,1]^d \times [-1,1]^d$, the corresponding bound is $2d$.

Given a context $\boldsymbol{x}$, the one-dimensional problem along $\boldsymbol{x}$ now involves a pair (S,B) rather than a single target. For a common price $p \in \mathbb{R}$, let $Y(p) = \mathbf{1}\{S \le p \le B\}$ denote the trade bit, so that the loss incurred by posting price p is

$$L(p) = \mathbb{E}\left[(B-S)_+(1-Y(p))\right]. \tag{15}$$

To quantify the information obtained from the trade bit, for each $Z \in \{S,B\}$, we define the standard deviation drop $D_Z(p) = \sigma_Z - \mathbb{E}\left[\mathrm{sd}(Z \mid Y(p))\right]$, so that the overall drop is

$$D(p) = D_S(p) + D_B(p). \tag{16}$$

Each $D_Z(p)$ is nonnegative by the law of total variance and the concavity of the square root. We now introduce the key lemma that shows the existence of a specific price p relating these two terms. We provide a proof sketch below, while the full proof is deferred to Appendix D.1.

Lemma 3.1 (Trade or information)**.** *For every pair (S,B) of square-integrable real random variables, there is a common price p such that*

$$L(p) \le 2^{11} D(p). \tag{17}$$

The price can be chosen as a function of the joint distribution.

Proof sketch. Let $\alpha(p) = \mathbb{P}(S \le p \le B)$ be the trade profile. For fixed valuations, $(B-S)_+$ is the length of the interval of prices permitting trade. Averaging this length and interchanging expectation and integration gives

$$G = \mathbb{E}\left[(B-S)_+\right] = \int_{\mathbb{R}} \alpha(p)\, dp.$$

Write $\mu_Z = \mathbb{E}[Z]$, $\sigma_Z = \mathrm{sd}(Z)$ for $Z \in \{S,B\}$, $\sigma = \sigma_S + \sigma_B$, and $m = \mu_B - \mu_S$. A deterministic valuation (i.e., $\sigma_Z = 0$) is handled by posting that value. Otherwise, the analysis has two cases.

Overlapping case ($m < 3\sigma$). In this case, the seller's and buyer's means are close to each other. To deal with this case, we put $\overline{D} = \sup_p D(p)$ and, since $L(p) \le G$, we bound the entire area under α by constructing the envelope H in Figure 4. The crucial ingredient that we use to construct H is the following information identity, which is proved in the appendix and holds for any $Z \in \{S,B\}$:

$$\mathbb{E}\left[\left(\mathbb{E}\left[Z \mid Y(p)\right] - \mu_Z\right)^2 + \left(\mathrm{sd}(Z \mid Y(p)) - \sigma_Z\right)^2\right] = 2\sigma_Z D_Z(p). \tag{Information Identity}$$

The left-hand side averages the squared change in the mean and standard deviation after observing rejection or trade. Thus a feedback outcome that substantially changes either quantity must provide information, weighted by its probability. In the tails of $\alpha(\cdot)$, a successful trade forces a large shift of the conditional mean, yielding inverse-square bounds. For instance, if $p < \mu_S$, trading implies $\mathbb{E}\left[S \mid Y(p) = 1\right] \le p$, thus providing a uniform upper bound on $\alpha(p)$ in terms of $D_S(p)$ (a similar derivation holds for the buyer's tail):

$$\alpha(p)(\mu_S - p)^2 \le 2\sigma_S D_S(p) \implies \alpha(p) \le \frac{2\sigma_S D(p)}{(\mu_S - p)^2}. \tag{Tail Bound}$$

For $|p - \mu_S| \le \sigma_S/3$, consider the values of sellers who trade at p: they either lie within $\sigma_S/3$ of their mean or fall below $\mu_S - \sigma_S/3$ (and would also permit trade at price $\mu_S - \sigma_S/3$). Trade at $\mu_S - \sigma_S/3$ forces a conditional mean shift of at least $\sigma_S/3$, so the information identity controls these outliers' contribution to the second moment. Thus either conditioning on trade at p substantially changes the mean or spread, or the outliers needed to preserve the spread provide information at $\mu_S - \sigma_S/3$. Combining the bounds from these two prices gives

$$\alpha(p) \le \frac{225\overline{D}}{2\sigma_S}. \qquad \text{(Close-to-mean Bound)}$$

The buyer's bound is symmetric. For values of p close to the average of μ_S and μ_B, we employ the trivial upper bound $\alpha(p) \le 1$. Integrating $H(p)$ over $\mathbb{R}$ provides the desired upper bound on G (and thus on $L(p)$) in terms of $\overline{D}$. Integrating the tail bounds yields $6\overline{D}$, while the close-to-mean regions have width $2\sigma_Z/3$ and height $225\overline{D}/(2\sigma_Z)$, for a total area of $162\overline{D}$. Finally, the central plateau has height one and length $(m - \sigma/3)_+$, hence

$$G \le \int_{\mathbb{R}} H(p)\, dp \le 162\overline{D} + (m - \sigma/3)_+.$$

Since $m \le G$, this implies $\min\{G, \sigma/3\} \le 162\overline{D}$. Also, Cauchy–Schwarz gives $G \le m_+ + \sigma < 4\sigma$, so

$$G \le 12 \min\{G, \sigma/3\} \le 1944\overline{D}.$$

If $\overline{D} = 0$, then $G = 0$ and any price works. Otherwise, since $1944 < 2^{11}$, a price sufficiently close to maximizing D satisfies $L(p) \le G < 2^{11} D(p)$.

Well-separated case ($m \ge 3\sigma$). In this case, the seller's and buyer's means are sufficiently far apart that we can post their midpoint as the price. A failed trade forces either $S - \mu_S > m/2$ or $\mu_B - B > m/2$; thus, conditional on rejection, the combined root-mean-square deviation is at least $m/2 \ge 3\sigma/2$, compared with at most σ before conditioning. This increase requires changes in the conditional means or standard deviations. Their probability-weighted squared changes are bounded by $2\sigma D(p)$, while the missed gain is controlled by the same valuation deviations. Comparing these bounds gives $L(p) \le 41 D(p)$. □

Our Bayesian algorithm posts the price supplied by the lemma.

Algorithm 3 Bayesian Contextual Bilateral Trade

Input: a known joint prior π on $(\boldsymbol{s}, \boldsymbol{b})$ with finite second moments.
Initialize the posterior $\pi_1 = \pi$.
For $t = 1, 2, \ldots$:

1. Receive $\boldsymbol{x}_t$ and let π_t be the current posterior of $(\boldsymbol{s}, \boldsymbol{b})$.
2. Apply Lemma 3.1 under π_t to $S = \boldsymbol{x}_t^\top \boldsymbol{s}$ and $B = \boldsymbol{x}_t^\top \boldsymbol{b}$ and post the resulting price p_t.
3. Observe Y_t and condition the joint posterior on this bit and the posted offer.

Theorem 3.2 (Bayesian Bilateral Trade)**.** *For every known joint prior π with finite second moments, Algorithm 3 satisfies the following bound at every horizon:*

$$\sum_{t=1}^{T} \mathbb{E}\left[\ell_t\right] \le 2^{11} \Psi_{\mathrm{tr}}(\pi). \tag{18}$$

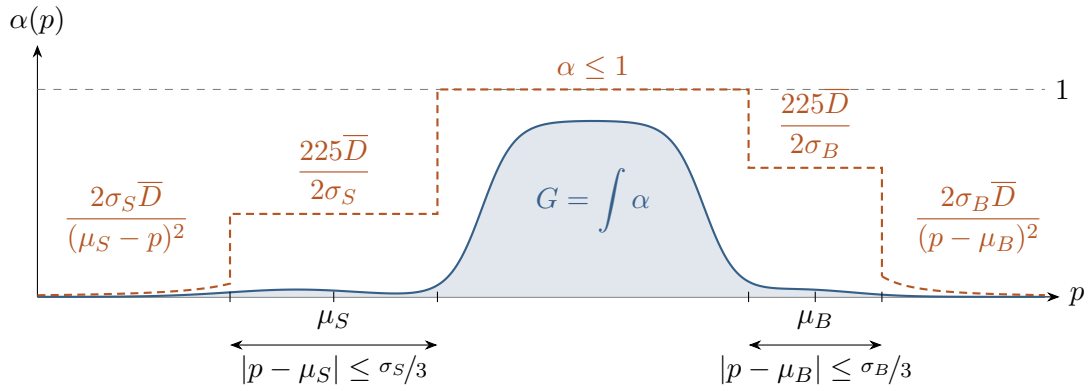


Figure 4: A schematic trade profile α (solid blue) and its piecewise upper bound H (dashed orange) in the overlapping case. The area under α is G. The tails and central intervals contribute at most $162\overline{D}$ under H; the remaining plateau has area $(m - \sigma/3)_+$. Claim D.4 establishes the bound.

Proof. Condition on the history and current context, and let π_t^+ be the posterior after the offer and trade bit. With $D(p_t)$ evaluated under π_t, the price choice and Lemma 1.2 give

$$\mathbb{E}_t[\ell_t] \overset{\text{Lemma 3.1}}{\le} 2^{11} D(p_t) \overset{\text{Lemma 1.2}}{\le} 2^{11}\big(\Psi_{\mathrm{tr}}(\pi_t) - \mathbb{E}_t \Psi_{\mathrm{tr}}(\pi_t^+)\big). \tag{19}$$

The second inequality applies the directional comparison of Lemma 1.2 to each marginal. It holds for any observed signal, including a trade bit that depends on both agents. By the context assumption in Section 1.4, the next context supplies no additional information about the hidden pair. Taking expectations and telescoping therefore yields

$$\sum_{t=1}^{T} \mathbb{E}[\ell_t] \le 2^{11}\big(\Psi_{\mathrm{tr}}(\pi) - \mathbb{E}\left[\Psi_{\mathrm{tr}}(\pi_T^+)\right]\big) \le 2^{11}\Psi_{\mathrm{tr}}(\pi).$$

The common price ensures per-round budget balance. The normalized bounds follow from (14). □

3.2 Adversarial Contextual Bilateral Trade

In the adversarial setting, no prior is given, and the hidden pair is fixed: $\boldsymbol{w} = (\boldsymbol{s}, \boldsymbol{b}) \in K = \mathbb{B}_2^d \times \mathbb{B}_2^d$. Our algorithm maintains a center $\boldsymbol{m} = (\boldsymbol{m}_S, \boldsymbol{m}_B) \in \mathbb{R}^{2d}$ and a block-diagonal positive-definite matrix $R = \mathrm{diag}(R_S, R_B) \in \mathbb{R}^{2d \times 2d}$, where both R_S and R_B belong to $\mathbb{R}^{d \times d}$. These three quantities induce the same potential as in search, but in dimension $2d$:

$$\Phi(\boldsymbol{m}, R, \boldsymbol{w}) = \mathrm{Tr}\, R + (\boldsymbol{w} - \boldsymbol{m})^\top R^{-1}(\boldsymbol{w} - \boldsymbol{m}) = \Phi(\boldsymbol{m}_S, R_S, \boldsymbol{s}) + \Phi(\boldsymbol{m}_B, R_B, \boldsymbol{b}). \tag{20}$$

Note that the learner never evaluates the potential at the unknown pair $\boldsymbol{w}$. In each time step, our algorithm designs a suitable *plan*, which specifies an offer p and two successor states $(\boldsymbol{m}^{(0)}, R^{(0)})$ and $(\boldsymbol{m}^{(1)}, R^{(1)})$, one for each possible feedback bit. Let $\boldsymbol{m}^+, R^+$ denote the random successor state induced by the observed feedback. Ideally, we would like to find a plan that balances *decreasing* the expected potential at $(\boldsymbol{m}^+, R^+, \boldsymbol{w})$ and incurring little regret $\ell(\boldsymbol{w})$. The following lemma shows that there exists a lottery over prices (i.e., a randomized plan) that ensures this property, up to some slack.

Lemma 3.3 (A valid lottery exists)**.** *For every $\boldsymbol{m} \in \mathbb{R}^{2d}$, every block-diagonal $R \succ 0$, every context $\boldsymbol{x}$ with $\|\boldsymbol{x}\|_2 \le 1$, and every $\gamma \in (0, 1]$, there is a finite lottery of plans satisfying*

$$\mathbb{E}\left[\ell(\boldsymbol{w}) + 2^{10}\Phi(\boldsymbol{m}^+, R^+, \boldsymbol{w})\right] \le 2^{10}\Phi(\boldsymbol{m}, R, \boldsymbol{w}) + \gamma \qquad \text{for every } \boldsymbol{w} \in K. \tag{21}$$

Each plan in the lottery posts the same price, drawn from a finite grid, to both agents. The lottery can be chosen as a function of the state and context.

Note that the expectation in the lemma is only over the lottery, and the successor is selected using the bit generated by the sampled plan. For a given slack $\gamma \in (0, 1]$, the lottery must satisfy the one-round bound in (21), which balances incurred regret against the decrease in the (unknown) potential and must hold for all possible hidden pairs. We provide here a proof sketch of the lottery lemma, and defer the detailed proof to Appendix D.2.

Proof sketch. The goal is to show that, for every fixed current state and context, we can first construct a good *plan* for each hypothetical prior, and then use minimax to obtain one *lottery* that works for every hidden pair. A *plan* a consists of a common price and two successor states, one for rejection and one for trade, all fixed before observing the bit. A *lottery* Q is a probability distribution over plans: the learner samples a plan, posts its price, and keeps the successor selected by the observed bit. We seek to find such a lottery and to ensure that it has finite support.

Let us write $\Phi_0(\boldsymbol{w}) = \Phi(\boldsymbol{m}, R, \boldsymbol{w})$ for the current quadratic potential from (6), and $\Phi_a^+(\boldsymbol{w})$ for the successor quadratic potential selected when the hidden pair is $\boldsymbol{w}$. Moreover, let us define the excess cost

$$f_a(\boldsymbol{w}) = \ell_a(\boldsymbol{w}) + 2^{10}\Phi_a^+(\boldsymbol{w}) - 2^{10}\Phi_0(\boldsymbol{w}), \tag{Excess cost}$$

where ℓ_a is the lost gain under plan a. The required inequality is therefore $\mathbb{E}_{a\sim Q}\left[f_a(\boldsymbol{w})\right] \le \gamma$ for every $\boldsymbol{w} \in K$. Indeed, f_a is the loss plus the scaled successor potential minus the scaled current potential, so this is exactly (21).

A plan for each prior. Temporarily suppose that the hidden pair W is sampled from a known prior π on K. Lemma 3.1, together with the directional comparison of Lemma 1.2, supplies a common price p satisfying

$$L(p) \le 2^{11}\left(\Psi_{\mathrm{tr}}(\pi) - \mathbb{P}_\pi(Y(p) = 0)\Psi_{\mathrm{tr}}(\pi_0) - \mathbb{P}_\pi(Y(p) = 1)\Psi_{\mathrm{tr}}(\pi_1)\right), \tag{Bayesian bound}$$

where $\pi_y = \mathcal{L}_\pi(W \mid Y(p) = y)$ is the posterior after observing bit y. Thus the scaled prior potential covers both the expected loss and the scaled posterior potential remaining after the bit. We want to replace these potentials by the quadratic potentials of the current and successor states.

For each bit, choose the posterior mean as the successor center and the square roots of the posterior marginal covariances as the matrix blocks, adding λI_d before taking the roots to ensure positive definiteness. These choices define a plan a_λ. We stress that both successors are fixed before posting the price, and the observed bit only selects which successor is kept. The variational identity (6) gives

$$\begin{aligned} \mathbb{E}_\pi\left[\Phi_0(W)\right] &\ge 2\Psi_{\mathrm{tr}}(\pi), \\ \lim_{\lambda\to 0^+} \mathbb{E}_\pi\left[\Phi_{a_\lambda}^+(W)\right] &= 2\left(\mathbb{P}_\pi(Y(p) = 0)\Psi_{\mathrm{tr}}(\pi_0) + \mathbb{P}_\pi(Y(p) = 1)\Psi_{\mathrm{tr}}(\pi_1)\right). \end{aligned} \tag{Potential bounds}$$

Combining these comparisons with the Bayesian bound cancels the posterior terms:

$$\begin{aligned} \inf_a \mathbb{E}_\pi\left[f_a(W)\right] &\le \lim_{\lambda\to 0^+} \mathbb{E}_\pi\left[f_{a_\lambda}(W)\right] \\ &\le 2^{11}\Psi_{\mathrm{tr}}(\pi) - 2^{10}\mathbb{E}_\pi\left[\Phi_0(W)\right] \le 0. \end{aligned} \tag{Average bound}$$

We have therefore constructed plans with arbitrarily small average excess cost for every prior, although the plans may depend on that prior.

One lottery for every hidden pair. It remains to turn a suitable plan for *each* prior into a *single* lottery that works for every hidden pair. The obstacle to applying minimax directly is that the trade bit jumps when a valuation crosses the posted price. In the analysis, we slightly contract the hidden pair toward the origin and add a small common random shift to both valuations. The contraction keeps the perturbed pair in K, while averaging over the shift makes the expected cost continuous. Corresponding changes to the price and successor states preserve the averaged bound up to a uniformly vanishing error. We then apply Claim D.5 together with Sion's minimax theorem (Theorem D.6 [Sio58]) to these continuous costs and then let the perturbation vanish, giving

$$\inf_{Q} \sup_{\boldsymbol{w} \in K} \mathbb{E}_{a \sim Q} \left[f_a(\boldsymbol{w}) \right] = \sup_{\rho \in \mathcal{P}(K)} \inf_{a} \mathbb{E}_{\rho} \left[f_a(W) \right] \leq 0. \tag{Minimax}$$

Above, we write $\mathcal{P}(K)$ for the set of all probability measures on K, representing the possible prior distributions of the hidden pair. At this point, Q ranges over arbitrary lotteries, not necessarily with finite support. Since the minimax value is at most zero, we can find a lottery with excess cost below any positive tolerance. We choose $\gamma/3$, leaving the remaining $2\gamma/3$ for the price and successor-state approximations below. Consequently, there is a lottery satisfying

$$\sup_{\boldsymbol{w} \in K} \mathbb{E}_{a \sim Q} \left[f_a(\boldsymbol{w}) \right] < \frac{\gamma}{3}. \tag{Uniform bound}$$

This bound holds for every fixed hidden pair, with expectation only over the sampled plan.

A finite lottery. The same claim lets us choose Q with a bounded price density and compact ranges of prices and successor states. Using this regularity and compactness, we round prices to a finite grid and approximate successor states by a finite net, making each error smaller than $\gamma/3$. Thus, the resulting lottery has finite support and satisfies the required bound with slack γ.

The detailed proof is deferred to Appendix D.2. □

Algorithm 4 Adversarial Contextual Bilateral Trade

Initialize $\boldsymbol{m}_1 = 0 \in \mathbb{R}^{2d}$ and $R_1 = I_{2d}/\sqrt{d}$.
For $t = 1, 2, \ldots$:

1. Receive $\boldsymbol{x}_t$ and set $\gamma_t = \frac{1}{t(t+1)}$.
2. Choose a lottery of plans as in Lemma 3.3 with current state $(\boldsymbol{m}_t, R_t)$, context $\boldsymbol{x}_t$, and slack γ_t.
3. Draw a plan from the finite lottery and post its price p_t.
4. Observe Y_t and retain the plan's successor state for that bit as $(\boldsymbol{m}_{t+1}, R_{t+1})$.

The algorithm uses only its state, the current context, its own randomization, and the observed bit. It needs neither a prior nor a horizon. We do not claim an efficient implementation of the optimization in step 2.

Theorem 3.4 (Adversarial Bilateral Trade)**.** *For every fixed $\boldsymbol{w} \in K$, every admissible adaptive context strategy with $\|\boldsymbol{x}_t\|_2 \leq 1$, and every horizon T, Algorithm 4 satisfies*

$$\sum_{t=1}^{T} \mathbb{E} \left[\ell_t \right] \leq 2^{12} \sqrt{d} + 1. \tag{22}$$

Proof. Fix $\boldsymbol{w} = (\boldsymbol{s}, \boldsymbol{b}) \in K$ and write $\Phi_t = \Phi(\boldsymbol{m}_t, R_t, \boldsymbol{w})$. Lemma 3.3 guarantees the existence of the lottery required by Algorithm 4 at every round. Conditional on the observed history and current context, before the lottery is drawn, (21) gives

$$\mathbb{E}_t[\ell_t] + 2^{10}\mathbb{E}_t[\Phi_{t+1}] \le 2^{10}\Phi_t + \gamma_t.$$

Since Φ_1 is finite and all losses and potentials are nonnegative, this inequality also establishes their integrability inductively. Taking full expectations and summing over $t = 1, \dots, T$ telescopes the potential terms:

$$\sum_{t=1}^{T} \mathbb{E}[\ell_t] + 2^{10}\mathbb{E}[\Phi_{T+1}] \le 2^{10}\Phi_1 + \sum_{t=1}^{T} \gamma_t.$$

The chosen slacks have total sum at most one, because

$$\sum_{t=1}^{T} \gamma_t = \sum_{t=1}^{T} \frac{1}{t(t+1)} = \sum_{t=1}^{T} \left(\frac{1}{t} - \frac{1}{t+1} \right) = 1 - \frac{1}{T+1} < 1.$$

Finally, the initialization and $\|\boldsymbol{s}\|_2, \|\boldsymbol{b}\|_2 \le 1$ give

$$\Phi_1 = \Phi(0, I_{2d}/\sqrt{d}, \boldsymbol{w}) = 2\sqrt{d} + \sqrt{d}(\|\boldsymbol{s}\|_2^2 + \|\boldsymbol{b}\|_2^2) \le 4\sqrt{d}.$$

Substituting these two bounds proves (22). □

Acknowledgment

Federico Fusco, Stefano Leonardi, and Matteo Russo acknowledge OpenAI for providing complimentary access through the ChatGPT for Academic Researchers program. The work of Federico Fusco and Stefano Leonardi is partly supported by AI-PHOQUS PN RIC 2021-2027, co-financed by EU Italian Ministry MIMIT and MUR. Anupam Gupta was supported in part by NSF awards CCF-2422926 and CCF-2608359.

Disclosure of AI Usage

The first proof of $O(\sqrt{d})$ regret for the adversarial version of contextual bilateral trade was suggested by GPT-6 Astra in conversations with the authors. Subsequently, we dedicated considerable effort to simplifying the proofs, to understanding and interpreting the potential Φ used in the proof, to exploring the (somewhat simpler) setting of contextual search, to identifying concrete connections between the Bayesian and adversarial versions of both problems, and to investigating a Gaussian setting of contextual search, which led to interpreting the parameters of the algorithm as posterior mean and covariance proxies. The authors have substantially rewritten the proofs (albeit with some help from Gemini, Claude, and ChatGPT) and take full responsibility for the content of this work.

All the proofs have been formalized in Lean4 and can be found at this url.

Appendix

A Notation

We summarize the main notations used in the paper in the table below. We generally use bold symbols for d-dimensional vectors, capital Roman letters for matrices and scalar random variables, and Greek letters for distributional quantities (such as moments) and potentials. Time dependence is usually indicated by a subscript; for example, $\Psi_t = \Psi(\Sigma_t)$ denotes the potential at the t-th round.

Quantity	Dimension d	Projection	Random variable
Contextual search			
Unknown target	$\boldsymbol{z}$	z	$\boldsymbol{z}, Z$
Distributional mean	$\mathbb{E}[\boldsymbol{z}] = \boldsymbol{\mu}$	μ	
Mean proxy	$\boldsymbol{m}$	m	
Distributional covariance	$\mathrm{Cov}(\boldsymbol{z}) = \Sigma$	σ^2	
Square root of covariance	$\sqrt{\mathrm{Cov}(\boldsymbol{z})} = \Sigma^{1/2}$	$\mathrm{sd}(Z) = \sigma$	
Square root covariance proxy	R	r	
Context	$\boldsymbol{x}$		
Price		$p \in \mathbb{R}$	$P \in \mathbb{R}$
Feedback		$y \in \{-1, 1\}$	$Y \in \{-1, 1\}$
Random sign (Rademacher)			$\varepsilon \in \{-1, 1\}$
Contextual bilateral trade			
Unknown seller/buyer valuations	$\boldsymbol{w} = (\boldsymbol{s}, \boldsymbol{b})$	S, B	$S, B \in \mathbb{R}$
Joint law	$\pi = \mathcal{L}(\boldsymbol{s}, \boldsymbol{b})$		
Potentials and functions			
Bayesian potential	$\Psi(\Sigma) = \mathrm{Tr}\sqrt{\Sigma}$		
Bayesian trade potential	$\Psi_{\mathrm{tr}}(\pi) = \Psi(\mathrm{Cov}(\boldsymbol{s})) + \Psi(\mathrm{Cov}(\boldsymbol{b}))$		
Worst-case potential	$\Phi(\boldsymbol{m}, R, \boldsymbol{z}) = \mathrm{Tr}\, R + (\boldsymbol{z} - \boldsymbol{m})^\top R^{-1}(\boldsymbol{z} - \boldsymbol{m})$		
Scalar potential	$\varphi(m, r, z) = r + (m - z)^2/r$		
Gap function	$g_{\boldsymbol{x}}(\Sigma) = \mathrm{Tr}\sqrt{\Sigma} - \sqrt{\boldsymbol{x}^\top \Sigma \boldsymbol{x}}$		

B Concavity of the Gap Function

This appendix proves the concavity and monotonicity of the gap function $g_{\boldsymbol{x}}$ used in Lemma 1.2. We do so by expressing $g_{\boldsymbol{x}}$ as an infimum of affine functions with positive-semidefinite slopes.

Lemma B.1 (Gap representation and directional comparison)**.** *Fix $\boldsymbol{x} \in \mathbb{R}^d$ with $\|\boldsymbol{x}\|_2 \le 1$ and define, for $\Sigma \succeq 0$,*

$$g_{\boldsymbol{x}}(\Sigma) = \mathrm{Tr}\sqrt{\Sigma} - \sqrt{\boldsymbol{x}^\top \Sigma \boldsymbol{x}}. \tag{23}$$

The function $g_{\boldsymbol{x}}$ admits the affine representation

$$2g_{\boldsymbol{x}}(\Sigma) = \inf_{R \succ 0} \left\{ \mathrm{Tr}\, R - \|R\boldsymbol{x}\|_2 + \mathrm{Tr}\left[\left(R^{-1} - \frac{\boldsymbol{x}\boldsymbol{x}^\top}{\|R\boldsymbol{x}\|_2} \right) \Sigma \right] \right\}. \tag{24}$$

Here $\boldsymbol{x}\boldsymbol{x}^\top/\|R\boldsymbol{x}\|_2$ is defined to be zero when $\boldsymbol{x} = 0$. If $\Sigma \succ 0$, the infimum is attained at $R = \Sigma^{1/2}$.

Each function inside the infimum in (24) is affine in Σ and has positive-semidefinite slope. Consequently, $g_{\boldsymbol{x}}$ is concave as a pointwise infimum of affine functions, and it is nondecreasing in the positive semidefinite order. Applied to $\mathbb{E}_t[\Sigma_{t+1}] \preceq \Sigma_t$ in the Bayesian search model, the lemma thus implies

$$\Psi_t - \mathbb{E}_t\left[\Psi_{t+1}\right] \geq \sigma_t - \mathbb{E}_t\left[\sigma_t'\right]. \tag{25}$$

Proof of Lemma B.1. Fix $R \succ 0$ and put $T = \Sigma^{1/2}$, $M = (T-R)^2 \succeq 0$, $r = \|R\boldsymbol{x}\|_2$, and $\sigma = \|T\boldsymbol{x}\|_2$. Expanding the square and using cyclicity of the trace gives

$$\operatorname{Tr}(R^{-1}M) = \operatorname{Tr}(R^{-1}\Sigma) - 2\operatorname{Tr} T + \operatorname{Tr} R. \tag{26}$$

For $\boldsymbol{x} \neq 0$, Cauchy–Schwarz in the R-inner product, followed by $\boldsymbol{x}^\top R\boldsymbol{x} \leq \|\boldsymbol{x}\|_2\|R\boldsymbol{x}\|_2 \leq r$, yields

$$R^{-1} \succeq \frac{\boldsymbol{x}\boldsymbol{x}^\top}{\boldsymbol{x}^\top R\boldsymbol{x}} \succeq \frac{\boldsymbol{x}\boldsymbol{x}^\top}{r}.$$

Thus $H_R = R^{-1} - \boldsymbol{x}\boldsymbol{x}^\top/r \succeq 0$, and the reverse triangle inequality gives

$$\operatorname{Tr}(R^{-1}M) \geq \frac{\boldsymbol{x}^\top M\boldsymbol{x}}{r} = \frac{\|(T-R)\boldsymbol{x}\|_2^2}{r} \geq \frac{(\sigma - r)^2}{r}.$$

Combining this inequality with (26) and rearranging proves

$$2\operatorname{Tr} T - 2\sigma \leq \operatorname{Tr} R - r + \operatorname{Tr}(H_R\Sigma). \tag{27}$$

For $\boldsymbol{x} = 0$, take $H_R = R^{-1}$; the same bound follows directly from (26). If $\Sigma \succ 0$, equality holds at $R = T$. For singular Σ, the same value is approached by $R_\eta = (\Sigma + \eta I)^{1/2}$ as $\eta \to 0^+$. □

C Supporting Proofs for Contextual Search

We have deferred two proofs in Section 2: Lemma 2.2 in the Bayesian analysis and Lemma 2.5 in the worst-case analysis. Both rely on the same one-dimensional potential decrease, stated in Lemma C.1 below. We first establish this common inequality, then use it to complete the two proofs.

C.1 A One-Dimensional Potential Lemma

Lemma C.1 (One-dimensional potential decrease)**.** *For $m, z \in \mathbb{R}$ and $r > 0$, define*

$$\varphi(m, r, z) = r + \frac{(m-z)^2}{r}.$$

Let ε be uniform on $\{-1, 1\}$, set

$$P = m + \frac{r}{2}\varepsilon, \quad Y = \operatorname{sign}(z - P),$$

and define

$$m^+ = m + \frac{rY}{3 - \varepsilon Y}, \qquad r^+ = \frac{3r}{3 - \varepsilon Y}.$$

Then

$$\varphi(m, r, z) - \mathbb{E}_\varepsilon\left[\varphi(m^+, r^+, z)\right] \geq \frac{1}{6}\mathbb{E}_\varepsilon\left[|z - P|\right] = \frac{1}{6}\max\left\{|z - m|, \frac{r}{2}\right\}. \tag{28}$$

In particular,

$$\mathbb{E}_\varepsilon\left[\varphi(m^+, r^+, z)\right] \leq \varphi(m, r, z) - \frac{r}{12}. \tag{29}$$

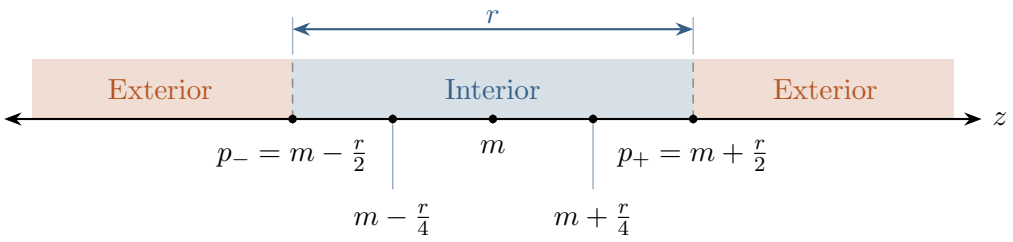


Figure 5: Prices and possible updated centers $m^+ \in \{m \pm r/4,\, m \pm r/2\}$ in Lemma C.1.

Proof. The identity for the expected loss follows by averaging over the two prices. We prove the potential bound by separating targets between the two prices from targets outside that interval, as illustrated in Figure 5.

Interior: $|z-m| < r/2$**.** Both answers point inward, so $\varepsilon Y = -1$ and $r^+ = 3r/4$. The new center is $m + r/4$ or $m - r/4$, each with probability $1/2$. Thus

$$\mathbb{E}_\varepsilon\left[(m^+ - z)^2\right] = (m-z)^2 + \frac{r^2}{16},$$

and

$$\begin{aligned}\mathbb{E}_\varepsilon\left[\varphi(m^+, r^+, z)\right] &= \frac{3r}{4} + \frac{4}{3r}\left((m-z)^2 + \frac{r^2}{16}\right) \\ &= \varphi(m, r, z) - \frac{r}{4} + \frac{(m-z)^2}{3r} + \frac{r}{12} \le \varphi(m, r, z) - \frac{r}{12}.\end{aligned}$$

Here $\mathbb{E}_\varepsilon\left[|z - P|\right] = r/2$, proving (28) in this case.

Exterior: $|z - m| > r/2$**.** By reflection, it suffices to consider $z > m + r/2$. Both prices yield the answer $Y = 1$, and the two equally likely successor states are

$$(m^+, r^+) = (m + r/4, 3r/4) \quad \text{and} \quad (m + r/2, 3r/2).$$

Consequently,

$$\mathbb{E}_\varepsilon\left[r^+\right] = \frac{9r}{8},$$

and

$$\begin{aligned}\mathbb{E}_\varepsilon\left[\frac{(z-m^+)^2}{r^+}\right] &= \frac{1}{2}\frac{(z-m-r/4)^2}{3r/4} + \frac{1}{2}\frac{(z-m-r/2)^2}{3r/2} \\ &= \frac{(z-m)^2}{r} - \frac{2}{3}(z-m) + \frac{r}{8}.\end{aligned}$$

Adding the two terms and using $z - m \ge r/2$ gives

$$\mathbb{E}_\varepsilon\left[\varphi(m^+, r^+, z)\right] = \varphi(m, r, z) + \frac{r}{4} - \frac{2}{3}(z-m) \le \varphi(m, r, z) - \frac{1}{6}(z-m).$$

Since $\mathbb{E}_\varepsilon\left[|z - P|\right] = z - m$, this proves the claim for exterior targets.

Boundary: $|z-m| = r/2$**.** Our convention $\operatorname{sign}(0) = 1$ gives one inward and one outward update when $z = m + r/2$, as in the exterior calculation. When $z = m - r/2$, both updates are inward, as in the interior calculation. Substituting into the corresponding expressions gives a decrease of exactly $r/12$ in either case, while the expected loss is $r/2$. This proves (28) at both endpoints. Finally, the expected loss is always at least $r/2$, which implies (29). □

C.2 Proof of Lemma 2.2

Lemma 2.2 (One-bit deviation contraction)**.** *Let Z be a square-integrable real random variable, with $\mu = \mathbb{E}[Z]$ and $\sigma = \operatorname{sd}(Z)$. Let ε be uniform on $\{-1, 1\}$ and independent of Z. Define*

$$P = \mu + \frac{\sigma}{2}\varepsilon, \qquad Y = \operatorname{sign}(Z - P).$$

Writing the posterior standard deviation as $\sigma' = \operatorname{sd}(Z \mid P, Y)$, we have $\mathbb{E}[\sigma'] \le \frac{23}{24}\sigma$.

Proof of Lemma 2.2. If $\sigma = 0$, then $Z = \mu$ almost surely and $\sigma' = 0$, so the result is immediate. Assume $\sigma > 0$. For any mean proxy $m \in \mathbb{R}$ and deviation proxy $r > 0$,

$$2\operatorname{sd}(Z) \le 2\sqrt{\mathbb{E}[(Z-m)^2]} \le r + \frac{\mathbb{E}\left[(Z-m)^2\right]}{r} = \mathbb{E}[\varphi(m, r, Z)]. \tag{30}$$

Equality holds at $m = \mu$ and $r = \sigma$. We apply this bound conditionally after observing the price and answer, with the proxies

$$m^+ = \mu + \frac{\sigma Y}{3 - \varepsilon Y}, \qquad r^+ = \frac{3\sigma}{3 - \varepsilon Y}.$$

Because $\sigma > 0$, the sign ε is determined by P, so these proxies are (P, Y)-measurable. Applying (30) to the conditional law yields

$$2\sigma' \le \mathbb{E}\left[\varphi(m^+, r^+, Z) \mid P, Y\right], \qquad 2\mathbb{E}\left[\sigma'\right] \le \mathbb{E}\left[\varphi(m^+, r^+, Z)\right].$$

For each fixed $Z = z$, Lemma C.1 with $m = \mu$ and $r = \sigma$ gives

$$\mathbb{E}\left[\varphi(m^+, r^+, Z) \mid Z = z\right] \le \varphi(\mu, \sigma, z) - \frac{\sigma}{12}.$$

Averaging over Z and using $\mathbb{E}[\varphi(\mu, \sigma, Z)] = 2\sigma$ gives

$$2\mathbb{E}\left[\sigma'\right] \le 2\sigma - \frac{\sigma}{12} = \frac{23}{12}\sigma.$$

Division by two proves the lemma. □

C.3 Proof of Lemma 2.5

Lemma 2.5 (One-round potential decrease)**.** *On every round of Algorithm 2,*

$$\mathbb{E}_t[\ell_t] \le 6\big(\Phi_t - \mathbb{E}_t[\Phi_{t+1}]\big). \tag{13}$$

Proof of Lemma 2.5. For $\boldsymbol{x}_t = 0$, the loss and potential change are both zero. Otherwise, condition on the history and current context, suppress time subscripts, and put

$$m = \boldsymbol{x}^\top \boldsymbol{m}, \quad z = \boldsymbol{x}^\top \boldsymbol{z}, \quad \boldsymbol{u} = R\boldsymbol{x}, \quad a = \boldsymbol{x}^\top R\boldsymbol{x}, \quad r = \|R\boldsymbol{x}\|_2.$$

Since $R \succ 0$ and $\|\boldsymbol{x}\|_2 \le 1$, we have

$$0 < a = \boldsymbol{x}^\top \boldsymbol{u} \le \|\boldsymbol{x}\|_2 \|\boldsymbol{u}\|_2 \le r. \tag{31}$$

For a realization of the random sign and the answer, let $\boldsymbol{m}_+, R_+$ be the updated state. The update formulas imply

$$m^+ = \boldsymbol{x}^\top \boldsymbol{m}_+ = m + \frac{Yr}{3-\varepsilon Y}, \qquad R_+ \boldsymbol{x} = \frac{3}{3-\varepsilon Y} R\boldsymbol{x}, \qquad r^+ = \|R_+ \boldsymbol{x}\|_2 = \frac{3r}{3-\varepsilon Y}.$$

Thus m^+, r^+ are exactly the successor quantities in Lemma C.1.

We first prove the identity

$$\Phi(\boldsymbol{m}_+, R_+, \boldsymbol{z}) - \Phi(\boldsymbol{m}, R, \boldsymbol{z}) = \frac{r}{a}\big(\varphi(m^+, r^+, z) - \varphi(m, r, z)\big). \tag{32}$$

The trace change is

$$\operatorname{Tr}(R_+ - R) = \frac{\varepsilon Y r^2}{a(3-\varepsilon Y)} = \frac{r}{a}(r^+ - r). \tag{33}$$

For the quadratic term, write $\delta = m^+ - m = Yr/(3-\varepsilon Y)$. The updates give

$$\boldsymbol{m}_+ - \boldsymbol{m} = \frac{\delta}{a}\boldsymbol{u}, \qquad R_+^{-1} = R^{-1} - \frac{\varepsilon Y}{3a}\boldsymbol{x}\boldsymbol{x}^\top.$$

The inverse identity follows directly by multiplication with R_+. Using $R^{-1}\boldsymbol{u} = \boldsymbol{x}$ and $\boldsymbol{u}^\top R^{-1}\boldsymbol{u} = a$, we obtain

$$\begin{aligned}
&(\boldsymbol{z} - \boldsymbol{m}_+)^\top R_+^{-1}(\boldsymbol{z} - \boldsymbol{m}_+) \\
&\quad = (\boldsymbol{z} - \boldsymbol{m})^\top R^{-1}(\boldsymbol{z} - \boldsymbol{m}) - \frac{2\delta(z-m)}{a} + \frac{\delta^2}{a} - \frac{\varepsilon Y}{3a}(z - m^+)^2 \\
&\quad = (\boldsymbol{z} - \boldsymbol{m})^\top R^{-1}(\boldsymbol{z} - \boldsymbol{m}) + \frac{3-\varepsilon Y}{3a}(z - m^+)^2 - \frac{(z-m)^2}{a} \\
&\quad = (\boldsymbol{z} - \boldsymbol{m})^\top R^{-1}(\boldsymbol{z} - \boldsymbol{m}) + \frac{r}{a}\left(\frac{(z-m^+)^2}{r^+} - \frac{(z-m)^2}{r}\right).
\end{aligned}$$

The second equality uses $z - m^+ = (z-m) - \delta$; the third uses $r^+ = 3r/(3-\varepsilon Y)$. Adding (33) proves (32).

Finally, take the conditional expectation and apply Lemma C.1:

$$\begin{aligned}
\Phi - \mathbb{E}_t[\Phi_+] &= \frac{r}{a}\big(\varphi(m, r, z) - \mathbb{E}_t\left[\varphi(m^+, r^+, z)\right]\big) \\
&\ge \frac{r}{6a}\mathbb{E}_t\left[|z - P|\right] \ge \frac{1}{6}\mathbb{E}_t\left[\ell_t\right],
\end{aligned}$$

where the last inequality is (31). This is the desired drift bound. □

D Supporting Proofs for Bilateral Trade

We first prove the trade-or-information inequality of Lemma 3.1, which directly yields the Bayesian guarantee. We then use it, together with quadratic bounds derived from Lemma B.1 and a minimax argument, to prove the lottery-existence statement of Lemma 3.3.

D.1 Proof of Lemma 3.1

Recall that $Y(p) = \mathbf{1}\{S \le p \le B\}$ indicates whether trade occurs at price p. The expected lost gain and the expected reduction in the sum of the two standard deviations are, respectively,

$$L(p) = \mathbb{E}\left[(B-S)_+(1-Y(p))\right],$$
$$D(p) = (\mathrm{sd}(S) - \mathbb{E}\left[\mathrm{sd}(S \mid Y(p))\right]) + (\mathrm{sd}(B) - \mathbb{E}\left[\mathrm{sd}(B \mid Y(p))\right]).$$

Before proving the lemma, we show that $D(p)$ is nonnegative, provide a useful integral representation of the gain from trade, and introduce a useful result on the information produced by binary feedback.

Claim D.1. *For every pair (S, B) of square-integrable real random variables, $D(p) \ge 0$ for all $p \in \mathbb{R}$.*

Proof. For each $Z \in \{S, B\}$, the law of total variance gives

$$\mathrm{Var}(Z) = \mathbb{E}\big[\mathrm{Var}(Z \mid Y(p))\big] + \mathrm{Var}\big(\mathbb{E}[Z \mid Y(p)]\big) \ge \mathbb{E}\big[\mathrm{Var}(Z \mid Y(p))\big].$$

Consequently, by Jensen's inequality and the concavity of the square root,

$$\begin{aligned}\sum_{y\in\{0,1\}} \mathbb{P}(Y(p)=y)\,\mathrm{sd}(Z \mid Y(p)=y) &= \mathbb{E}\left[\sqrt{\mathrm{Var}(Z \mid Y(p))}\right] \\ &\le \sqrt{\mathbb{E}\big[\mathrm{Var}(Z \mid Y(p))\big]} \\ &\le \mathrm{sd}(Z).\end{aligned}$$

Thus, assuming finite variances, each summand in the definition of $D(p)$ is nonnegative, and hence $D(p) \ge 0$. Here, outcomes y with $\mathbb{P}(Y(p) = y) = 0$ are omitted from the sum. □

Trade profile and gain from trade. Here we give an integral representation of the optimal expected gain from trade under (S, B), in terms of the trade-profile function $\alpha(p) = \mathbb{P}\left(p \in [S, B]\right)$. Denote the optimal expected gain from trade by $G = \mathbb{E}\left[(B-S)_+\right]$, where $(\cdot)_+$ denotes the positive part. The claim does not require independence and also holds when both sides are infinite.

Claim D.2 (Gain-from-trade decomposition)**.** *The following integral decomposition holds:*

$$G = \int_{\mathbb{R}} \alpha(p)\, dp.$$

Proof. For every realization (s, b) of (S, B), we can write the corresponding gain from trade as follows:

$$(B-S)^+ = \int_{\mathbb{R}} \mathbf{1}\{S \le p \le B\}\, dp,$$

since the integral is the length of the interval of prices at which trade is possible. Because the integrand is nonnegative, Tonelli's theorem gives

$$\mathbb{E}[(B-S)^+] = \mathbb{E}\left[\int_{\mathbb{R}} \mathbf{1}\{S \le p \le B\}\, dp\right] = \int_{\mathbb{R}} \mathbb{P}(S \le p \le B)\, dp = \int_{\mathbb{R}} \alpha(p)\, dp.$$ □

Information as displacement. For any event E of positive probability, we introduce two notions of "distance" to measure the information provided by the occurrence of E about S and B.

$$\delta_{Z|E}^2 = (\mathbb{E}[Z \mid E] - \mu_Z)^2 + (\mathrm{sd}(Z \mid E) - \sigma_Z)^2, \qquad \varrho_{Z|E}^2 = \mathbb{E}\left[(Z - \mu_Z)^2 \mid E\right].$$

The first quantity $\delta_{Z|E}^2$ measures the distance from the posterior point $(\mathbb{E}[Z \mid E], \mathrm{sd}(Z \mid E))$ to (μ_Z, σ_Z), while the second one $\varrho_{Z|E}^2$ is the conditional second moment around the unconditional mean μ_Z.

Claim D.3 (Information is displacement)**.** *For a square-integrable Z and a binary signal Y,*

$$\sum_{y\in\{0,1\}} \mathbb{P}(Y = y)\delta_{Z|\{Y=y\}}^2 = 2\sigma_Z \left(\sigma_Z - \sum_{y\in\{0,1\}} \mathbb{P}(Y = y)\,\mathrm{sd}(Z \mid Y = y)\right). \tag{34}$$

For either positive-probability branch E, we also have

$$|\mathbb{E}[Z \mid E] - \mu_Z| \le \delta_{Z|E}, \qquad |\varrho_{Z|E} - \sigma_Z| \le \delta_{Z|E}. \tag{35}$$

If $|\mathbb{E}[Z \mid E] - \mu_Z| \ge \eta > 0$, then $\varrho_{Z|E} \le (1 + \sigma_Z/\eta)\delta_{Z|E}$.

Proof. By the law of total variance,

$$\sum_{y\in\{0,1\}} \mathbb{P}(Y = y)\left(\big(\mathbb{E}[Z \mid Y = y] - \mu_Z\big)^2 + \mathrm{sd}(Z \mid Y = y)^2\right) = \sigma_Z^2.$$

Expanding the square in the definition of $\delta_{Z|\{Y=y\}}^2$ and using this identity gives

$$\begin{aligned}
&\sum_{y\in\{0,1\}} \mathbb{P}(Y = y)\delta_{Z|\{Y=y\}}^2 \\
&= \sum_{y\in\{0,1\}} \mathbb{P}(Y = y)\Big[\big(\mathbb{E}[Z \mid Y = y] - \mu_Z\big)^2 + \mathrm{sd}(Z \mid Y = y)^2 - 2\sigma_Z\,\mathrm{sd}(Z \mid Y = y) + \sigma_Z^2\Big] \\
&= 2\sigma_Z^2 - 2\sigma_Z \sum_{y\in\{0,1\}} \mathbb{P}(Y = y)\,\mathrm{sd}(Z \mid Y = y) \\
&= 2\sigma_Z \left(\sigma_Z - \sum_{y\in\{0,1\}} \mathbb{P}(Y = y)\,\mathrm{sd}(Z \mid Y = y)\right),
\end{aligned}$$

which proves (34).

Fix an event E of positive probability. The first inequality in (35) follows directly from

$$\delta_{Z|E}^2 = \big(\mathbb{E}[Z \mid E] - \mu_Z\big)^2 + \big(\mathrm{sd}(Z \mid E) - \sigma_Z\big)^2 \ge \big(\mathbb{E}[Z \mid E] - \mu_Z\big)^2.$$

To prove the second inequality, first note that

$$\begin{aligned}
\varrho_{Z|E}^2 &= \mathbb{E}\big[(Z - \mu_Z)^2 \mid E\big] \\
&= \mathbb{E}\big[(Z - \mathbb{E}[Z \mid E])^2 \mid E\big] + \big(\mathbb{E}[Z \mid E] - \mu_Z\big)^2 \\
&= \mathrm{sd}(Z \mid E)^2 + \big(\mathbb{E}[Z \mid E] - \mu_Z\big)^2.
\end{aligned}$$

Here the cross term vanishes because $\mathbb{E}[Z - \mathbb{E}[Z \mid E] \mid E] = 0$ and, in particular, $\varrho_{Z|E} \geq \mathrm{sd}(Z \mid E)$. Therefore,

$$\begin{aligned}\big(\varrho_{Z|E} - \sigma_Z\big)^2 &= \varrho_{Z|E}^2 + \sigma_Z^2 - 2\sigma_Z \varrho_{Z|E} \\ &\leq \varrho_{Z|E}^2 + \sigma_Z^2 - 2\sigma_Z \,\mathrm{sd}(Z \mid E) \\ &= \big(\mathbb{E}[Z \mid E] - \mu_Z\big)^2 + \big(\mathrm{sd}(Z \mid E) - \sigma_Z\big)^2 \\ &= \delta_{Z|E}^2.\end{aligned}$$

Taking square roots proves $|\varrho_{Z|E} - \sigma_Z| \leq \delta_{Z|E}$, and hence (35).

Finally, if $|\mathbb{E}[Z \mid E] - \mu_Z| \geq \eta > 0$, the first inequality in (35) gives $\delta_{Z|E} \geq \eta$. Combining this with the second inequality yields

$$\varrho_{Z|E} \leq \sigma_Z + \delta_{Z|E} \leq \frac{\sigma_Z}{\eta}\delta_{Z|E} + \delta_{Z|E} = \left(1 + \frac{\sigma_Z}{\eta}\right)\delta_{Z|E}. \qquad \square$$

We now proceed with the proof of Lemma 3.1.

Lemma 3.1 (Trade or information)**.** *For every pair (S, B) of square-integrable real random variables, there is a common price p such that*

$$L(p) \leq 2^{11} D(p). \tag{17}$$

The price can be chosen as a function of the joint distribution.

Proof. We write $\mu_Z = \mathbb{E}[Z]$ and $\sigma_Z = \mathrm{sd}(Z)$ for $Z \in \{S, B\}$ to denote the mean and standard deviation of valuation Z, respectively. As a first-order measure of the gain from trade achievable for a given pair (S, B), we use $m = \mu_B - \mu_S$, while we use $\sigma = \sigma_S + \sigma_B$ as a proxy for the spread of the two variables. We start by addressing some simple corner cases, and then move to the most challenging regime.

Deterministic values. If either σ_S or σ_B is zero, then one of the two variables is deterministic and posting that value yields the optimal gain from trade. If $\sigma_S = 0$, then posting μ_S is optimal and yields $L(\mu_S) = 0 \leq D(\mu_S)$ (by Claim D.1); similarly, posting μ_B when $\sigma_B = 0$ yields the desired inequality.

Well-separated case. This case is characterized by a large distance m between the expected values of S and B, relative to their standard deviations, i.e., $m \geq 3\sigma$, and a price satisfying the statement of the lemma is $\overline{p} = (\mu_S + \mu_B)/2$. Let $F = \{Y(\overline{p}) = 0\}$ be the event that trade fails at price $\overline{p}$. If $\overline{p}$ always results in a trade, then Claim D.1 yields the statement. We may therefore assume that $\mathbb{P}(F) > 0$. Conditioning on F, at least one of $S - \mu_S$ and $\mu_B - B$ exceeds $m/2$. Hence

$$N = \sqrt{(S - \mu_S)^2 + (B - \mu_B)^2} > m/2, \qquad (B - S)_+ \leq m + \sqrt{2}N \leq (2 + \sqrt{2})N.$$

Put $\nu = \sqrt{\mathbb{E}[N^2 \mid F]} \geq m/2 \geq 3\sigma/2$. Then

$$L(\overline{p}) \leq (2 + \sqrt{2})\mathbb{P}(F)\nu. \tag{36}$$

Claim D.3 applied to each agent under F yields

$$\mathbb{P}(F)\left[(\varrho_{S|F} - \sigma_S)^2 + (\varrho_{B|F} - \sigma_B)^2\right] \leq 2\sigma_S D_S(\overline{p}) + 2\sigma_B D_B(\overline{p}) \leq 2\sigma D(\overline{p}).$$

The vector $(\varrho_{S|F}, \varrho_{B|F})$ has norm ν, while (σ_S, σ_B) has norm at most σ. The quantity in brackets is therefore at least $(\nu - \sigma)^2 \geq \nu^2/9$. It follows that

$$D(\overline{p}) \geq \frac{\mathbb{P}(F)\nu^2}{18\sigma} \geq \frac{\mathbb{P}(F)\nu}{12}, \qquad L(\overline{p}) \leq 12(2+\sqrt{2})D(\overline{p}) \leq 41 D(\overline{p}).$$

Overlapping case. We are left with the last case, i.e., $m < 3\sigma$, which is arguably the most intricate: the seller's and buyer's expected values are relatively close, and we want to find a "good" price. We tackle this case by providing a uniform upper bound on the trade profile $\alpha(p)$, which yields, by Claim D.2, the desired result. We refer to Figure 4 for a visualization of this bound. Define the supremum of D on $\mathbb{R}$ by $\overline{D} = \sup_{p\in\mathbb{R}} D(p)$. Note that $0 \leq \overline{D} \leq \sigma$. We first bound the trade profile uniformly over all prices in terms of the auxiliary piecewise function H :

$$H(p) = \begin{cases} \dfrac{2\sigma_S \overline{D}}{(\mu_S - p)^2}, & p < \mu_S - \sigma_S/3, \\ \dfrac{225\overline{D}}{2\sigma_S}, & \mu_S - \sigma_S/3 \leq p \leq \mu_S + \sigma_S/3, \\ 1, & \mu_S + \sigma_S/3 < p < \mu_B - \sigma_B/3, \\ \dfrac{225\overline{D}}{2\sigma_B}, & p > \mu_S + \sigma_S/3, \\ & \mu_B - \sigma_B/3 \leq p \leq \mu_B + \sigma_B/3, \\ \dfrac{2\sigma_B \overline{D}}{(p - \mu_B)^2}, & p > \max\{\mu_S + \sigma_S/3, \mu_B + \sigma_B/3\}. \end{cases}$$

If the central intervals overlap, the seller's bound is used through $\mu_S + \sigma_S/3$ and the buyer's bound thereafter.

Claim D.4 (Uniform bound on the trade profile)**.** *For every $p \in \mathbb{R}$, we have $\alpha(p) \leq H(p)$.*

Proof. For any positive-probability branch E of $Y(p)$, Claim D.3 gives

$$\mathbb{P}(E)\delta^2_{Z|E} \leq 2\sigma_Z D_Z(p) \leq 2\sigma_Z \overline{D}, \qquad Z \in \{S, B\}.$$

We prove the seller's tail and central bounds first. If $\alpha(p) = 0$, both are immediate, so assume that $E = \{S \leq p \leq B\}$ has positive probability. For $p < \mu_S$, we have $\mathbb{E}[S \mid E] \leq p$, and hence

$$\alpha(p)(\mu_S - p)^2 \leq \alpha(p)\delta^2_{S|E} \leq 2\sigma_S \overline{D}.$$

This proves the upper bound in the first piece of H.

For $\mu_S - \sigma_S/3 \leq p \leq \mu_S + \sigma_S/3$, set $\eta = \sigma_S/3$, $\alpha = \alpha(p)$, and $\varrho = \varrho_{S|E}$. Sellers trading at p with $S \geq \mu_S - \eta$ are within η of their mean. Those with $S < \mu_S - \eta$ also trade at the lower price $\mu_S - \eta$, since $B \geq p \geq \mu_S - \eta$. Thus, with $E' = \{S \leq \mu_S - \eta \leq B\}$,

$$\alpha\varrho^2 \leq \eta^2\alpha + \mathbb{E}\left[(S - \mu_S)^2 \mathbf{1}_{E'}\right]. \tag{37}$$

If $\mathbb{P}(E') > 0$, then $\mathbb{E}[S \mid E'] \leq \mu_S - \eta$. By Claim D.3,

$$\mathbb{P}(E')\varrho^2_{S|E'} \leq 16\mathbb{P}(E')\delta^2_{S|E'} \leq 32\sigma_S\overline{D}.$$

The same bound on the last expectation in (37) is immediate if $\mathbb{P}(E') = 0$. On E itself, Claim D.3 also gives $\alpha(\sigma_S - \varrho)^2 \le 2\sigma_S\overline{D}$. Consequently,

$$\sigma_S\sqrt{\alpha} \le \sqrt{\eta^2\alpha + 32\sigma_S\overline{D}} + \sqrt{2\sigma_S\overline{D}} \le \frac{\sigma_S}{3}\sqrt{\alpha} + 5\sqrt{2\sigma_S\overline{D}}.$$

Rearranging and squaring yields $\alpha(p) \le 225\overline{D}/(2\sigma_S)$, proving the second piece.

Applying these two bounds to $(-B, -S, -p)$ gives the buyer's central bound for $\mu_B - \sigma_B/3 \le p \le \mu_B + \sigma_B/3$ and tail bound for $p > \mu_B + \sigma_B/3$. These prove the fourth and fifth pieces. The third piece is simply $\alpha(p) \le 1$. The five cases cover $\mathbb{R}$, regardless of the ordering of the two means. □

By Claim D.2, we can now integrate the uniform bound in Claim D.4 over all of $\mathbb{R}$. The seller's tail and central pieces have integrals

$$\int_{-\infty}^{\mu_S-\sigma_S/3} \frac{2\sigma_S\overline{D}}{(\mu_S - p)^2}\,dp = 6\overline{D}, \qquad \int_{\mu_S-\sigma_S/3}^{\mu_S+\sigma_S/3} \frac{225\overline{D}}{2\sigma_S}\,dp = 75\overline{D}.$$

The buyer's tail and central interval have the same respective integrals; restricting them to $p > \mu_S + \sigma_S/3$ can only decrease their contribution. The plateau has length $(\mu_B - \sigma_B/3 - \mu_S - \sigma_S/3)_+ = (m - \sigma/3)_+$. Therefore,

$$G = \int_{\mathbb{R}} \alpha(p)\,dp \le \int_{\mathbb{R}} H(p)\,dp \le 162\overline{D} + (m - \sigma/3)_+. \tag{38}$$

Since $m = \mathbb{E}[B - S] \le G$, subtracting $(G - \sigma/3)_+$ gives

$$\min\{G, \sigma/3\} \le 162\overline{D}. \tag{39}$$

Finally, $m < 3\sigma$, Cauchy–Schwarz, and the triangle inequality imply

$$G \le m_+ + \mathbb{E}[|(B - \mu_B) - (S - \mu_S)|] \le m_+ + \sigma < 4\sigma.$$

Thus $G \le 12\min\{G, \sigma/3\} \le 1944\overline{D}$. If $\overline{D} = 0$, every price has zero lost gain. Otherwise, because $1944 < 2^{11}$, choose any price with $2^{11}D(p) > 1944\overline{D}$. Then $L(p) \le G \le 2^{11}D(p)$. □

D.2 Proof of Lemma 3.3

Recall the quadratic variational expression for Φ, given in (6):

$$\Phi(\boldsymbol{m}, R, \boldsymbol{z}) = \operatorname{Tr} R + (\boldsymbol{z} - \boldsymbol{m})^\top R^{-1}(\boldsymbol{z} - \boldsymbol{m}).$$

For a random vector $U \in \mathbb{R}^n$ with mean $\boldsymbol{\mu}$ and covariance Σ, expanding around the mean and applying Lemma B.1 with $\boldsymbol{x} = 0$ gives, for every $\boldsymbol{m} \in \mathbb{R}^n$ and $R \succ 0$,

$$\mathbb{E}[\Phi(\boldsymbol{m}, R, U)] \ge 2\Psi(\Sigma),$$

with equality when $\boldsymbol{m} = \mathbb{E}[U]$ and $R = \sqrt{\operatorname{Cov}(U)} \succ 0$.

Lemma 3.3 (A valid lottery exists)**.** *For every $\boldsymbol{m} \in \mathbb{R}^{2d}$, every block-diagonal $R \succ 0$, every context $\boldsymbol{x}$ with $\|\boldsymbol{x}\|_2 \le 1$, and every $\gamma \in (0, 1]$, there is a finite lottery of plans satisfying*

$$\mathbb{E}\left[\ell(\boldsymbol{w}) + 2^{10}\Phi(\boldsymbol{m}^+, R^+, \boldsymbol{w})\right] \le 2^{10}\Phi(\boldsymbol{m}, R, \boldsymbol{w}) + \gamma \qquad \text{for every } \boldsymbol{w} \in K. \tag{21}$$

Each plan in the lottery posts the same price, drawn from a finite grid, to both agents. The lottery can be chosen as a function of the state and context.

Proof. Fix the current state $(\boldsymbol{m}, R)$, context $\boldsymbol{x}$, and slack $\gamma \in (0, 1]$, and write $\Phi_0(\boldsymbol{w}) = \Phi(\boldsymbol{m}, R, \boldsymbol{w})$. If $\boldsymbol{x} = 0$, there is nothing to prove as the gain from trade is zero at any price, so assume $\boldsymbol{x} \neq 0$. Let $\mathcal{A}$ be the set of plans consisting of a common price p_a and two successor states $(\boldsymbol{m}_a^{(y)}, R_a^{(y)})$ for $y \in \{0, 1\}$, with each $R_a^{(y)}$ positive definite and block diagonal.

We recall some useful notation. For $\boldsymbol{w} = (\boldsymbol{s}, \boldsymbol{b}) \in K$, the valuations S and B are given by $S = \boldsymbol{x}^\top \boldsymbol{s}$ and $B = \boldsymbol{x}^\top \boldsymbol{b}$. For the price p_a associated with the generic plan $a \in \mathcal{A}$, we denote by $Y_a(\boldsymbol{w})$ the indicator of whether S and B permit trade at p_a, and by $\ell_a(\boldsymbol{w})$ the corresponding loss in terms of the gain from trade:

$$Y_a(\boldsymbol{w}) = \mathbf{1}\{S \leq p_a \leq B\}, \qquad \ell_a(\boldsymbol{w}) = (B - S)_+(1 - Y_a(\boldsymbol{w})).$$

We denote by $\Phi_a^+(\boldsymbol{w})$ the value of the potential under the realized feedback $Y_a(\boldsymbol{w})$ for the generic plan $a \in \mathcal{A}$ with price p_a and successor states $(\boldsymbol{m}_a^{(y)}, R_a^{(y)})$ for $y \in \{0, 1\}$:

$$\Phi_a^+(\boldsymbol{w}) = \Phi(\boldsymbol{m}_a^{(Y_a(\boldsymbol{w}))}, R_a^{(Y_a(\boldsymbol{w}))}, \boldsymbol{w}).$$

Finally, we introduce the function f_a, which is central to our proof:

$$f_a(\boldsymbol{w}) = \ell_a(\boldsymbol{w}) + 2^{10}\Phi_a^+(\boldsymbol{w}) - 2^{10}\Phi_0(\boldsymbol{w}).$$

The function f_a is defined in such a way that inequality (21) in the statement is equivalent to the existence of a finite lottery Q satisfying $\mathbb{E}_{a\sim Q}[f_a(\boldsymbol{w})] \leq \gamma$ for every $\boldsymbol{w} \in K$. We first show that, for every hypothetical prior, the infimum of the expected payoff over plans is nonpositive, and then use minimax and discretization to obtain a single such lottery for all hidden pairs.

A plan for each prior. Fix any valuation distribution $\pi \in \mathcal{P}(K)$ over K, where $\mathcal{P}(K)$ is the set of all possible prior distributions of the hidden pair on K. We assume that S and B are induced by sampling W according to π and using the fixed context $\boldsymbol{x}$. For any such distribution, Lemma 3.1 supplies a common price p with $L(p) \leq 2^{11}D(p)$. For simplicity, let Y denote the bit $\mathbf{1}\{S \leq p \leq B\}$. Write $\pi_y = \mathcal{L}_\pi(W \mid Y = y)$ whenever $\mathbb{P}_\pi(Y = y) > 0$. Here, the latter is the conditional law of W given $Y = y$. Lemma 1.2, applied to both marginals, bounds $D(p)$ by the decrease in Ψ_{tr}, so

$$\mathbb{E}_\pi[\ell_p(W)] = L(p) \leq 2^{11}\left(\Psi_{\mathrm{tr}}(\pi) - \sum_{y\in\{0,1\}} \mathbb{P}_\pi(Y = y)\Psi_{\mathrm{tr}}(\pi_y)\right).$$

Here and below, sums over y omit zero-probability branches. For $j \in \{S, B\}$, let $\boldsymbol{\mu}_{y,j}$ and $\Sigma_{y,j}$ be the mean and covariance of W_j under π_y. For $\lambda > 0$, define a plan a_λ with common price p and successor states

$$\boldsymbol{m}_{a_\lambda}^{(y)} = (\boldsymbol{\mu}_{y,S}, \boldsymbol{\mu}_{y,B}), \qquad R_{a_\lambda}^{(y)} = \mathrm{diag}\Big((\Sigma_{y,S} + \lambda I_d)^{1/2}, (\Sigma_{y,B} + \lambda I_d)^{1/2}\Big).$$

Choose arbitrary admissible successor states on zero-probability branches. We add the perturbation λI_d to ensure that both $\Sigma_{y,S} + \lambda I_d$ and $\Sigma_{y,B} + \lambda I_d$ are positive definite, so the inverses of their square roots are well defined. Applying the quadratic variational identity (6) separately to the two marginals gives

$$\mathbb{E}_\pi[\Phi_0(W)] \geq 2\Psi_{\mathrm{tr}}(\pi), \qquad \lim_{\lambda\to 0^+} \mathbb{E}_\pi\left[\Phi_{a_\lambda}^+(W)\right] = 2\sum_{y\in\{0,1\}} \mathbb{P}_\pi(Y = y)\Psi_{\mathrm{tr}}(\pi_y).$$

To compute the limit, diagonalize each marginal covariance: an eigenvalue $\tau \geq 0$ contributes $\sqrt{\tau+\lambda}+\tau/\sqrt{\tau+\lambda}$ to the corresponding conditional expected successor potential, which tends to $2\sqrt{\tau}$ as $\lambda \to 0^+$, including when $\tau = 0$. Combining the preceding bounds yields

$$\inf_{a\in\mathcal{A}} \mathbb{E}_\pi\left[f_a(W)\right] \leq \lim_{\lambda\to 0^+} \mathbb{E}_\pi\left[f_{a_\lambda}(W)\right] \leq 2^{11}\Psi_{\mathrm{tr}}(\pi) - 2^{10}\mathbb{E}_\pi\left[\Phi_0(W)\right] \leq 0.$$

Since π was arbitrary, the "Bayesian value" V_{B} satisfies

$$V_{\mathrm{B}} = \sup_{\pi\in\mathcal{P}(K)} \inf_{a\in\mathcal{A}} \mathbb{E}_\pi\left[f_a(W)\right] \leq 0.$$

The following claim exchanges the order of choice and provides the regularity needed to discretize the resulting lottery. Its proof involves a nontrivial application of Sion's theorem and is deferred to the end of the section. We introduce the inf sup value:

$$V_{\mathrm{W}} = \inf_{Q} \sup_{\boldsymbol{w}\in K} \mathbb{E}_{a\sim Q}\left[f_a(\boldsymbol{w})\right].$$

Claim D.5. *Allowing arbitrary probability lotteries Q on $\mathcal{A}$, we have*

$$\sup_{\pi\in\mathcal{P}(K)} \inf_{a\in\mathcal{A}} \mathbb{E}_\pi\left[f_a(W)\right] = \inf_{Q} \sup_{\boldsymbol{w}\in K} \mathbb{E}_{a\sim Q}\left[f_a(\boldsymbol{w})\right].$$

Moreover, for every $\eta > 0$, there is a lottery Q with $\sup_{\boldsymbol{w}\in K} \mathbb{E}_{a\sim Q}\left[f_a(\boldsymbol{w})\right] < V_{\mathrm{W}} + \eta$ whose price distribution has bounded density and compact support, and whose successor states range over a compact set of positive-definite block-diagonal states.

A finite lottery on a finite grid. Apply Claim D.5 with $\eta = \gamma/3$. Since $V_{\mathrm{W}} \leq 0$, it gives a lottery whose payoff is uniformly less than $\gamma/3$, with the regularity stated in the claim. Round its common price to the nearest multiple of $1/n$, retaining both successor states. Compact price support ensures that only finitely many grid points occur. The feedback can change only if the original price lies within $1/n$ of S or B. The bounded density makes this event have probability $O(1/n)$ uniformly over $\boldsymbol{w} \in K$. The branch payoffs are uniformly bounded, so the rounding changes the expected payoff uniformly by $o(1)$.

Next replace the two successor states by representatives of a sufficiently fine finite net in their compact range. Uniform continuity of Φ on the product of that range and K allows the additional payoff error to be made uniformly as small as desired; choosing representatives from the range preserves positive definiteness and block diagonality. Take each of these two errors smaller than $\gamma/3$. The resulting lottery has finite support, satisfies (21), and uses common prices on a finite rational grid. □

Sion's theorem. Before proving Claim D.5, we recall Sion's minimax theorem in its general semicontinuous form and record the affine version used in its proof. Recall that K is defined as $\mathbb{B}_2^d \times \mathbb{B}_2^d \subset \mathbb{R}^{2d}$ and is endowed with the Euclidean metric, so that it is a nonempty compact metric space. Quasiconcavity means that every superlevel set is convex, and quasiconvexity means that every sublevel set is convex.

Theorem D.6 (Sion's minimax theorem [Sio58])**.** *Let M and N be nonempty convex subsets of real topological vector spaces, and suppose that at least one of M and N is compact. Let $F : M \times N \to \mathbb{R}$ satisfy the following conditions:*

1. *For every $n \in N$, the function $F(\cdot, n)$ is upper semicontinuous and quasiconcave on M.*
2. *For every $m \in M$, the function $F(m, \cdot)$ is lower semicontinuous and quasiconvex on N.*

Then

$$\sup_{m \in M} \inf_{n \in N} F(m, n) = \inf_{n \in N} \sup_{m \in M} F(m, n).$$

Corollary D.7 (Separately continuous affine payoffs)**.** *Let $C(K)$ be the space of continuous real-valued functions on K. Let $\mathcal{G} \subset C(K)$ be nonempty and convex, equipped with the uniform norm. Equip the space $\mathcal{P}(K)$ of Borel probability measures on K with the weak topology. Suppose that $F : \mathcal{G} \times \mathcal{P}(K) \to \mathbb{R}$ is separately continuous in these topologies and affine in each variable. Then*

$$\inf_{g \in \mathcal{G}} \sup_{\pi \in \mathcal{P}(K)} F(g, \pi) = \sup_{\pi \in \mathcal{P}(K)} \inf_{g \in \mathcal{G}} F(g, \pi).$$

We now turn to the proof of Claim D.5.

Proof of Claim D.5. We start by proving the "easy" direction of the equality, via weak duality. For any prior π and lottery Q, draw $W \sim \pi$ and $a \sim Q$ independently. Then

$$\inf_{a \in \mathcal{A}} \mathbb{E}_\pi \left[f_a(W) \right] \leq \mathbb{E}_{a \sim Q} \left[\mathbb{E}_{W \sim \pi} \left[f_a(W) \right] \right] = \mathbb{E}_{W \sim \pi} \left[\mathbb{E}_{a \sim Q} \left[f_a(W) \right] \right] \leq \sup_{\boldsymbol{w} \in K} \mathbb{E}_{a \sim Q} \left[f_a(\boldsymbol{w}) \right].$$

The first and last inequalities use that an average is at least the infimum and at most the supremum. The exchange of expectations follows from Tonelli's theorem after adding the constant $2^{10} \max_K \Phi_0$, which makes the integrand nonnegative. Taking the supremum over π and then the infimum over Q gives $V_{\mathrm{B}} \leq V_{\mathrm{W}}$. Both values are finite: $f_a \geq -2^{10} \max_K \Phi_0$, and retaining the current state gives a bounded payoff.

We now move to the other direction. Unfortunately, we cannot directly invoke Sion's theorem, as the function $f_a(\boldsymbol{w})$ has the discontinuities induced by $Y_a(\boldsymbol{w})$. To smooth it out, we introduce a random perturbation of the posted price. More precisely, we first average the discontinuous payoffs over a small random perturbation, apply minimax to the resulting continuous functions, and then let the perturbation vanish.

Fix any $0 < \varepsilon < 1/2$ and let U be sampled uniformly at random in $[-\varepsilon \|\boldsymbol{x}\|_2, \varepsilon \|\boldsymbol{x}\|_2]$ (recall that $\boldsymbol{x}$ is the nonzero context we are given). For any realization u of U, we define the perturbed vector $T_u \boldsymbol{w}$ as follows:

$$T_u \boldsymbol{w} = (1 - \varepsilon) \cdot \boldsymbol{w} + u \cdot \boldsymbol{v}, \quad \text{where } \boldsymbol{v} = (\boldsymbol{x}/\|\boldsymbol{x}\|_2^2, \boldsymbol{x}/\|\boldsymbol{x}\|_2^2).$$

By construction, the "perturbed" values are just the original S and B, scaled by $1 - \varepsilon$ and shifted by u:

$$\boldsymbol{x}^\top \left[(1 - \varepsilon) \cdot \boldsymbol{s} + u \cdot \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} \right] = (1 - \varepsilon) S + u, \quad \boldsymbol{x}^\top \left[(1 - \varepsilon) \cdot \boldsymbol{b} + u \cdot \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} \right] = (1 - \varepsilon) B + u.$$

Moreover, T_u maps K into itself, i.e., $T_u K \subseteq K$. To see this, we consider the first d components of $T_u(\boldsymbol{w})$; the other case is analogous:

$$\left\| (1 - \varepsilon) \cdot \boldsymbol{s} + u \cdot \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} \right\| \leq (1 - \varepsilon) \|\boldsymbol{s}\|_2 + \frac{|u|}{\|\boldsymbol{x}\|_2} \leq (1 - \varepsilon) + \varepsilon = 1.$$

For any base plan a, draw U and implement the common price and successor states

$$p_{a,U} = \frac{p_a - U}{1-\varepsilon}, \qquad \boldsymbol{m}_{a,U}^{(y)} = \frac{\boldsymbol{m}_a^{(y)} - U\boldsymbol{v}}{1-\varepsilon}, \qquad R_{a,U}^{(y)} = \frac{R_a^{(y)}}{1-\varepsilon}.$$

Denote this actual plan by a_U^ε. By construction, the original valuations S and B permit trade at the perturbed price $p_{a,u}$ if and only if the perturbed valuations permit trade at the original price, for any plan a and realization u of U:

$$\mathbf{1}\left\{S \le \frac{p_a - u}{1-\varepsilon} \le B\right\} = \mathbf{1}\{(1-\varepsilon)S + u \le p_a \le (1-\varepsilon)B + u\}.$$

The transformation T_u is used only in the analysis; the learner implements the displayed price and successor states using its sampled u. For either branch y, substitution into the quadratic potential gives

$$\Phi\left(\frac{\boldsymbol{m}_a^{(y)} - u\boldsymbol{v}}{1-\varepsilon}, \frac{R_a^{(y)}}{1-\varepsilon}, \boldsymbol{w}\right) = \frac{1}{1-\varepsilon}\Phi(\boldsymbol{m}_a^{(y)}, R_a^{(y)}, T_u\boldsymbol{w}).$$

Also, $((1-\varepsilon)B + u - (1-\varepsilon)S - u)_+ = (1-\varepsilon)(B-S)_+$, so the lost gain has the same scaling. Combining these identities and subtracting the unchanged current potential gives

$$g_a^\varepsilon(\boldsymbol{w}) = \mathbb{E}_U\left[f_{a_U^\varepsilon}(\boldsymbol{w})\right] = \frac{1}{1-\varepsilon}\mathbb{E}_U\left[f_a(T_U\boldsymbol{w})\right] + 2^{10}\mathbb{E}_U\left[\frac{\Phi_0(T_U\boldsymbol{w})}{1-\varepsilon} - \Phi_0(\boldsymbol{w})\right].$$

The crucial fact is that now, for each fixed a, the averaged payoff g_a^ε is continuous in $\boldsymbol{w}$. To see this, consider any sequence $\boldsymbol{w}_n \to \boldsymbol{w}$. Except at the two noise values $u = p_a - (1-\varepsilon)S$ and $u = p_a - (1-\varepsilon)B$, the feedback at $\boldsymbol{w}_n$ eventually agrees with that at $\boldsymbol{w}$, and the corresponding branch payoff converges. These exceptional values have probability zero under U. Both branch payoffs are bounded on the compact set of hidden pairs ($\boldsymbol{w} \in K$) and noise values ($|u| \le \varepsilon\|\boldsymbol{x}\|_2$), so dominated convergence applies, including when $S = B$.

Moreover, the absolute value of the difference between Φ_0 and its scaled version converges uniformly to 0:

$$\delta_\varepsilon = 2^{10} \sup_{\substack{\boldsymbol{w}\in K \\ |u|\le\varepsilon\|\boldsymbol{x}\|_2}} \left|\frac{\Phi_0(T_u\boldsymbol{w})}{1-\varepsilon} - \Phi_0(\boldsymbol{w})\right| \longrightarrow 0.$$

Indeed, $\|T_u\boldsymbol{w} - \boldsymbol{w}\|_2 \le 2\sqrt{2}\,\varepsilon$, and the current potential Φ_0 is uniformly continuous on K. In particular, the correction is controlled independently of the base plan a.

For a prior π, let π_ε be the law of $T_U W$ when $W \sim \pi$ and U is independent. Since $T_u K \subseteq K$, this is another prior in $\mathcal{P}(K)$. Averaging the payoff identity and using the definition of V_{B} yields

$$\inf_a \mathbb{E}_\pi\left[g_a^\varepsilon(W)\right] \le \frac{1}{1-\varepsilon}\inf_a \mathbb{E}_{W'\sim\pi_\varepsilon}\left[f_a(W')\right] + \delta_\varepsilon \le \frac{V_{\mathrm{B}}}{1-\varepsilon} + \delta_\varepsilon.$$

We are now ready to apply the version of Sion's theorem presented in Corollary D.7. To this end, we consider the space $C(K)$ of continuous functions from K to $\mathbb{R}$, equipped with the uniform norm; in particular, we focus on $\mathcal{G}_\varepsilon \subset C(K)$, which is defined as the convex hull of the functions g_a^ε. By definition, every $g \in \mathcal{G}_\varepsilon$ has the form

$$g = \sum_{i=1}^N \alpha_i g_{a_i}^\varepsilon, \qquad \alpha_i \ge 0, \qquad \sum_{i=1}^N \alpha_i = 1,$$

for finitely many base plans $a_1, \dots, a_N$. To implement g, first sample an index i with probabilities $(\alpha_i)_{i=1}^N$, then independently sample U as above and execute the plan $(a_i)_U^\varepsilon$. The resulting lottery Q_g on $\mathcal{A}$ satisfies, for every hidden pair,

$$\mathbb{E}_{a\sim Q_g}\left[f_a(\boldsymbol{w})\right] = \sum_{i=1}^{N} \alpha_i \mathbb{E}_U\left[f_{(a_i)_U^\varepsilon}(\boldsymbol{w})\right] = g(\boldsymbol{w}).$$

Here the mixture uses finitely many base plans, but Q_g itself need not have finite support because U has a continuous distribution. Such lotteries are allowed in the definition of V_{W}; finite support will be obtained in the subsequent discretization step. Thus minimizing over $g \in \mathcal{G}_\varepsilon$ restricts the admissible lotteries and can only increase the infimum defining V_{W}.

Define the payoff pairing

$$F(g,\pi) = \int_K g(\boldsymbol{w})\, d\pi(\boldsymbol{w}).$$

Recall that $\mathcal{G}_\varepsilon$ is equipped with the uniform norm $\|g\|_\infty = \max_{\boldsymbol{w}\in K} |g(\boldsymbol{w})|$, while $\mathcal{P}(K)$ is equipped with the topology of weak convergence of probability measures. Since K is compact, $\mathcal{P}(K)$ is compact in this topology; it is also convex. The set $\mathcal{G}_\varepsilon$ is convex by construction, and F is affine in each argument separately.

Explicitly, for $t \in [0,1]$, $g, h \in \mathcal{G}_\varepsilon$, and $\pi, \pi_1, \pi_2 \in \mathcal{P}(K)$, linearity of integration gives

$$F(tg + (1-t)h, \pi) = tF(g,\pi) + (1-t)F(h,\pi),$$
$$F(g, t\pi_1 + (1-t)\pi_2) = tF(g,\pi_1) + (1-t)F(g,\pi_2).$$

These are exactly the affine identities required in Corollary D.7. We next check separate continuity, meaning continuity in either variable with the other held fixed.

For a fixed prior π and any $g, h \in \mathcal{G}_\varepsilon$,

$$|F(g,\pi) - F(h,\pi)| \le \int_K |g-h|\, d\pi \le \|g-h\|_\infty.$$

For a fixed $g \in \mathcal{G}_\varepsilon$, weak convergence $\pi_n \to \pi$ gives

$$F(g,\pi_n) = \int_K g\, d\pi_n \longrightarrow \int_K g\, d\pi = F(g,\pi),$$

because g is continuous and bounded on K. This is precisely where continuity of the smoothed payoffs is used.

We can therefore apply Corollary D.7 with $\mathcal{G} = \mathcal{G}_\varepsilon$ and the pairing $F(g,\pi) = \int_K g\, d\pi$. The hidden-pair set K is a nonempty compact metric space, $\mathcal{G}_\varepsilon \subset C(K)$ is nonempty and convex with the uniform norm, and $\mathcal{P}(K)$ carries the weak topology. The two preceding calculations verify separate affinity and continuity. We obtain

$$\inf_{g\in\mathcal{G}_\varepsilon} \sup_{\pi\in\mathcal{P}(K)} F(g,\pi) = \sup_{\pi\in\mathcal{P}(K)} \inf_{g\in\mathcal{G}_\varepsilon} F(g,\pi).$$

For each fixed g, an average under any prior is at most $\max_K g$. Conversely, g attains this maximum at some $\boldsymbol{w}_g \in K$, and the point mass $\delta_{\boldsymbol{w}_g}$ attains the same value in the prior formulation. Hence

$$\sup_{\pi\in\mathcal{P}(K)} F(g,\pi) = F(g,\delta_{\boldsymbol{w}_g}) = \max_{\boldsymbol{w}\in K} g(\boldsymbol{w}).$$

For each fixed prior π, linearity gives, for every finite mixture,

$$F\left(\sum_{i=1}^{N}\alpha_i g^{\varepsilon}_{a_i},\pi\right)=\sum_{i=1}^{N}\alpha_i\mathbb{E}_\pi\left[g^{\varepsilon}_{a_i}(W)\right]\geq\inf_{a\in\mathcal{A}}\mathbb{E}_\pi\left[g^{\varepsilon}_{a}(W)\right].$$

Since each individual g^{ε}_a also belongs to the convex hull, the reverse inequality holds after taking the infimum over g. Thus

$$\inf_{g\in\mathcal{G}_\varepsilon}F(g,\pi)=\inf_{a\in\mathcal{A}}\mathbb{E}_\pi\left[g^{\varepsilon}_{a}(W)\right].$$

Combining the implementation of g, the point-mass identity, Sion's exchange, and the convex-hull identity, in that order, gives

$$\begin{aligned}V_{\mathrm{W}}&\leq\inf_{g\in\mathcal{G}_\varepsilon}\max_{\boldsymbol{w}\in K}g(\boldsymbol{w})\\&=\inf_{g\in\mathcal{G}_\varepsilon}\sup_{\pi\in\mathcal{P}(K)}F(g,\pi)\\&=\sup_{\pi\in\mathcal{P}(K)}\inf_{g\in\mathcal{G}_\varepsilon}F(g,\pi)\\&=\sup_{\pi\in\mathcal{P}(K)}\inf_{a\in\mathcal{A}}\mathbb{E}_\pi\left[g^{\varepsilon}_{a}(W)\right]\leq\frac{V_{\mathrm{B}}}{1-\varepsilon}+\delta_\varepsilon.\end{aligned}$$

The last inequality is the bound proved above for every prior π, whose right-hand side does not depend on π. This estimate holds for every $0<\varepsilon<1/2$. Since V_{B} is finite, letting $\varepsilon\to0^+$ gives $V_{\mathrm{W}}\leq V_{\mathrm{B}}$. Together with weak duality, this proves equality: $V_{\mathrm{W}}=V_{\mathrm{B}}$. The limit is taken only in this numerical bound, so no convergence of the implementing lotteries is needed.

For the additional assertion, choose ε small enough and then an element of $\mathcal{G}_\varepsilon$ whose maximum is less than $V_{\mathrm{B}}+\eta$; no attainment of the infimum is required. This element is implemented by a finite mixture of the smoothed plans a^{ε}_U. Each common price is uniform on a bounded interval of positive length, and each successor state is a continuous function of U on a compact interval. The finite mixture therefore has bounded price density, compact price support, and a compact range of positive-definite block-diagonal successor states.

More explicitly, every price interval has length $2\varepsilon\|\boldsymbol{x}\|_2/(1-\varepsilon)$, so the mixture's density is bounded by $(1-\varepsilon)/(2\varepsilon\|\boldsymbol{x}\|_2)$. The choice of $\varepsilon>0$ and the finite mixture is fixed before the subsequent grid approximation. □

Proof of Corollary D.7. Since K is compact, $\mathcal{P}(K)$ is a compact convex subset of the space of finite signed Borel measures equipped with the weak topology. Apply Theorem D.6 with $M=\mathcal{P}(K)$, $N=\mathcal{G}$, and the transposed payoff $\widetilde{F}(\pi,g)=F(g,\pi)$. For fixed g, continuity in π implies upper semicontinuity, and affinity implies concavity and hence quasiconcavity. For fixed π, continuity in g implies lower semicontinuity, and affinity implies convexity and hence quasiconvexity. Thus all the hypotheses hold, with compactness supplied by the prior space $\mathcal{P}(K)$, and the theorem gives the stated equality. □

E Lower Bounds

In this section, we present standard lower bounds for both contextual search and contextual bilateral trade. First, we recall the standard coordinate-wise lower-bound argument for contextual search [LPLS21].

Proposition E.1 (Lower bound for contextual search)**.** *For every $d \geq 1$, every $T \geq d$, and every possibly randomized algorithm, there exists a fixed target $\boldsymbol{z} \in \mathbb{R}^d$ with $\|\boldsymbol{z}\|_2 = 1$ and a deterministic sequence of context vectors of norm at most one such that*

$$\sum_{t=1}^{T} \mathbb{E}\left[\ell_t\right] \geq \sqrt{d}.$$

Proof. We consider an adversary that chooses the coordinates of $\boldsymbol{z}$ independently and uniformly from $\{-1/\sqrt{d}, 1/\sqrt{d}\}$, so that $||\boldsymbol{z}|| = 1$. For the first d rounds, the adversary lets $\boldsymbol{x}_t$ be the t-th basis vector, and sets $\boldsymbol{x}_t = 0$ afterward. At round $t \leq d$, $z_t = \boldsymbol{x}_t^\top \boldsymbol{z}$ is independent of any previous observation made by the algorithm, and thus of the posted price P_t. Therefore,

$$\mathbb{E}\left[\ell_t \mid P_t\right] = \frac{1}{2}\left|P_t - \frac{1}{\sqrt{d}}\right| + \frac{1}{2}\left|P_t + \frac{1}{\sqrt{d}}\right| \geq \frac{1}{\sqrt{d}}.$$

Summing the loss over $t \in \{1, \dots d\}$ provides the desired lower bound. □

We now turn to the matching lower bound for contextual bilateral trade. The coordinate construction in [CFG+26, Proposition 4.3], rescaled to the Euclidean unit ball, gives the following bound. Here, we allow the learner to post two distinct prices, as long as budget balance is enforced, i.e., the price p_t offered to the seller is at most the price q_t offered to the buyer.

Proposition E.2 (Lower bound for contextual bilateral trade)**.** *For every $d \geq 1$, every $T \geq d$, and every possibly randomized learner satisfying $p_t \leq q_t$, there are a fixed pair $(\boldsymbol{s}, \boldsymbol{b}) \in K$ and deterministic contexts of norm at most one for which*

$$\sum_{t=1}^{T} \mathbb{E}\left[\ell_t\right] \geq \frac{\sqrt{d}}{6}.$$

The same lower bound holds in the Bayesian setting for the prior whose coordinates are independent and for which each pair $(\sqrt{d}\,\boldsymbol{s}_j, \sqrt{d}\,\boldsymbol{b}_j)$ is uniform on $\{(0, 1/3), (2/3, 1)\}$.

Proof. This prior is supported on K, since $\|\boldsymbol{s}\|_2^2 \leq 4/9$ and $\|\boldsymbol{b}\|_2^2 \leq 1$. Use the t-th coordinate context on rounds $1 \leq t \leq d$, and zero contexts afterwards. Each round has gain $1/(3\sqrt{d})$, but its trading interval is equally likely to be $[0, 1/(3\sqrt{d})]$ or $[2/(3\sqrt{d}), 1/\sqrt{d}]$. Conditional on the learner's random seed and previous coordinates, the current interval remains independent and uniform. The nonempty interval $[p_t, q_t]$ can lie inside at most one of these two disjoint intervals, so trade occurs with probability at most $1/2$. Each of the first d rounds therefore loses at least $1/(6\sqrt{d})$ in expectation. Summing gives at least $\sqrt{d}/6$, and later losses are nonnegative. Averaging over this finite prior selects a fixed pair with at least the same expected loss, even when the learner is given the prior. □

Both marginal covariances of this prior are $I_d/(9d)$, so $\Psi_{\mathrm{tr}}(\pi) = 2\sqrt{d}/3$. Thus the Bayesian lower bound is $\Psi_{\mathrm{tr}}(\pi)/4$, and the dependence on the trace-square-root potential cannot be improved by more than a constant factor uniformly over priors. The dimension dependence in Theorems 3.2 and 3.4 is consequently optimal.

References

[ARS14] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Repeated contextual auctions with strategic buyers. In *NIPS*, pages 622–630, 2014.

[BCCF24] M. Bernasconi, M. Castiglioni, A. Celli, and F. Fusco. No-regret learning in bilateral trade via global budget balance. In *Proceedings of the ACM Symposium on Theory of Computing (STOC)*, 2024.

[BFL25] O. Besbes, Y. Fonseca, and I. Lobel. Contextual inverse optimization: Offline and online learning. *Operations Research*, 73(1):424–443, 2025.

[BPS17] A. Bärmann, S. Pokutta, and O. Schneider. Emulating the expert: Inverse optimization through online learning. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 70 of *Proceedings of Machine Learning Research*, pages 400–410, 2017.

[CBCC+21] N. Cesa-Bianchi, T. R. Cesari, R. Colomboni, F. Fusco, and S. Leonardi. A regret analysis of bilateral trade. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2021.

[CBCC+24] N. Cesa-Bianchi, T. R. Cesari, R. Colomboni, F. Fusco, and S. Leonardi. Bilateral trade: A regret minimization perspective. *Mathematics of Operations Research*, 49(1):171–203, 2024.

[CCC+24] Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. Bilateral trade: A regret minimization perspective. *Math. Oper. Res.*, 49(1):171–203, 2024.

[CFG+26] R. Cosson, F. Fusco, A. Gupta, S. Leonardi, R. Paes Leme, and M. Russo. Contextual online bilateral trade. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2026. arXiv:2602.12903.

[CLPL20] M. C. Cohen, I. Lobel, and R. Paes Leme. Feature-based dynamic pricing. *Management Science*, 66(11):4921–4943, 2020.

[GBC+25] S. Gaucher, M. Bernasconi, M. Castiglioni, A. Celli, and V. Perchet. Feature-based online bilateral trade. In *International Conference on Learning Representations (ICLR)*, 2025. arXiv:2405.18183.

[GDFS25] Simone Di Gregorio, Paul Dütting, Federico Fusco, and Chris Schwiegelshohn. Nearly tight regret bounds for profit maximization in bilateral trade. In *FOCS*, pages 1570–1594. IEEE, 2025.

[GGK+21a] S. Gollapudi, G. Guruganesh, K. Kollias, P. Manurangsi, R. Paes Leme, and J. Schneider. Contextual recommendations and low-regret cutting-plane algorithms. In *Advances in Neural Information Processing Systems*, 2021.

[GGK+21b] Sreenivas Gollapudi, Guru Guruganesh, Kostas Kollias, Pasin Manurangsi, Renato Leme, and Jon Schneider. Contextual recommendations and low-regret cutting-plane algorithms. *Advances in Neural Information Processing Systems*, 34:22498–22508, 2021.

[GGL+26] Anupam Gupta, Guru Guruganesh, Honghao Lin, Vahab Mirrokni, Renato Paes Leme, and David P Woodruff. Optimal and efficient online inverse optimization. *arXiv preprint arXiv:2610.08735*, 2026.

[Jav17] Adel Javanmard. Perishability of data: Dynamic pricing under varying-coefficient models. *J. Mach. Learn. Res.*, 18:53:1–53:31, 2017.

[KL03a] R. Kleinberg and T. Leighton. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *Proceedings of the IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 594–605, 2003.

[KL03b] Robert D. Kleinberg and Frank Thomson Leighton. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *FOCS*, pages 594–605. IEEE Computer Society, 2003.

[LPLS21] A. Liu, R. Paes Leme, and J. Schneider. Optimal contextual pricing and extensions. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1059–1078, 2021. arXiv:2003.01703.

[LPLV18] I. Lobel, R. Paes Leme, and A. Vladu. Multidimensional binary search for contextual decision-making. *Operations Research*, 66(5):1346–1361, 2018.

[MS83] R. B. Myerson and M. A. Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281, 1983.

[PLS18] R. Paes Leme and J. Schneider. Contextual search via intrinsic volumes. In *Proceedings of the IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 268–282, 2018.

[Sak26] S. Sakaue. Tight regret bound for online inverse linear optimization via multiscale matrix weights. arXiv preprint arXiv:2609.26978, 2026.

[Sho77] Naum Z Shor. Cut-off method with space extension in convex programming problems. *Cybernetics*, 13(1):94–96, 1977.

[Sio58] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958.

[STBO25] S. Sakaue, T. Tsuchiya, H. Bao, and T. Oki. Online inverse linear optimization: Efficient logarithmic-regret algorithm, robustness to suboptimality, and lower bound. In *Advances in Neural Information Processing Systems*, 2025.

[VB04] Lieven Vandenberghe and Stephen Boyd. *Convex optimization*, volume 1. Cambridge university press, Cambridge, 2004.

[YN76] David B Yudin and Arkadi S Nemirovskii. Informational complexity and efficient methods for the solution of convex extremal problems. *Matekon*, 13(2):22–45, 1976.